Solidarity as an Antecedent of Consensus Decision-Making: A Mixed-Mode Study

Submitted to Regent University

School of Business & Leadership

In partial fulfillment of the requirements

for the degree of

Doctor of Philosophy in Organizational Leadership

Susan Hares

December 2021

School of Business & Leadership

Regent University

This is to certify that the dissertation prepared by:

Susan Hares

titled

SOLIDARITY AS AN ANTECEDENT OF CONSENSUS DECISION-MAKING: A MIXED-MODE STUDY

Has been approved by her committee as satisfactory completion of the dissertation requirement for the degree of Doctor of Philosophy

Approved By:

Corné J. Bekker, D. Litt. et Phil., Chair
School of Business & Leadership

Doris Gomez, Ph.D., Committee Member
School of Business & Leadership

Emily Cabanda, Ph.D., Committee Member
School of Business & Leadership

December 2021

Abstract

Research by Koster and Sanders (2006) into alternatives to Organizational Citizen Behavior (OCB) found solidarity was a possible alternate concept to OCB. This longitudinal mixed-mode study examined whether solidarity and conflict were antecedents to effective team consensus decision-making in the Internet Engineering Steering Group (IESG), a virtual top management team (TMT) of the IETF. The IETF is a standards development organization (SDO) that has developed Internet technology standards since 1987. This research examined the IESG consensus decision-making in the 28 years of consensus decision-making from 1989 to 2016 from five different viewpoints (three different types of historical data and two surveys). Historiometric best practices guided the collection and IPA analysis of the historical data and the open-ended questions from the surveys. IPA analysis examined 3,458 group consensus decisions for group behaviors, with over 39.186 individual behavioral patterns consisting of 1 or more of the behaviors studied. The researcher surveyed all IESG members in 2013 and 2017. Every step of the five-strand methodology used triangulation (data triangulation, data collection, method triangulation in the analysis process [within-method and between-method], and theory triangulation). This web of triangulation used a majority strategy to resolve differences. This research concluded that solidarity was an antecedent to effective team consensus decision-making. Increases in solidarity behaviors in the TMT might increase the effectiveness of consensus decisions in team consensus decision-making. This research examined conflict based on the Jehn's (1999) Intragroup Conflict Scale, concluding that conflict (task and relationship) correlated to effective consensus decision-making. However, the exact nature between conflict and effective consensus decision-making was complex. Relationship conflict was minimized in the publicly available IETF historical records, slightly dampened in the survey, and expressed in-depth in a set of open-ended questions on conflict. Future researchers may continue to examine the interplay of solidarity and conflict in the IESG and other TMTs in other organizations.

Dedication

This work is dedicated to God and His glory, with thanks to God's guiding hand and his gift of my husband, David; my professors; my family; my encouraging friends; and my IETF community of professionals.

Soli Deo Gloria.

Acknowledgments

My deepest thanks go to the following professors at Regent University who have guided the creation of this dissertation: Dr. Corne Bekker, Dr. Doris Gomez, and Dr. Emily Cabanda. The faithful support of Dr. Bekker and Dr. Gomez through the entire process has enabled me to persevere in this lengthy Ph.D. research. In addition, Dr. J. Lyle Story's exegetical analysis of the Jerusalem Council in Acts 15 inspired me to investigate horizontal solidarity and vertical solidarity in leadership decisions involving conflict. Finally, Dr. Bruce Winston's guidance helped me with the initial exploratory investigation into the IESG.

My deepest gratitude goes to my husband; my daughters, Ruth and Deborah; my extended family; my Christian friends, Jill and Joyce; my research inter-raters, Eileen Aveni Sappington and Pam DeVries Grzech; and my church family at Saline United Methodist Church. For all the missed meals and delayed vacations, thank you. Your prayers and support kept me going through the long days and nights on this research.

I also want to thank the IETF chairs and IESG members who participated in the surveys and chatted with me about the results. Thank you for your help and patience as I have journeyed toward the results. The journey of research into solidarity in the IETF community is just beginning. I hope this research will pave the way for future IESG cohorts that effectively help the IETF continually push for a better Internet through open standards.

Table of Contents

List of Tables

List of Figures

Chapter 1 – Introduction

Volunteers staff and lead social organizations such as disaster relief organizations, religious organizations, political organizations, and open information and computer technology (ICT) standard development organizations (SDOs). In 1899, de Tocqueville (1899/1945) marveled that Americans formed voluntary "associations for the smallest undertakings" (Chapter 5, para. 4), but in the 21st-century, volunteer organizations impact the daily life of citizens in most countries of the world. Leaders of volunteer organizations use consensus decision-making at all levels to decide organizational activities and increase volunteer buy-in on decisions. These volunteer organizations impact the daily lives of citizens, but one may question if the leader's use of consensus decision-making to guide the discretionary efforts of volunteers can be linked to a volunteer organization having effective outcomes.

ICT SDOs create standards for information technology (IT) devices for businesses, consumers, schools, and research institutions. The effectiveness of the volunteer leadership of these SDOs directly impacts the quality and timeliness of the IT standards developed. Three SDOs that create standards for Internet technology are of the World Wide Web (W3C), the Institute of Electrical and Electronic Engineers (IEEE), and the Internet Engineering Task Force (IETF, 2021b). The W3C standardizes web technology, and the IEEE sets standards for Ethernet and Wi-Fi devices. The IETF (2021b) standardizes how digital information moves across many types of platforms connected to the Internet around the globe. Gençer (2012) points out that "most actors in Software and Internet technology" (p. 17) had embraced these open ICT standards for inclusion in products, and delays with these standards would result in delays in new ICT products. These delays in ICT standards may be caused by the volunteers, whose discretionary efforts within ICT organizations create these standards and lead the standards activities through consensus decisions. Simcoe (2007) suggests that the slowdown of IETF volunteer activities creating standards in 1992 to 2000 slowed down the development and delivery of new ICT products to ICT consumers. The IETF (2021a) makes all the decisions based on consensus decision-making, so the ineffective consensus decision-making delays the creation of IETF standards and ICT products. Understanding what constitutes good antecedents of effective consensus decision-making in the IETF can help the IETF and other ICT SDOs speed up standards creation for ICT products.

Volunteer organizations, such as churches, political organizations, and ICT standards groups, may have used consensus decision-making for decades. A leadership team for a volunteer organization changes over time as volunteer staff members leave or new members are added to handle new tasks. A volunteer's discretionary effort to create ICT standards or leading standards aligns with Organ’s (1997) definition of organizational citizenship behavior (OCB) but stresses one of OCB’s theoretically weak areas. Organ (1997) defines OCB as an individual’s “discretionary” efforts as efforts indirectly “or explicitly recognized by the formal reward system, that in aggregate, promote organizational goals” (p. 86).

A secondary concept that may explain volunteer efforts over time is solidarity. Hechter’s (1987) definition of solidarity suggests that an individual who contributes effort toward a group or a person has greater solidarity with the group (horizontal solidarity [HS]) or the leader (vertical solidarity [VS]). Suppose solidarity and OCB describe the discretionary behaviors of organizations. In that case, these behaviors should exist in different leadership teams within organizations, including the decision-making in the top management teams (TMT) of volunteer organizations. The first problem this research examined was the following: *How are OCB and solidarity related to effective consensus decision making as antecedents, and does solidarity predict effective consensus decision making better than OCB in the leadership of volunteer organizations?*

Yukl (2010) notes that the effective decision-making of strategic management teams (SMT) depends on the background and strengths of the leader and the members of the SMT, plus their ability to interact to make and implement decisions. Stewart, Manz, and Sims (1999) and Yukl (2010) suggest that a team’s effective collective decision-making requires that team members to do the following: share a commitment to the team’s tasks; have the appropriate level of task interaction for the team’s tasks; have a high level of mutual trust, cohesiveness, and cooperation; operate with accurate and shared mental models; and believe the team is capable of collective efficacy. If a team has interpersonal conflict, task conflict (TC), or a lack of interaction, the SMT's decision-making ability may be lessened due to these antecedent causes (Yukl, 2010). This current researcher also examined a second problem to refine the first problem: *How do task interaction levels, task-conflict, or relationship conflict (RC) moderate the ability of the solidarity or OCB antecedents to predict effective consensus decision-making in the leadership of volunteer organizations?*

Purpose of the Research

The purpose of this study was to advance understanding of antecedents of effective consensus decision-making in strategic management leadership teams. The researcher sought to determine if horizontal solidarity (HS), vertical solidarity (VS), and organizational citizenship behavior (OCB) were antecedents of effective consensus decision-making and if HS and VS predicted effective consensus decision making better than OCB. This researcher also determined if the task conflict (TC) or relationship conflict (RC) moderated these antecedents when controlling for task interdependence (TI). This research adds to the empirical leadership research reconsidering other empirically valid concepts for extra-role discretionary behaviors of OCB when roles within an organization are rapidly changing. The study was also part of a multiple-phase investigation undertaken by this author during graduate studies into the antecedents of effective consensus decision-making in the leadership of volunteer standards organizations, especially those in the ICT area.

Since 2000, organizations have evolved rapidly due to economic and technological changes due to the widespread adoption of Internet technology. Internet technology allows organizations to become virtual organizations, with portions of the organization and leaders located in geographically dispersed locations. This research examined whether solidarity was an improved concept for OCB in top leadership teams of volunteer organizations in an environment of rapid changes and virtual organizations that could predict effective consensus decision-making. Application of these improved OCB theories may aid research into consensus decision-making in public leadership of volunteer organizations. Improvements in the effectiveness of public leadership of volunteer ICT standards organizations may result in millions or billions of additional revenue in the ICT sector. Effective leadership in volunteer organizations helping the survivors of disasters such as the Red Cross may save money and lives.

Research Questions

The research questions included the following:

1. Are horizontal solidarity, vertical solidarity, and OCB direct antecedents of the consensus decision making in a top management team (TMT) in a volunteer organization such that an increase in horizontal solidarity, vertical solidarity, or OCB increases effective consensus decision-making?
2. Does horizontal solidarity and vertical solidarity predict effective consensus decision-making more accurately than OCB in a volunteer organization TMT?
3. Does interdependency of tasks, task conflict, or relationship conflict moderate the effect of horizontal solidarity, vertical solidarity, and OCB on effective consensus decision-making in a TMT in a volunteer organization?

Theory and Key Variables

Organ (1997); Van Dyne, Cummings, and Parks (1995); and Fields (2002) suggested that the view of OCB as extra-role behaviors (ERBs) versus in-role behaviors (IRBs) runs into difficulty when roles and contexts rapidly change. For example, the volunteer efforts in creating new open ICT standards in IETF (2021a, 2021b), IEEE, and W3C often involve rapidly evolving roles for individuals where ERBs may rapidly change to IRBs. Determining what antecedent behaviors slow down or speed up the efforts of volunteers to create open ICT standards requires determining the following: (a) what are good consensus decision-making behaviors in the volunteers, (b) what behaviors within volunteers cause good decision-making, and (c) what behaviors within volunteer leaders inspire volunteers to these behaviors? In addition, individuals participating in ICT standards often lead efforts to create new technology within their companies and circles of influence, so volunteer leaders within open ICT standards groups may have leadership experience within commercial organizations.

Fields (2002) noted that OCB also came from theories founded on civic virtue and the virtue of participating in organizational governance. Many people who participate in open ICT standards, such as IETF (2021b), IEEE, or W3C, state that their leadership combines ethical civic duty enhanced by virtuous participation in the open ICT standards, plus work-related efforts. Therefore, OCB behaviors may be commonplace in the leadership of these ICT SDOs instead of being an indicator of leaders who make effective decisions.

Koster and Sanders (2006) determined that HS with co-workers and vertical solidarity with supervisors better fit the OCB discretionary behaviors of extra-role or in-role. Koster and Sanders (2006) used solidarity to define cooperative behavior within a group, stating "solidarity involves at least two people who choose [either] to cooperate or not" (p. 253). HS involves two or more people within a group who cooperate by putting extra effort toward the cooperative goal. Vertical solidarity involves solidarity between a team leader and a team member (Sanders & Schyns, 2006a). Koster and Sanders (2006) used one-time survey research to examine seven companies and 674 employees to determine supervisor's and co-worker's solidarity. The researchers found that HS and VS behaviors were a better descriptor of OCB discretionary behaviors than the general OCB construct. This current researcher examined whether solidarity was a better predictor of effective consensus decision-making in volunteer organizations than OCB.

The research on consensus decision-making (Amason, 1996; Dess & Origer, 1987; Jehn, 1995; Koster & Sanders, 2006; Kotlyar, Karakowsky, & Ng, 2011; Sanders & Schyns, 2006a) shows that team decisions with strong consensus are quality decisions, which include the following: (a) agreed to by all team members, (b) perceived as fairly arrived at through deliberation of all pros and cons, (c) implemented by all members the decision requires, and (d) occurred without delays or cynicism. These researchers also found that consensus decision-making requires team cohesiveness, cooperation, and TI. Stewart et al. (1999) suggested that teams, with members who had successfully worked through the stages of forming, storming, and norming, had resolved TC and interpersonal RCs. This lack of conflict allows the team to cooperate cohesively on tasks. Koster and Sanders (2006) showed that organizational solidarity instruments provided a better definition of the cooperative interactions than OCB instruments because they presented both HS and VS measurements instead of OCB-generalized compliance and altruism.

Problems With Past Research

Organ (1997) and Podsakoff, Paine, and Bachrach (2000) questioned whether ERB OCB was distinct from in-role behavior (IRB). Solidarity is an alternative construct that may explain both extra-role and in-role OCB. The proof for solidarity as an alternative construct must show whether it explains group and organizational behaviors better than OCB for some circumstances. However, solidarity has fewer empirical research studies into the impact of solidarity on group behaviors in leadership teams. Koster and Sanders (2006) conducted an empirical study and examined how solidarity impacted group cohesiveness. However, Koster and Sanders did not examine how group cohesiveness related to collective efficacy (group efficacy). Koster and Sanders's research had limited external validity due to its method of a one-time survey of an employee's relationship with first-line supervisors in businesses with data gathered only from employees in seven enterprises. A longitudinal study of different teams in the same organization would provide better external validity than Koster and Sanders's research. Even though Yukl (2010) indicated that cohesiveness was one of the characteristics of an effective SMT, Sanders and Schyn's (2006) results did not show an antecedent relationship between HS and VS and effective decision making.

Researchers have examined the IETF (2021a, 2021b) as an ICT SDO when considering organizational processes. Gençer (2012), Simcoe (2007, 2012; Rysman

& Simcoe, 2010), Nickerson and Muehlen (2006), and Russell (2006) investigated the processes inside the IETF. Gençer (2012) noted that most studies on ICT examined how ICT standardized technology diffused into products, and few studied the process. Nickerson and Muehlen (2006), Russell (2006), Gençer (2012), and Simcoe (2007) conducted organizational process studies using the concepts of life-cycle and ecology to study the processes that formed the IETF SDO and the processes within the IETF that created published standards. Gençer (2012) examined how standardizing a unique technology required a different process than standardizing a linked technology, which refined an existing unique technology. Simcoe (2007) investigated the variance in the time each component of the IETF's standardized process took to create an IETF standard for technology. Simcoe (2007) examined the IETF process that began with creating an IETF working group (WG) to define the technology standards and ended with the final approval of the documents for the technology by the Internet Engineering Steering Group (IESG, 2020).

The IESG (2020) is the IETF’s TMT in charge of the final review of standards and submission for publication by the IETF’s publication series editor. Simcoe (2007) examined whether different technology complexity (task issues), working group TC and RC, and different IESGs (one TMT per year) created standards at different paces. Simcoe (2012) found that intellectual property rights (IPR) issues (regarding technology also caused delays in standards creation. In this research, Simcoe (2007) modeled the IESG review and approval as a constant delay per year without considering any variation in the IESG’s actions due to yearly changes in IESG’s membership or the impact of the IETF chair’s leadership of the IESG as a TMT. Simcoe (2007) assumed the IESG review and approval of standards had a constant delay based on technology type and complexity based on theory rather than empirical evidence. This current researcher must determine if IESG time for reviewing and approving the technology was constant or varied and whether leadership behaviors caused the variances. The following is restating Simcoe’s (2007) conclusion regarding IETF processes using concepts of consensus decision-making: The antecedents of timely decisions were low complexity of the technology, no IPR claims, and low conflict in the WG. IETF organizational governance standards, such as IETF (2016a, 2016b, 2016c) process documents, IETF organizational Request for Comments (RFCs), IESG statements, and liaison reports, were assigned low technological complexity and no IPR claims. Simcoe's (2007) assumptions on organizational documents are unproven and not supported by leadership theory.

This author's unpublished preliminary research in 2012 and 2013 suggested that HS and VS as antecedents of effective decision making might explain the effectiveness of IESG (2019, 2020) decisions rather than lower technological complexity or lower IPR rights. In 2012, as part of the Regent's Ph.D. on organizational research, this researcher took on a consulting project, conducting exploratory mixed-mode research into the effectiveness of the IESG team consensus decision-making when the IESG was under the leadership of three IETF chairs (denoted as IETF Chairs 4, 5, and 6 in this study). This exploratory research examined the minutes of five IESG biweekly meetings held during April to June in one of the years that each of these IETF chairs led the IESG. This exploratory interpretive phenomenological analysis (IPA) research examined the text of 5 IESG minutes from 2003 for IETF Chair 4, 5 IESG minutes from 2006 for IETF Chair 5, and 5 IESG minutes from 2011 for IETF Chair 6.

The effectiveness of the IESG (2019, 2020) consensus decision-making for an IESG was defined for an exploratory study by the total number of IESG decisions that created a measurable result. The IESG measurable result fell into the following categories: approval to publish documents as an IETF standard, WG actions, IETF management actions, or liaison actions. Based on this definition of effective IESG consensus decision-making, 2003 IESG and 2011 IESG TMTs were more effective than the 2006 IESG TMT for the IESG minutes studied. The researcher in this exploratory research posited that high percentages of collaborative behavioral interactions in the IESG decision-making would result in more effective consensus decision-making. The incidents of collaborative behavioral interactions found in the IESG formal minutes were higher in 2003 (87% of total) and 2011 (78% of total) versus 2006 (67%). This 2012 unpublished

IPA research also indicated that collaboration differed in the IESG TMTs in 2003, 2006, and 2010. Appendix D provides the consultant report on the exploratory research given to the IETF Chair 6 and a re-examination of the 2012 data based on this current research.

This researcher investigated in 2013 whether HS and VS were antecedents of the perception of effective consensus decision-making using survey research that queried present and past IESG members. Appendix Q.4 contains a write-up of this research. The 2013 survey found that HS and VS combined explained 28.1% (1989 to 2013) to 29.3% (1991 to 2013) of the variance in the effectiveness of consensus decision-making per IESG cohort.

Past IETF chairs and past IESG members reviewed the results of this researcher's unpublished research in 2012 to 2013 and suggested changes. These reviewers pointed out that interpersonal conflicts and substantial changes in the direction of IETF technology or organizational processes also impacted the effectiveness of IESG. The reviewer's detailed descriptions of the situations, where interpersonal conflict caused issues with the IESG's consensus decision-making process, suggested to the research that RC might moderate solidarity's relationship to effective decision-making. These IESG members and IETF chairs pointed out that the formal and narrative minutes might show these interpersonal conflicts and changes in direction. Based on the review comments on this researcher's unpublished research in 2012 and 2013, this researcher examined TC and RC in addition to solidarity. This researcher did not include the IPR related conflict (denoted by Simcoe's (2012) as "distributional conflicts"). Simcoe's (2012) antecedents of "time-period and technology class" (p. 316) were considered in the 2013 survey and this researcher by examining the results per IESG cohort year and the type of IETF RFC (standard or informational).

The size of the survey responses is a concern when adding a moderating impact to a research model. The 2013 survey of the IESG received 41 valid IESG responses and four valid IETF chair responses. The 41 responses in 2013 represent 54% of the active members in 2013 from the IESG cohorts from 1989 to 2013. Only six IETF chairs were active participants in 2013, so the four IETF chair responses represented 67% of potential responses from the IETF chairs. Even at 60% of the total IESG participants, the IESG survey had relatively few possible participants. Villa, Howell, Dorfman, and Daniel (2003) suggested that detecting moderator effects might require appropriate sample sizes to determine the existence of a moderator using multiple regression modeling techniques. Due to this recommendation from Villa et al. (2003) and the small number of potential survey responses, this researcher used two theoretical models (full model and reduced model). The full theoretical model had four independent variables (HS, VS, RC, and TC), four independent variables from the moderator effect (HSxRC, VSxRC, HSxTC, VSxTC), one control variable (TI), and one criterion (dependent) variable (effective consensus decision-making). The full theoretical required enough survey responses to support eight independent variables and one control variable. The reduced model had two independent variables (solidarity and conflict), one independent variable from the moderator effect (SxC), one control variable (TI), and one criterion (dependent) variable. The reduced theoretical model must have enough survey responses to support three independent variables. Appendix F provides additional details on the required sample sizes.

Proposed Consensus Decision-Making Model

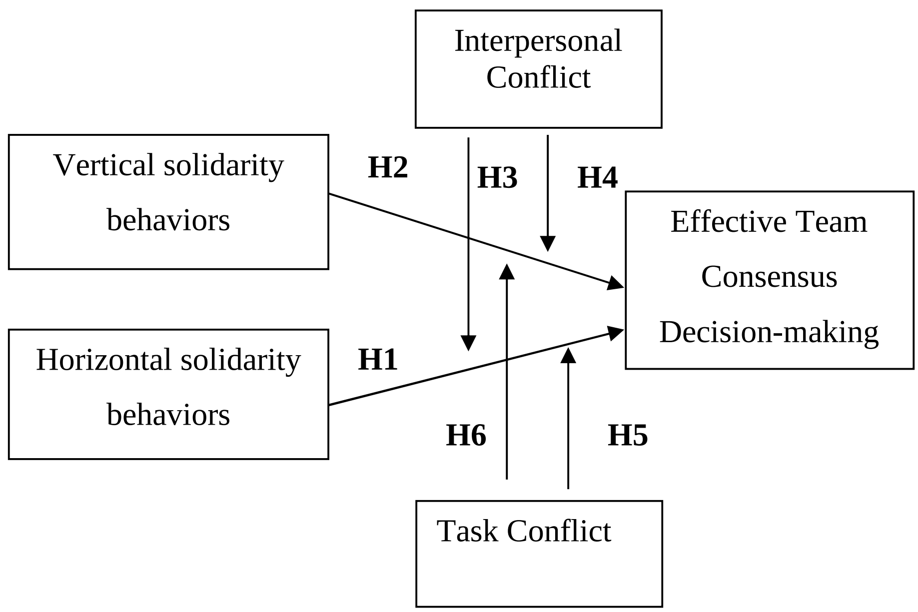
This researcher used a theoretical model for consensus decision-making model involving behaviors of solidarity, conflict, and TI shown in Figure 1. This theoretical model was used to posit that HS and VS were direct antecedents to effective team consensus decision-making. It was also used to suggest that task and relationship conflict (TC and RC) within the team moderated the effect of the solidarity antecedents on effective consensus decision-making within the team. Finally, the TI behavior moderated the impact of the antecedent HS behavior.

The key independent variables, dependent variables, and control variables for this research were group-level variables. The key independent variables for the full theoretical model included VS, HS, TC, and RC measured per yearly leadership team of the IESG (denoted as IESG cohort). These group-level independent variables were defined as (a) composites of individual interactions within the group in recorded historical artifacts or (b) composites of individual perceptions reported in the survey. The dependent variable was effective team consensus decision-making. The group-level effective decision-making was a composite variable comprised of the sum: the number of RFCs approved for publication within a year, the number of WG actions per year, and the number of IETF management actions. WG actions included scheduling pre-working meetings (denoted as Birds of a Feather meetings [BOFs]), creating new WGs, managing existing WGs, and closing WGs. The IETF management actions included administrative actions, liaison reports, IPR management, appeals processed, change to IETF processes, discussion of technical directions, technical advice for the Internet Assigned Numbers Authority (IANA, 2020), writing and publishing IESG statements, and writing “how-to” information on the IESG wiki. The control variables that influenced the hypotheses were TI and cohort year. Since an IESG cohort represented a unique TMT, evaluating the theoretical model focused on the IESG behaviors summarized per cohort year in the historical records and perceptions reported in the survey. The analysis of the historical record summarizes detected group behaviors as total counts of group-level behaviors per IESG cohort year. The analysis of the survey responses summarized the IESG member’s perception of the group behaviors as mean scores for the IESG cohort.

The solidarity behavioral construct was an improvement on the OCB construct. One way to determine that solidarity was a unique construct with discriminant validity was to replace the solidarity construct in the model with OCB constructs. The OCB constructs chosen for this research to replace VS and HS in the full theoretical model were OCB-generalized compliance (OCB-GC) and OCB-altruism (OCB-A). The reduced model was used to combine the OCB-GC and OCB-A scores to create a single OCB construct. Each of these OCB constructs was measured as follows: (a) individual behaviors in the historical records and (b) individual survey scores per cohort slot for an IESG member. The group-level OCB variables in the historical records were sums of the total individual counts per consensus decision, per meeting, or for all decisions made by an IESG cohort per year. The group-level OCB variables in the survey represented the perception of the

IESG as a group. This researcher defined the OCB group-level variables as the mean value of all individual IESG responses scores for each OCB variable (OCB-GC, OCB-A, OCB).

Figure 1: Full theoretical model for solidarity antecedents of effective decision-making.



Hypotheses

This model implied the following hypotheses:

H1: An increase in horizontal solidarity behaviors in a team will increase the effectiveness of the consensus decision made in team consensus decision-making.

H2: An increase in vertical solidarity behaviors will increase the effectiveness of the consensus decision made in team consensus decision-making.

H3: An increase in interpersonal conflict will moderate the positive effect of horizontal solidarity on effective consensus decision-making, such that an increase in interpersonal conflict will lessen the strength of the positive effect of horizontal solidarity on effective team consensus decision-making.

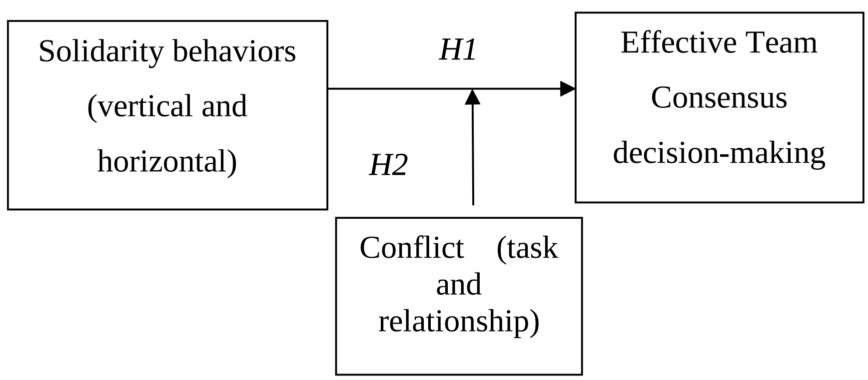
H4: An increase in interpersonal conflict will moderate the positive effect of vertical solidarity on effective consensus decision-making, such that an increase in interpersonal conflict will decrease the strength of the positive effect of vertical solidarity on effective team consensus decision-making.

H5: Controlling for task interdependence, an increase in task conflict will moderate the positive effect of horizontal solidarity on effective consensus decision-making, such that an increase in task conflict will lessen the strength of the positive effect of horizontal solidarity on effective team consensus decision-making.

H6: Controlling for task interdependence, an increase in interpersonal conflict will moderate the positive effect of vertical solidarity on effective consensus decision-making, such that an increase in interpersonal conflict will decrease the strength of the positive effect of vertical solidarity on effective team consensus decision-making.

Reduced Model Moderator testing

Figure 2: Reduced model for antecedents of consensus decision-making.



Reduced Model Hypotheses

This reduced model implied the following hypotheses:

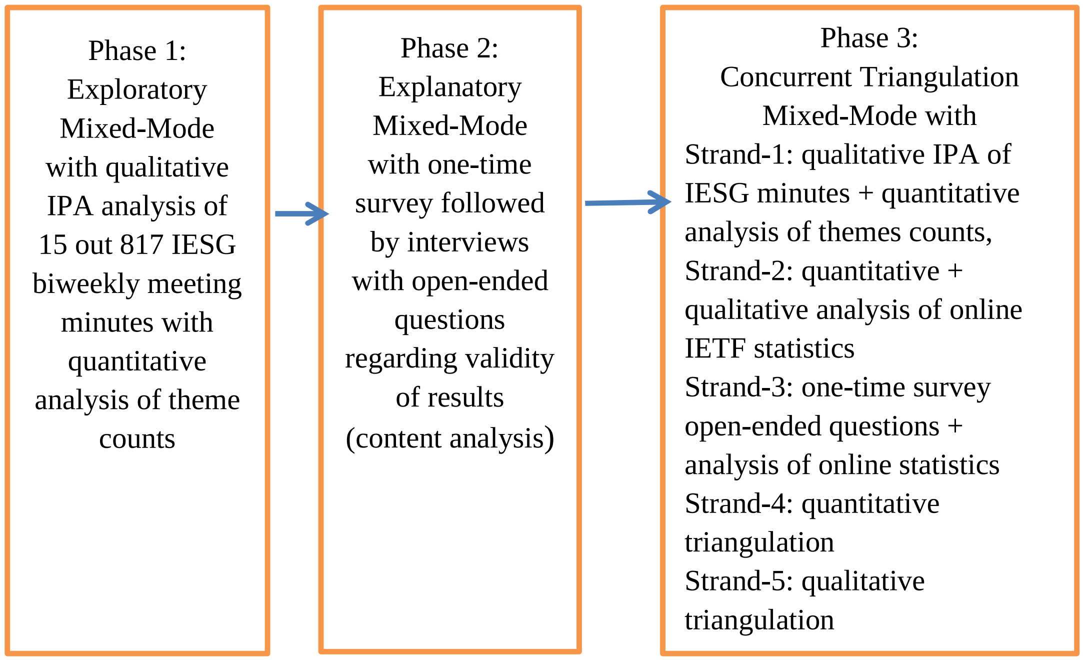
H1: An increase in solidarity behaviors in a team will increase the effectiveness of the consensus decision made in team consensus decision-making.

H2: Controlling for task interdependence, an increase in solidarity behaviors moderated by conflict will increase the effectiveness of the consensus decision made in team consensus decision-making.

Scope

Three things fascinated this researcher for years: leadership in volunteer organizations, consensus decision-making in volunteer organizations, and open ICT standards (https://open-stand.org/) in the IETF (2021a, 2021b), IEEE, and the W3C. This fascination led this author to weave several pieces of the research engaged in the doctoral program into a three-phase mixed-mode study investigating consensus decision-making in the IESG (2020), the IETF's (2020a, 2020b, 2020c) TMT. The IESG (2020) had existed as a TMT of the IETF (1991) since 1989. The membership of the IESG (2019) changes each year as volunteers join or leave the IESG. The three phases of mixed-mode research in this dissertation included (a) exploratory qualitative research with IPA and quantitative counts of themes, (b) explanatory mixed mode with survey research followed up by interviews, and (c) mixed-mode research using concurrent triangulation to expand earlier qualitative and quantitative research to a 28-year longitudinal study (1989 to 2016). This section describes the problems encountered in Phase 1 and Phase 2 of this research that helped define the scope for Phase 3 research.

Figure 3: Context in multiple phase research.



Scope of Phases 1 and 2

The first phase of the multiple phases was a qualitative exploratory study on whether the IETF chair's leadership of the IESG impacted the effective consensus decision-making of the IESG (Hares, 2012). Phase 1's exploratory mixed-mode study examined the formal minutes of five of the biweekly meetings of the IESG during the tenure of three different IETF leaders (IETF Chairs 4, 5, and 6) in 2003, 2006, and 2011 (see Appendix D for details). Each year, the IESG met biweekly, and the minutes of these meetings were posted online (http://www.ietf.org/iesg/minutes.html).

The qualitative researcher used IPA to analyze each meeting's minutes utilizing two encodings: IESG actions with results (encoded as results) and leader-group cooperative interactions (encoded as dyadic or multiple-person). A mixed-mode analysis was used on the theme count totals for the IESG actions, dyadic interactions, and multiple-person interactions for the IESG meetings analyzed per year. The IESG actions with results indicated effective consensus decision-making.

The exploratory examination showed fewer IESG actions with results in 2006 than in 2003 or 2011, plus a lower ratio of dyadic to multiple person interactions in 2006 than in 2003 or 2011. Fielder's (1964) least preferred co-worker (LPC) contingency theory explains that the situational change of leader-member relationships could cause this different group relationship in the IESG in 2006. The weakness of this research was the small sample size (5 out of ~20 per year) and the weak indication of results. The three chairs reviewed the results of the exploratory IPA analysis and provided feedback on the research. IETF Chair 5 (who led the IESG in 2005 to 2006) pointed out that the change in leader-group dynamics was due to the reorganization of areas in the IESG in 2005 to 2006 to create the Real-Time Application and Infrastructure (RAI) area. This effort increased task and RC. The Phase 3 research found similar task and RCs after reorganizing IETF areas during 2013 to 2015.

The Phase 2 explanatory mixed-mode research combined a "one-time" survey and post-survey interviews with participants. The survey instrument for Phase 2 included Koster and Sanders's (2006) survey instrument and the researcher's questions on IESG effective decision-making. Appendix Q provides a full copy of the survey in Section Q.4.9. Koster and Sanders's (2006) survey instrument contained questions on solidarity, OCB, and TI. This researcher's questions on IESG effective decision-making asked how effective an IESG cohort was to publish RFCs, create and manage WGs, and manage the IETF.

This researcher sent the 2013 survey to all past and current IESG available via email and received 46 valid survey responses (41 IESG member responses and five IETF chair responses). Each of the 46 responses contained data linked to multiple IESG cohorts. The analysis process normalized the survey response data to a data analysis format with data stored per IESG member per cohort year. This research denoted these data as IESG cohort slot data. The 41 valid IESG responses contain data for 129 IESG cohort slots. The five valid IETF chair responses had data for 14 valid IESG cohort slots. The survey instruments were reliable, with Cronbach alpha values of 0.849 to 0.908 for solidarity (S, HS, VS), TI, and the perceived results scores. The OCB-GC and OCB-A instruments were on the lower range of reliability, with Cronbach alpha values of 0.707 to 0.772, and OCB Cronbach alpha scores were unreliable (0.656). Appendix Q and Chapter 4 provide the full results of the reliability tests on the 2013 survey data.

The survey analysis found a correlation between a supportive solidarity atmosphere (in which the members of the IESG each perceived VS from the IETF chair and HS from the other members) and the perception of IETF (2019) effective consensus decision-making for all functions (RFC publication, WG management, and IETF [2020d] organizational management). The analysis of the survey responses hierarchical regression modeling (HRM) found that the solidarity explained 28% to 29% of the variance in the perceptions of effective consensus decision-making per IESG Cohort (1989 to 2013). Alternate models based on OCB did not explain any of these variances.

The Phase 2 research held post-survey interviews with eight individuals who served on the IESG (2019, 2020). These individuals reported the weakness of the Phase 2 study were (a) failure to link the perception of the IESG effectiveness to IETF statistics, (b) poor response in some years, (c) reports of confusion with the instrument, and (d) a lack of consideration of the impact of task or RCs. The Phase 2 study intended to link IESG perceptions to the actual yearly statistics of the IESG, but the errors in available IETF online statistics in 2013 caused this effort to be abandoned. The researcher’s feedback to the administrators of the IETF web caused the IETF’s yearly statistical data to improve. Each IESG had between 10 to 15 people. The IESG data for IESG cohorts in 1989, 1990, 1993, 1998, 2004, and 2005 have less than three IESG cohort responses raising concern about whether the cohort mean represents the IESG cohort’s perceptions. The survey data from the IESG cohorts in 1989 to 1993 had smaller responses due to the mortality issues due to IESG members aging out of the active workforce.

The postsurvey interviews indicated that the IESG (2019) cohorts in 1998, 2004, and 2005 had significant conflicts in the IESG, which might have caused a low survey response rate. The postsurvey interviews also indicated the following: (a) The survey form was confusing for IESG members who wanted to indicate a difference between years, (b) IESG relationship varied by year, and (c) the IESG working on substantial changes to the IETF reorganization caused TC or RC. The reviewers suggested that an in-depth review of the biweekly minutes would be the most accurate snapshot of the IESG meetings with the task and interpersonal relationships.

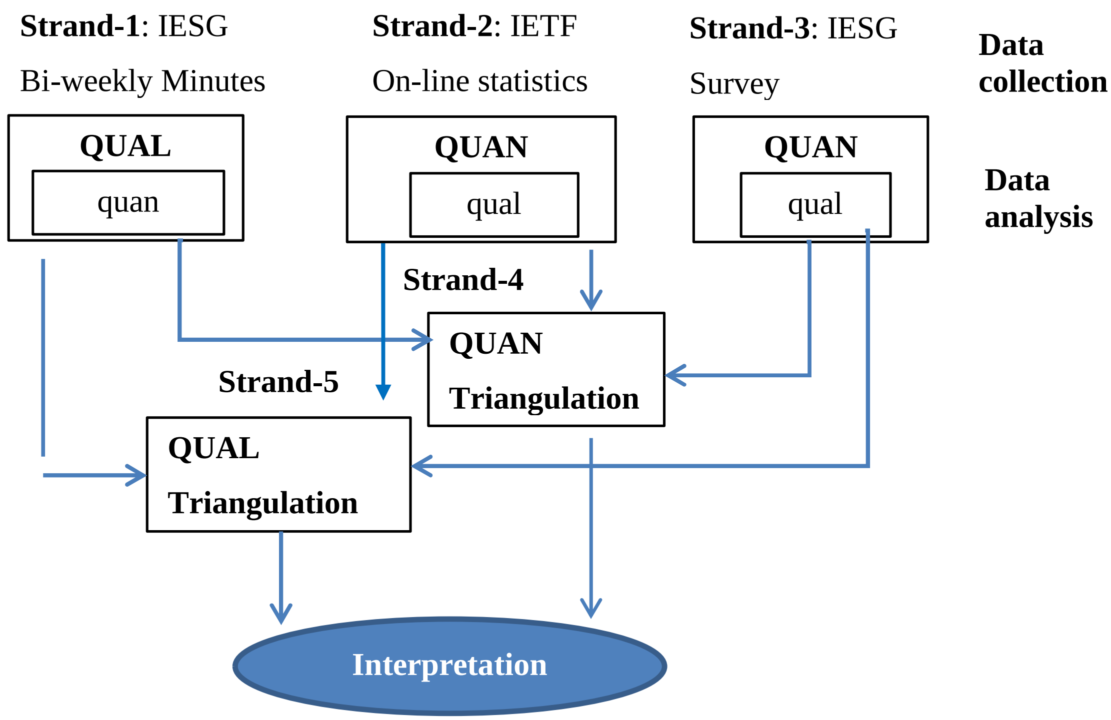
Scope of Phase 3

Phase 3 sought to resolve Phase 2 research issues with three adjustments to the Phase 2 methodology. The first adjustment added a complete qualitative IPA of the formal IESG minutes and the informal IESG (1992, 2020) minutes from 1991 to 2016 using the survey parameters to encode HS, VS, TC, RC, OCB (generalized compliance and altruism), and TI. This first adjustment added a historical data source to balance the view from the IESG survey. The second adjustment added the collection of IETF (1986, 1987, 1991, 1992, 2016a, 2016b, 2016c) statistics on the actual results of IESG consensus decisions from 1989 to 2016. This second adjustment enabled the Phase 3 analysis process to determine if the perceptions of each IESG cohort on the effectiveness of their consensus decision-making matched reality. The analytical tests on the theoretical models used the historical data on IESG (2000) and the perceived effectiveness of the IESG consensus decision-making from the survey. The third adjustment expanded the 2013 survey instrument to include Jehn's (1995) Intragroup Conflict Scale (ICS) and two open questions on conflict. Each of these actions was considered a "strand" of the mixed-mode analysis (see Strands-1–3 in Figure 4).

The feedback in the reviews of Phase 1 and Phase 2 indicated a lack of triangulation in data collection, data analysis, and interpretation of the results. This mixed-mode research triangulated the data collected from the biweekly minutes, online statistics, and survey per IESG cohort during the data collection phase. The theme counts from the Strand-1 mixed-mode IPA qualitative analysis of the biweekly meetings were analyzed using quantitative methods and triangulated with Strand-2 online statistics and Strand-3 survey quantitative data. Strand-4 provided the quantitative analysis that triangulated the data from Strands-1–3. After the quantitative analysis of Strand-1's IPA theme data, a qualitative analysis of the results was used to triangulate the two views of the biweekly IESG minutes data.

Strand-2's methodology also included qualitative analysis of the quantitative results. Strand-3 analysis of the survey data required an IPA qualitative analysis of the open-ended questions, a quantitative of all other survey questions, and a qualitative analysis of all results. Strand-5 qualitative analysis triangulated the qualitative data from Strands-1–3.

Figure 4: Mixed-mode design with triangulation.



Phase 3's research did not expand certain portions of Phase 2's study. Phase 3 did not develop the theoretical model to explain how task-interdependence interacted with solidarity or investigate an IESG member's management of their WGs. The Phase 2 survey showed that TI did correlate with HS and VS per cohort; thus, Strands-1–5 included it as a control variable. The Phase 3 research did not expand Simcoe's (2012) analysis of the IETF process for creating standards by considering how leadership worked in WGs or IETF areas. WG chairs lead the technical work in WGs. IETF area directors (ADs) coordinate the work of the WGs and WG chairs. Effective leadership by an AD could inspire WG chairs to prepare better quality standards, which would take less time to review at the IESG. An effective AD helps participants create new WGs in new technological areas, producing more documents to publish as RFCs. This researcher did not examine these important antecedents to RFC publication and WG actions.

Mixed-Mode Methodology (Phase 3)

This researcher used a longitudinal concurrent triangulation mixed-mode methodology to capture how 28 different TMTs had utilized consensus decision-making to create ICT standards in the IETF SDO. Each of these TMTs operated as the IESG for the IETF for 1 year (March to March). This researcher defined the group of people in the yearly TMT for the IESG as an IESG cohort and the year the IESG cohort operated in as the IESG cohort year. This concurrent triangulation study used the following five mixed-mode strands: Strand-1 – IESG biweekly minutes, Strand-2 – IETF online statistics, Strand-3 – IESG Re-survey, Strand-4 – Quantitative Triangulation of Strands-1–3, and Strand-5 – Qualitative Triangulation of Strands-1–3. The interpretation of the study arose from the triangulation analyses in Strand-4 and Strand-5. The triangulation analysis combined the multiple sources of data in the data collection and the mixed-mode analysis data analysis in Strands-1–5 shown in Figure 4. The researcher added triangulation to address concerns about the reliability and consistency of the results.

The original methodology of the research did not provide consistent results between Strand-1 and Strand-2. However, the researcher did not stop at the failure of the original methods but discovered a reliable alternate research methodology through cycles of data collection and data analysis with triangulation between Strands-1–3. This section provides an overview of the original methodology, the journey to the alternate methodology, and the alternate methodology. Each methodology description contains the theoretical assumptions, the methods for data collection, data analysis, and data triangulation checkpoints. The journey of discovery chronicles the cycles of data analysis, triangulation, and changes that led the researcher to revise the original theoretical assumptions, methods, and scope of the research.

Original Methodology for Analysis of Strands 1 to 3

The original methodology developed from scholarly research on survey instruments for solidarity, conflict, TI, and OCB; the exploratory IPA research done in Phase 1; and the survey research done in Phase 2 on the IESG. This combination of this research led the researcher to two assumptions. The first assumption was that the IESG minutes contained enough information to determine the group behaviors of the IESG cohorts during each act of consensus decision-making. A consensus decision-making action would either be a discussion or a measurable result. A single consensus decision might require several consensus decision-making actions to result in a decision. A consequence of this assumption was that the researcher expected the formal and narrative minutes to record the same IESG consensus decision-making actions with different levels of detail due to their purpose. The purpose of the formal minutes was to record official decisions. The goal of the narrative minutes was to record conversations in the meeting.

The second assumption was that the yearly online statistics would correlate with the results of the IESG cohort's consensus decision-making for that year. The original methodology contained two comparisons to test these assumptions during the analysis phase. The Strand-1 analysis method compared the themes counts for behaviors and consensus decision-making from formal and narrative minutes. The Strand-2 analysis compared the IESG decisions with results per year to the yearly statistics on the IETF (2020b, 2020c, 2020d) website. These early comparisons ensured that accurate Strand-2 data for IESG consensus decision-making results were used in Strand-3 and Strand-4. The descriptions of the data collections methods and data analysis below provide an overview of the original data collection methods and analysis.

Original data collection methods. Strand-1 and Strand-2 collected historical data from online sources for analysis. Strand-3's survey collected perceptions from IESG members on historical events. This section provides an overview of those collection processes.

Strand-1 collected and analyzed the minutes of the biweekly meetings of the IESG, which are historical records of the interactions of the IESG members available online at http://www.ietf.org/iesg/minutes.html. The online minutes existed for the period studied by the research (July 30, 1991 to March 16, 2017). These historical records existed in two forms: formal and narrative. The formal IESG minutes recorded all actions of the IESG within the meeting and between IESG meetings. The narrative IESG minute recorded actions IESG actions taken in the biweekly meeting and IESG conversations regarding the IESG actions in written and verbal forms. Strand-2 collects data from the IETF website (www.ietf.org) and the IETF's (2016c) online database on results of IESG decisions, such as RFCs published, WGs' actions, and IETF (2020m, 2020n, 2020o) management. Strand-3 expanded the survey instrument from the Phase 2 survey by including Jehn's (1999) ICS as reformat by Pearson, Easley, and Amason (2002) and two open-ended questions on IESG Conflict. The researcher sent this revised survey to all active IESG members and IETF chairs. This survey was used to solicit self-reported data per IESG term on the independent variables (HS and VS, RC, and TC), dependent variable (components of effective IESG consensus decision-making), and statistical control variables. In addition, this expanded survey included statistical control variables for cohort year, demographics (age, gender, and education), TI, and OCB generalized compliance and altruism. [1]

Strand-3's survey instrument had five parts: IESG term, IESG team behaviors, attitudes about the IETF job, IESG consensus decision-making, and demographics. In addition, the survey instrument used two online forms. The first format was for an IESG member, and the second format was for an IETF chair. Therefore, this section first describes the IESG member format of the survey instrument (see Appendix A for the IESG Survey instrument). Then, this section describes the alternate text used in the IETF chair format (see Appendix B for the IETF chair format).

The IESG term part of the survey queried the respondent for the IESG cohort(s) the person served on and their willingness to participate in the survey.

---

[1] Demographics were not included in the revised methodology (see chapter 4 for details).

The IESG team behavior part of the survey queried the respondent for the IESG contains the Lambooij et al. (2003) questions on horizontal and vertical organizational solidarity (OS). The final part of the IESG team was the instrument from Jehn's (1999) ICS in the six question format recommended by Pearson et al.'s (2002) analysis of Jehn's (1999) 9-item scale as "best capturing relationship and task conflict" (p. 110). These six items contained three questions on RC and three questions on TC.

The final portion of the IESG Team behavior contained two open-ended questions on intra-group conflict. The two open-ended questions included the following:

1. *Tasks conflict*: In some years, the IESG undertook significant tasks that caused conflict. If this happened during your term(s), could you describe the link between the task and the conflict?
2. *Relationship conflict*: In some years, some IESG members have reported more conflict in interpersonal relationships than the IESG member expected. If this happened during your term(s), could you describe how this occurred and how it impacted you or your work?

These two open-ended questions considered information beyond information asked in Jehn's (1999) ICS. The section on IETF attitudes about the IETF job included two instruments: MacKenzie, Podsakoff, and Fetter's (1991) OCB test for generalized compliance and altruism and Van der Vegt, Emans, Van De Vliert's (1998) test for TI. The form of solidarity, OCB, and TI instruments aligned with the format of these instruments used in Koster and Sanders's (2006) survey research. Because Koster and Sanders's (2006) survey research adapted this general text of these instruments per organization surveyed, this survey also adopts these changes to fit the IETF. The alignment aimed to maximize the ability to triangulate between this study and Koster and Sanders's (2006) research.

The IESG consensus decision-making part of the survey queried the IESG member whether their IESG cohort made effective decisions during a particular year in their tenure. This section asks whether an IESG cohort made effective decisions on standards (proposed or Internet) and information standards, WG and

BOFs, and IETF management items. The term BOF denotes the pre-WG formation meetings called “Birds of a Feather” meeting (Hoffman, 2012, section 1,Acronyms and Abbreviations) IETF management items included protocol registry decisions, meeting planning, and other administrative duties for the technology. This section also asks if the IETF chair positively impacts consensus decision-making. The demographic portion of the survey was optional, asking for age, gender, and education level.

The reliability of the instruments in the IESG team behavior section and the attitudes about the job section in published research had Cronbach alphas ranging from of 0.70 to 0.91 (see Table 54). The Phase 2 survey (in 2013) found similar reliability for the instruments in these sections of 0.70 to 0.88. The reliability tests on solidarity questions adapted for the IESG found Cronbach alpha between 0.85 to 0.88. The reliability tests on the TI instrument had a Cronbach alpha of 0.85, and the OCB instruments (OCB-GC and OCB-Altruism) had a Cronbach alpha of 0.70 to 0.76. Pearson et al. (2002) validated the 6-item version of the ICS scale in two surveys, which showed the reliability of these scales had a Cronbach alpha’s of between 0.73 to 0.91 for TC and 0.86 to 0.89 for RC. One of these surveys in this research by Pearson et al. (2002) was sent to 148 top managers of 48 firms and found the TC scale had a Cronbach alpha of 0.91 and 0.86 for RC. The Phase 2 survey showed that these instruments had face validity, content validity, construct validity, and predictive validity when used with IESG members. The Phase 2 survey also found that horizontal and VS constructs had concurrent validity and discriminant validity as different from OCB constructs of generalized compliance and altruism.

The survey format for the IETF chairs contained the same five sections, but the sections on IESG team behaviors and attitudes about the IETF job had unique wording for the IETF chair format. The IESG term of the survey was common to the format for IESG members and IETF chair formats. Logic with the IESG term section allowed the respondent to indicate the person’s role as an IESG member or an IETF chair. Based on the person’s role, the survey logic automatically directed the IESG member format or the IETF chair format for questions in the IESG team behaviors and attitudes about the IETF job sections. The recasting of the logic in the solidarity questions was used to address the difference in horizontal and solidarity for the IETF chair role. The OCB-GC Question 1 changed slightly to state the “IESG/IETF chair job description” instead of the “IESG description for the AD role.” The rest of the OCB questions had the same in both formats. The questions in the survey sections on IESG consensus decision-making and demographics sections were the same in the IESG Member format and the IETF chair format.

Strand-4 of this research collected the quantitative Strand-1, Strand-2, and Strand-3 for comparative analysis and triangulation of the historical and survey data. Strand-1’s source data had strong historical validity since each of the IESG minutes (formal and narrative) was approved as accurate within 2 to 6 weeks of creation. Strand-2’s data provided a second historical source for online statistics on the IESG consensus decision-making results per year. The yearly theme counts generated from the IESG minutes should correlate with the online statistics per year. Strand-1’s theme totals for IESG consensus decision-making results for all IPA analysis should also correlate with Strand-2’s historical results for all the IESG cohorts from 1991 to 2016. After comparing these two sets of historical data, Strand-4 was used to triangulate these data with the quantitative results from the survey to determine the best quantitative results using a majority triangulation method. A few sections describe the results from the quantitative for Strands-1–4 and then discusses the results in full.

Strand-5 was used to collect qualitative data specific to Strands-1–3 and the summary code memos written at the end of the analysis for Strands-1–4. The unique Strand-1 data collected for Strand-5 analysis included themes charts, weighted node diagrams, and code memos. The Strand-2 qualitative analysis created unique summary write-ups on the progression of technology developed in WGs per IETF areas and the strategic focus of each IETF Chair. Strand-3’s qualitative analysis of the open-ended conflict questions created summary code nodes on conflict for the IESG cohorts per IETF chair. Finally, Strand-5 was used to collect four data viewpoints of each IESG cohort to determine behavioral patterns and results common to all 25 TMTs and behavioral patterns unique to one or more IESG cohorts.

Original methodology for data analysis. The Strand-1, Strand-2, and Strand-3 analyses occurred concurrently. This subsection describes the mixed-mode analysis planned for these strands to provide quantitative and qualitative results. The final step of the quantitative analysis of Strand-1, Strand-2, and Strand-3 was to forward quantitative data to Strand-4's for comparative quantitative analysis. The last step in the qualitative analysis in these three strands was to write a summary note reviewing both quantitative and qualitative data analyzed and forward the summary note and the qualitative results to Strand-5. After reviewing the planned methodology, this subsection indicates why the original methods for data analysis in Strands-1–3 were alternate methodologies.

Strand-1's analysis began with an IPA on the formal and narrative IESG minutes (1991 to 2016 cohorts). The IPA analysis was used to encode themes based on survey questions during this IPA analysis, and the researcher analyzed the theme content via qualitative tools (theme grids and weighted node diagrams). The IPA analysis used the MAXQDA (2016; MAXQDA-12 through MAXQDA-2020) to record the theme encoding and tabulate total theme counts for IESG behaviors and actions of consensus decision-making. After tabulating the theme count totals, the researcher transferred the theme counts from MAXQDA (2016) to IBM SPSS for statistical analysis using quantitative tools (descriptive statistics, correlation, and HRM). After completing these qualitative and quantitative analyses, the researcher wrote a summary note after completing this analysis to summarize the results. The results from the qualitative analysis forwarded from Strand-1 included the theme grids, weighted theme diagrams, and the summary note.

Strand-2's original analysis methodology was used to analyze online data collected on IESG consensus decisions with results and demographics of the IESGs. The original Strand-2 quantitative analysis included descriptive statistics on statistics collected on IETF online sources on RFC published, WG-related actions, and IETF management actions per IESG cohort year (March to March). The WG-related actions included WG creation, WG closure, and pre-WG actions (BOFs proposed and BOFs held). This quantitative analysis assumed that the yearly statistics for IESG consensus decision-making results would directly correlate to the number of successful IESG consensus decisions, culminating in a measurable result. Strand-2's methods validated results per year by triangulating these results against Strand-1 theme counts for IESG decision-making results before using these data in Strand-3 analysis.

Strand-2's original quantitative analysis method was used to run descriptive statistics on the demographics of people attending the IESG meeting. The people who attended the IESG meetings included IESG members, liaisons, and service members who handled administrative functions. Therefore, the researcher assumed that these online statistics on planned attendees should correlate to the number of people recorded in the IESG minutes for all meetings per year. The Strand-2 qualitative analysis was used to examine the Strand-2 data to determine if the leadership of an IESG cohort or a group of IESG cohorts under an IETF chair successfully helped the IETF fulfill its mission. The IETF (2021b) stated its mission was "to make the Internet work better by producing high quality, relevant technical documents that influence the way the people design, use, and manage the internet" (para. 1).

The Strand-3 analysis was used to examine the quantitative data from the survey (Parts 1 to 5) and the qualitative data in the open-ended questions (Part 2). The researcher analyzed the quantitative data from the survey instrument using the scale reliability tests, descriptive statistics, correlation, and HRM tools from IBM's Statistical Package for the Social Sciences (SPSS). Scale reliability tests for the behavioral instruments within the survey used all survey respondents. Descriptive statistics were used to analyze the behavioral scales and the perceptions of all responses per cohort year.

Based on the survey response data's suitability for correlation and HRM modeling, the researcher analyzed for correlations between IESG group behaviors and between these group behaviors and the results (real and perceived). Lastly, the researcher conducted hierarchical multiple regression modeling tests based on the theoretical models. The qualitative data from the open-ended questions were analyzed using the IPA encoding codebook of the text devised in Strand-1 and mixed-mode descriptive analysis of the theme counts. The qualitative data from the survey responses were adjusted to preserve anonymity before IPA analysis. The researcher's Strand-3 IPA used the MAXQDA (2016) tools to record encoding and generate theme count totals.

Strand-4 and Strand-5 triangulated the analytical results from the historical sources (Strand-1 and Strand-2) and the survey (Strand-3 results). In addition, Strand-4 triangulated the data results gathered from the descriptive statistics, reliability tests, suitability tests for correlation and HRM, the correlation tests, and HRM model tests for the IESG minutes and the survey. The researcher sought the most reliable quantitative view of the IESG behaviors and results. Strand-5 was used to examine the qualitative data from a phenomenological viewpoint per year, across years led by a single IETF chair, and across all years to determine common and unique patterns of behaviors. The researcher sought to understand the difference between the historical record and the individual perception of the TMT.

Known challenges in original methodology. The researcher took on three challenges in the Phase 3 research: model expansion, creating encoding methods for an IPA analysis based on survey instruments, and obtaining enough responses for the 2017 survey. The theoretical data model used for this current research expanded the Phase 2 theoretical model in two ways. First, this model augmented the antecedents of VS and HS to include these two antecedents moderated by a composite conflict variable which includes TC and RC. Secondly, the researcher redefined the theoretical model's effective team consensus decision-making variable. Phase 2's theoretical model defined this variable as perceived results. Phase 3's theoretical model defined this variable as real or perceived results. The real results of IESG decisions per IESG cohort were summary counts created from data collected from online IETF statistics. However, this data had errors during 2010 to 2015, so Phase 3's methods validated these online statistics. This validation compared the real results generated from online summary counts against the real results generated from theme counts for IESG decisions with results from IPA analysis of IESG minutes. This validation occurred after 5% of the IESG minutes had been analyzed and after all IESG minutes had been analyzed.

The second challenge entailed creating an accurate codebook to govern consistent IPA encoding. Strand-1 uses mixed-mode theme coding using the survey questions as theme categories in the IPA analysis of formal and informational minutes. Using this method, the theme counts were input into quantitative statistical analysis for Strand-1 and Strand-4. The researcher addressed this challenge in the methodology by doing the IPA analysis with two other experts and comparing the encoding. The researcher and the two raters discussed any differences in the encodings until an acceptable interrater agreement on encodings was reached. After the researcher completed the IPA analysis, the same experts evaluated 10% of the IESG minutes encoded.

The third challenge was whether the data collected in Strand-1 and Strand-3 had a sufficient sample size for the statistical analysis methods planned. Statistical analysis required 10 to 15 samples per independent variable for simple statistical analysis and ~40 samples per independent variable for HRM tests to have good statistical conclusion validity. The full theoretical model had six hypotheses, where four of the hypothesis had two predictor variables (HS, VS), two moderators (TC, RC), and four control variables (cohort year, TI, OCB-GC, and OCB-A). The reduced model for moderator testing had two hypotheses, with one predictor variable (solidarity), one moderator (conflict), one criterion (effective consensus decision-making), and two controlling variables (TI and cohort year). If analysis methods calculated the results per IESG cohort and all IESG cohorts, then there were three control variables in the full model and one in the reduced model.

The challenge to the quantitative analysis was whether the IESG minutes contained enough decisions per IESG cohort year, and the survey had sufficient responses per IESG cohort year to support either the full or reduced model. The IESG formal minutes record average (and median) 23 meetings per year for IESG cohorts from 1991 to 2016 (July 7, 1991, to March 17, 2017), with 598 formal minutes for all IESG meetings for the 1991 to 2016 IESG cohorts. The IESG added narrative minutes to the formal minutes starting September 9, 2005, so IESG narrative minutes only existed for the IESG cohorts from 2005 to 2016 (246 minutes total minutes). Volunteers recorded narrative minutes for some IESG minutes for the IESG cohort for 2005 (12 meetings) and IESG cohort for 2006 (16 minutes) but most meetings for IESG cohorts (2007 to 2016). The narrative minutes recorded an average of 22 meetings per IESG cohort year during this period (2005 to 2016). If IESG made 30 decisions per meeting, then each year would have 690 decisions in the formal minutes and 660 decisions per IESG Cohorts. Appendix E has the full details on the number of formal and narrative IESG minutes per cohort year.

The Strand-3 survey had a small response poll. The total number of people in the 28 IESG cohorts (1989 to 2016) was 98 IESG members and seven IETF chairs. However, mortality issues by 2017 (death and aging out of active email participation) reduced this possible total population to 82 IESG members and five IETF chairs. Defining an IESG cohort slot as a survey response from one IESG member for one IESG cohort year, these 82 IESG members had a potential of 319 cohort slot responses, and the five IETF chairs had a possibility of 23 cohort slot responses. If the actual cohort response rate was 50% of the total responses per year, the total cohort slot responses was ~160 or four to seven cohort slot responses per year. At a 50% response rate for the survey, the IESG response only had enough cases for the reduced model to have enough cases for statistical validity for Strand-3 analysis. If the statistical analysis in Strand-3 could only use the reduced model, then Strand-4's quantitative triangulation only used the reduced model.

The Journey Toward the Alternate Model

Strand-1's original methodology for the IPA analysis followed the exploratory analysis of 5% of the IESG minutes with a validation of the minutes against the online statistics gathered. The journey toward an alternate model started when this validation failed. This section discusses the initial failure and investigations into the Strand-1 and Strand-2 methodology that led to revising the original assumptions.

To understand the validation failure, one must start the purpose of the exploratory analysis was to set the encoding codebook for the IPA analysis. Due to this focus, the researcher carefully chose a 5% data sample representing all IESG cohorts and variants of IESG minutes. The researcher analyzed the 5% sample and wrote the codebook. Two other researchers analyzed the 5% sample of the IESG minutes using the proposed codebook. Finally, the researcher and the two raters compared the details of the IPA analysis and reached agreements on changes to encodings. The researcher used mixed-mode analysis to create totals of themes encoded per IESG Cohort and across all IESG cohorts. These theme count totals included components of IESG decisions with results (RFC approved for publication, WG actions, and IETF management actions). The encoding of themes for each component of IESG decisions with results was one theme per decision. The researcher assumed the original methodology assumed the 5% sample would represent ~5% of the total online statistics per year (by category). The validation failed.

The researcher went through cycles of refining the original process to an alternate methodology. Strand-1 and Strand-2 examined the historical data; thus, the researcher used the guidelines for historiometric research for leadership suggested by Ligon, Harris, and Hunter (2012) to revise the mixed-mode methodology. Ligon et al. (2012) recommended that the mixed-mode analysis regarding historical data needed careful “sample plan formation,” “content-coding scheme development,” “material preparation,” and “code logistics” so that the quantitative analysis using “descriptive statistics and multivariate” (p. 20) would remain accurate. The benefit of historiometric methods was that the mixed-mode quantitative analysis could use descriptive and multivariate analysis.

The researcher first evaluated if the 5% sample plan was valid. The researcher selected 30 of the 598 formal IESG minutes (1991 to 2016) and 12 of the 246 narrative minutes (2005 to 2016) for the 5% sample. The selection criteria for the IESG minute was to pick the minutes from the IESG meeting closest to May 31st of each year and to add four additional formal IESG minutes from 1992 to 1996. The researcher chose this meeting because it was midway between the spring and summer IETF meetings. The researcher and the two other researchers believed that this sampling choice was valid, so the researchers re-evaluated the two assumptions behind the content-coding scheme, the material preparation, and coding logistics.

In this evaluation, the researcher asked the following:

1. Do the IESG minutes contain all the behavioral data per decision?
2. Do the formal and narrative minutes have the same number of decisions?
3. Do the details of the online statistics on RFCs published and WG actions correlate with IESG consensus decisions on RFCs and WG actions?

Each question required the researcher to compare the detailed accounts on the IETF site carefully with information in the IESG minutes. The researcher found that the answer to these questions was a resounding "no." The IESG formal and narrative minutes did not contain descriptions of all the group behaviors per IESG decisions, but the minutes were simply pointers to online data. The IESG formal minutes contained the IESG decisions within the meeting and between meetings. The IESG narrative minutes only had decisions within meetings. The researcher found that the IESG operated as a virtual TMT with continuous consensus decision-making via Internet communication and teleconferences. The IESG minutes contained a shorthand form of minutes that reference the IESG conversations in the meetings and online forums.

The third question challenged the researcher's original assumption that the yearly online statistics (adjusted for the cohort year) would correlate with the record of IESG decisions that culminated in measurable actions in the IESG minutes. The researcher investigated when the RFCs approved by IESG in the 5% sample were published and found post-approval delays before publication. These delays included post-approval editing by the authors and delays in the RFC generation and publication process. Due to these delays, the online RFC statistics per year did not correlate with the IESG actions.

The researcher also considered the WG actions in the 5% sample of the IESG minutes versus the online details for BOFs (proposed or held), WGs created, and WG closed. The researcher found that the IESG biweekly minutes included little information on the BOFs. The BOF approval meetings were not open to the public from 1991 to 2015, so minutes for these meetings did not exist in the public IETF repository. Records of proposed BOF exist in a BOF wiki (IETF, 2020b) for 2006 to 2016, and the IETF (2020d) proceedings for IETF meetings contained the list of BOFs held. The researcher concluded that Strand-2 methods needed to take the statistics for the IESG (2020) decisions (real and approved) from formal minutes (1991 to 2004) or a hand-merge of the formal and narrative minutes (2005 to 2016). The IESG minutes in the repository needed to be augmented by pseudo-minutes for three BOF approval calls.

The researcher reviewed these methodology changes with the two researchers who had done the 5% IPA analysis. The researchers suggested using the alternate methods on 10% of the IESG minutes. The 10% of the IESG minutes should consist of the original 5% sample, a second 5% sample, and 1 BOF of the 3 BOF calls per year. The research chose the second 5% sample as the IESG meeting closest to December 15.per year.

After the IPA analysis, the researcher and the two researchers reviewed the encodings, the time required for the analysis, and the results. During the IPA, the researcher found the descriptive statistics on the behaviors and the IESG decisions (actions and results) indicate the methodology had reliable results for the defined behaviors. However, the researcher found the IESG minutes for the IESG cohorts from 1991 to 2005 lacked the detail to confirm the IESG discovered variables. The researcher discussed these results during the review with the other two researchers and determined the researcher should do a second test of the methodology using a sample of all IESG minutes (formal, narrative, and BOF call) during 4 months of 2016 from April to July in 2016. After this third sample, the research and other two experts finalized Strand-1 and Strand-2 methodology and associated changes for Strands-3–5. The new methodology required a significant increase in manual processing for Strand-1 and Strand-2, so the researcher reduced the scope of the IPA analysis. The IPA Analysis examines 10% sample all IESG minutes and 100% sample for the IESG for the 1991 to 2016 IESG cohorts and 100% for the 2015 to

2016 sample. The 100% sample would confirm if the 10% sample had realistic results.

Alternate Methodology

The foundation for the alternate methodology arose based on two assumptions. The first assumption was that IESG operated as a virtual TMT with continuous consensus decision-making via Internet communication and teleconferences. The IESG minutes (formal and narrative) were pointers to the group discussions (online and in-meeting) regarding consensus decisions that occur within the period since the last biweekly IESG meeting. The second assumption was that the IETF online statistics presented the statistics of IETF as SDO rather than statistics that correlated to the IESG decisions per year. These two assumptions were different from those foundational to the original methodology, so the researcher revised the original methodology based on these new assumptions. This section describes the adjustments made to each strand to create this alternative methodology.

The researcher changed Strand-1 and Strand-2 data collection methods and data analysis methods in the alternate methodology. The data collection methods for this research for these two strands included collecting information, materials preparation, and coding logistics. The new Strand-1 method also gathered information about BOFs proposed and approved by the IESG during the 1991 to 2016 IESG cohorts. Before an IETF meeting, the IESG held a meeting to discuss and approve BOFs, denoted as BOF calls. Scribes did not record minutes for the BOF calls during the IESG cohorts from 1991 to 2015 but did record minutes for BOF calls for IESG cohorts 2016 to 2021. The alternate methodology included material preparation, where the researcher created pseudo-BOF call minutes for the 1991 to 2016 cohorts. Then, the researcher compared the pseudo-BOF call minutes generated in 2016 against BOF-call minutes during IPA analysis. The new Strand-2 method still gathered data on the technology progression of WGs within the IETF, but statistics collection occurred as part of the new material preparation cycle.

The new material preparation for the IESG minutes for Strand-1 and Strand-2 use involved three steps. The first step was to hand-merge formal and narrative minutes into a consistent record for any biweekly sessions having formal and narrative minutes during 2005 to 2016. The second step entailed collecting the online data pointed to by the hand-merged IESG minutes on IESG Behaviors (Strand-1) and IESG decisions and results (Strand-2). Finally, the third step entailed recording an online reference validating each IESG decision with a measurable result. These references included website links to publish RFCs, WG charters, IETF meetings. The IESG members' use of consensus decision-making between meetings meant that even the hand-merged IESG minutes did not correctly capture all IESG decisions or behaviors related to an IESG decision. This third step in the data collection process showed these missing decisions. The online material gathered was placed in code notes linked to the description of the IESG decision in the IESG minutes. The coding logistics placed theme encodings for all the behaviors and actions in the IESG minutes. Unfortunately, the data collection methods for the IPA analysis involved substantially more labor by the researcher. Therefore, to stay within 5 years for the dissertation research, the researcher reduced the sample size of the IESG minutes analyzed by IPA (Strand-1 and Strand-2). The sample size for this research was 10% of all IESG minutes from 1991 to 2016 and 100% of all minutes from 2016.

The quantitative data analysis methods for Strand-1 and Strand-2 changed slightly as the statistics for the components of results of IESG decision came from the IPA analysis of the IESG minutes. The mixed-mode analysis of the IESG minutes (formal, narrative, and BOF-call) generated theme counts for IESG behaviors studied, the IESG decisions, and the IESG decision with results. Due to the encoding methodology, Strand-2 theme counts were the statistics for the IESG decisions. Strand-3 used the theme counts for IESG results per year as the results. The quantitative analyses in Strands-1–4 used the same descriptive statistics and multivariate statistics (correlation and hierarchical regression model) proposed in the original. Strand-5 qualitative analysis changed to consider the IESG as a virtual TMT for IESG cohorts studied (1989 to 2016).

Limitations

Limitations in this concurrent research methodology involved selecting data to examine in the data model and from the methods within the research (data collection and data analysis). This section first considers the limitations related to the data model. The data model considered the TI of the IESG member as a control variable. The theoretical model's lack of consideration for the complexity of the TI model could have impacted the validity of the HRM modeling tests. The TI of the IESG member as a TMT member had a horizontal component with other IESG members and a vertical element with the IETF chair. The IESG member also had a vertical component with his constituency in the IETF Area the IESG member leads. This study also did not include this vertical component of TI, which Simcoe (2012) factored in how fast such WGs would produce those standards. Simcoe (2012) examined how the time it took to create quality technical specifications was impacted by the class of technology, type of RFC, time it took the WG to come to a consensus on standard, delays due to distribution conflict, and type of participants. The WG leader(s) and the IESG member responsible for these leaders might have influenced distribution conflicts within the WG. This current researcher ignored the complexity of the task interactions and interdependence by excluding this vertical component of the AD's task interaction on WG processes. Future researchers should consider this complexity.

Each strand of Strands 1-5 has limitations based on the data examined and the data collection and analysis methods. Table 1 lists these limitations for Strands-1–5 and the interpretation phase. The interpretation phase included asking IESG members and IETF chairs to review the results of this study. The qualitative analysis did not collect or consider IETF data on WG conflict (e.g., mail list reviews) that Simcoe's (2007, 2012) examined, so a direct comparison with Simcoe's (2007, 2012) results was impossible. Instead, the researcher used technology life-cycle concepts of Gençer's (2012) research and Simcoe's (2007, 2012) qualitative analysis in Strand-5.

Table 1: Limitations of Research Methodology Specific to the Strands or Interpretation

Strand	Model data not examined	Data collection methodology	Data analysis methodology
Strand-1 – IESG minutes	Strand-1 only examines the demographics of attendance.	The recording of narrative minutes only covers 2005 to 2016. Alternative Methodology: The IESG minutes (formal and narrative) did not contain all group behaviors regarding a decision. Instead, the IESG minutes referenced some group behavioral data by pointing to online data sources with these group behaviors. The alternate methodology collected data from online data sources and stored those data in inline code notes. IESG consensus decision-making was continuous, consisting of acts of decision-making in the biweekly meetings and between these biweekly meetings. The formal minutes contained all of the IESG acts and results of IESG consensus decision-making. The narrative minutes recorded only the IESG decision-making actions and results completed in a biweekly meeting. A hand-merge of data in and data pointed to by formal minutes and narrative minutes is necessary to describe the group behaviors per IESG decision per meeting. The alternate methodology was labor-intensive, so the scope IPA of the IESG minutes was reduced. The reduced scope of the IPA analysis was a 10% sample of the IESG minutes (formal and narrative) from 1991 to 2016 and a 100% sample for 2015 to 2016.	IPA analysis was limited by the role of language, suitability of the account, and accounts that explain rather than describe perceptions. Primary research did 90% of the analysis without interrater validation. The IESG formal and narrative minutes were analyzed separately (per meeting, per year, and for all years). Strand-1 IPA analysis encoded themes detected for individual behaviors in the group behaviors per decision. Alternate Methodology: Strand-1 IPA analysis encoded themes on data collected from minutes, online sources, and hand-merge.
Strand-2 – statistics	Strand-2 only looks at demographic statistics and consensus-decision-making statistics.	Strand-2 collected online statistics on results of IESG decisions based on an IESG Cohort year (March to March). Alternate methodology:	The research used descriptive statistics to create totals per IESG cohort year of components of the results of decisions (RFC published, WG actions, IETF management actions). Then, the researcher compared these

Strand	Model data not examined	Data collection methodology	Data analysis methodology
		The Strand-2 collected statistics on the IESG decisions (actions and results) in the formal IESG minutes augmented with code notes for online data. IESG decisions (actions and results) were encoded as theme counts. The Strand-1 IPA analysis generated the Strand-2 statistics per meeting, year, and annual IESG formal minutes.	totals with similar totals from the IPA analysis of the IESG minutes. Alternate methodology: The researcher ran descriptive statistics on the theme counts for IESG decisions (actions and results) per IESG cohort.
Strand-3 – survey	Strand three examines only the perception of the IESG members.	Mortality of IESG members meant some IESG were dead or no longer responding to IETF email. Maturation of the perceptions of how an IESG TMT interacted in a past IESG (1 to 25 years ago) caused the reality to be foggy in the IESG member's recollections. The researcher must transform the survey data after transferring it from the website hosting the survey (surveymonkey.com)	Strand-3's quantitative analysis might suffer from missing data or errors in data due to transformation. Due to the number of potential responses, this strand used the reduced model rather than the full model. Strand-3's IPA analysis of the open-end questions (part 6) explained rather than described perceptions. The primary research did 90% of the analysis without interrater validation.
Strand-4 – Quantitative triangulation	None. All variables considered.	The IPA analysis collected data on behaviors and IESG decisions (actions and results) as theme counts (per IESG cohort and all IESG cohorts). Survey data on perceived behaviors and perceived IESG effectiveness were scale scores (1-7) or mean scale scores (per IESG Cohort or all respondents). Historiometric leadership research used both scores regarding group behavior, but these scores were not directly comparable.	Data analysis using correlation and HRM modeling must have data suitable for these statistical methods. The data in Strand-1, Strand-2, and Strand-3 might have different levels of data suitability for correlation and HRM. If this occurred, a comparison between the Strand-1 and Strand-3 results might be limited to comparing results. Strand-3 used the reduced model, which limited the triangulation of Strand-3 data to the reduced model.

Strand	Model data not examined	Data collection methodology	Data analysis methodology
Strand-5 – Qualitative triangulation	None. All themes related and code-memos related to variables and discovered themes were analyzed.	During Strand-5 data collection, the researcher collected qualitative summary notes for Strands-1–4 and qualitative data specific to each strand. However, the researcher did not collect data on conflicts specific to WG decisions (actions or results).	Due to survey response size, the Strand-3 data analysis were restricted to using the reduced model. Therefore, Strand-4's quantitative data analysis were also restricted to the reduced model. Strand-5's qualitative analysis used this research's qualitative model and an alternate theoretical model on leadership during an organizational change to provide theoretical triangulation.
Interpretation	The interpretation did not consider complex TI or WG conflict.	Interpretation considered the theoretical model, triangulation in mixed-mode methods, advancement of knowledge on solidarity.	

Ethical Concerns in Data Collection

Strand-1 and Strand-2 analyzed the IESG minutes and other data on the IETF's public website. The data from these two strands were not subject to confidentiality. The survey in Strand-3 disclosed the intent of the survey work and indicated the material's confidentiality. Strand-4 and Strand-5's triangulation data built off results from Strand-1 and Strand-2 plus Strand-3 results, which anonymized data to retain the confidentiality of the survey. Any publication of the results of this research must take care to maintain the level of confidentiality Strand-3 requires.

The survey form used in Strand-3 disclosed the primary and second intent of the research, voluntary nature of participation, the confidential nature of data, and the "first-review" right of the participants in the introduction (see Appendices A and B for example text). The primary use was to provide general research insights on consensus decision-making in leadership teams in voluntary standards communities (e.g., the IETF). The secondary use of this research was to aid the IESG, IAB, and IETF nominations committee. The participant volunteered to participate by checking the survey's box in part-1. This permission could have been removed by either sending email or accessing the survey form and filling in a key. During result review sessions, any participant had the right of the first review. The original methodology planned these sessions between November 28th, 2016, and December 8th, 2016. The alternate methodology held these sessions after the complete text of the dissertation is completed (February 2022).

Timeline and Budget

The original timeframe for the data collection and analysis was December 1, 2016 to January 31, 2017. The original timeframe for the interpretation was 2 weeks. The long interpretation phase allowed for a staged review of the data by IESG members who were respondents, then the current IETF TMTs (IESG and IAB), and then a broader audience in the IETF. The extensive rework of the mixed-mode methodology and the time-intensive work caused this research to take 5 years (December 2016 to December 2021). The total budget for this research was $5,820, but only the $5,500 cost to fund inter-raters was expended by the primary researcher.

Table 2: Schedule and Budget for the Project

	Original plan		Alternate Plan				
Strand	Data collection	Analysis	Data Collection	Analysis	Personnel	Equipment	Budget
Strand-1 – minutes	12/1/2016 to 4/20/2017 *[2]	12/1/2016 to 6/30/2017	12/22/2016 to 1/15/2019	1/1/2017 To 3/15/2019	Primary research + 2 additional coders	MAXQDA-11	$245 *[2] $5000 inter-raters
Strand-2 – statistics	2/1 to 2/28	1/1/2017 to 6/30/2017	1/1/17 to 1/15/2019	1/1/2017 to 6/30/2019	Primary researcher	SPSS	$105*[2] Already acquired
Strand-3 – survey	4/3/2017 to 6/3/2017	6/4 to 6/6/2017	3/25/2017 5o 6/30/2017	7/1 to 7/30/2017	Primary researcher + 2 additional coders	Survey-monkey site, SPSS, email, MAXQDA-11,	$85/month for survey-monkey*[2], $15/mail*[2], $500 raters
Strand-4 – Quantitative triangulation	7/1	7/1 to 7/10	7/1/2019 to 2/15/2021	7/1/2019 to 2/15/2021	Primary researcher	SPSS, MAXQDA-11	Already required
Strand-5 – Qualitative triangulation	7/1	7/1 to 7/10	2/15/2021 to 11/3/2021	2/15/2021 to 11/3/2021	Primary researcher	SPSS, MAXQDA-11	Already required
Interpretation	-	2/1/2017 to 2/10/2017	-	11/3/2021 To 12/3/2021	Primary researcher	SPSS, MAXQDA-11	Already acquired
Totals	12/1/2016 to 1/12/2017	12/1/2016 to 2/10/2017			Primary research + 2 additional IPA coders	Email, web-site, survey-monkey site, SPSS, MAXQDA-11	$5820

*1 – Final 2016 IESG minutes were not approved and loaded on website until 4/20/2017.
*2 - Acquired by the primary research before the study.

Significance of Work

The Internet weaves together society's social and business threads within countries and globally with the W3C, email, voice calls, and instant communication. The IETF standards create the interoperable technology standards that allow a myriad of devices to pull onto the Internet and carry data within a house, a business, or worldwide. Without timely and relevant standards, the pace of innovations in Internet technology will slow, causing ICT businesses to spend more to deploy their products or lose money. Simcoe (2012) estimated that the IETF standardization slow-down losses were in the millions of dollars in 1995 to 2000. The volunteer management of creating IETF technology standards by the IESG TMT can either enable the timely development of technically excellent standards or delay these technical standards. Understanding the antecedents of effective consensus decision-making in IESG and the moderating effects of TC and RC can help reduce the risk of IESG delaying these standards. Therefore, the significance of this research to the IETF and the ICT industry is high.

Expansion of this research to other volunteer organizations using consensus decision-making such as religious or social aid organizations is potentially significant. For example, Christian churches theologically support consensus decision-making (Acts 15) and use a church TMT (parish council, elders, or administrative council) to make decisions. Church TMTs may decide about current and future social projects, budgets, and evangelism. Mainline Protestant churches face tremendous changes as their membership declines. These local religious TMTs must make tough decisions that the local church must enact. Understanding the antecedents of effective consensus decision-making in these churches and the moderating effects of TC and RC will help the volunteer TMTs within churches make effective decisions regarding tough subjects. These two examples demonstrate the broad applicability of this fundamental research that advances understanding of the antecedents of consensus decision-making and the moderating effect of conflict on those antecedents.

Summary of Chapter 1

This introductory chapter brought the reader through the personal journey of the primary researcher who started wondering about the antecedents of effective consensus-decision making in the leadership of volunteer organizations and ended up working on a multiple phase research project. This chapter began by examining the theories and research on OCB and discretionary behavior (Organ, 1997) and solidarity (Hechter, 1987; Koster & Sanders, 2006), which suggested that OCB and solidarity behavior might be antecedents for effective consensus decision-making in teams. This researcher found that increases in OCB and solidarity antecedents (horizontal and vertical) increased effective decision making, moderated by TI and conflict (task and relationship). In addition, Koster and Sanders (2006) found that increases in solidarity predicted effective consensus decision-making in teams better than OCB generalized compliance and OCB altruism. Solidarity was a possible improvement for OCB; thus, Koster and Sanders (2006) tracked OCB (generalized compliance and altruism) as a control variable. This previous research led the current researcher to question the following:

1. Does the increase in the strength of the antecedents (OCB, solidarity) increase the effective consensus decision making?
2. Does the solidarity better predict effective consensus decision-making than OCB?
3. Does the interdependency of task actions and conflict moderate the effect of antecedents of consensus decision-making?

This research is a three phase mixed-mode research seeking to answer these questions by examining consensus decision making within the IESG, one of the TMTs of the IETF (2021a, 2021b). The IETF is an SDO staffed by volunteers creating ICT standards for the Internet since 1986. The IETF's policy dictated that the IESG should make all decisions by consensus decision-making. The IETF measures the effectiveness of the IESG by its ability to publish standards, create and manage workgroups to develop technology standards, and handle administrative duties relating to IETF technology. The IESG meets biweekly as a virtual team in teleconferences and three times per year in physical meetings. The

IESG minutes have recorded these biweekly meetings of the IESG since 1991. The membership of the IESG had changed yearly, and this longitudinal study examined the dynamics within the yearly TMTs (denoted as cohorts) and the similarities between cohorts. These factors, plus the primary researcher's familiarity with the IETF and the IESG, made it a productive research environment.

Phase 1 of this research examined whether the leader of the IESG made a difference in the TMT's consensus decision-making by reviewing the 20% of biweekly minutes of three IESG cohorts (2003, 2006, 2011) using an exploratory mixed-mode method with qualitative IPA analysis with thematic counts (Hares, 2012). Phase 1 found differences in the effectiveness of the IESG consensus decision-making between these three cohorts led by different leaders. Fielder's (1964) LPC contingency theory was used to explain this result in that leadership actions (low LPC/task, high LPC/relationship) impacted group performance moderated by situational aspects of leader-member relationships, leadership position power, and task structure. In addition, post-research discussions with the three leaders suggested a fourth situational aspect of the relationship between co-workers that solidarity might explain.

Phase 2 of this research sought to determine if leaders and co-workers impacted effective consensus decision-making rather than just the leader. Phase 2 research combined an explanatory mixed-mode method of a one-time survey of the members of 25 cohorts of the IESG from 1986 to 2013 followed up with semi-structured interviews to evaluate the results. The researcher received survey responses from 48% of the IESG members and 57% of the IETF chairs. This researcher concluded that IESG member perceptions of solidarity (horizontal and vertical) and OCB (generalized compliance and altruism) were antecedents of effective consensus decision-making (perceived and real) in all 25 yearly cohorts. This researcher also found that solidarity was a better predictor of the perception of effectiveness in consensus decision-making than OCB. However, post-survey interviews revealed potential moderating impacts of conflict and issues with online statistics. Due to these limitations, Phase 3 research augmented Phase 2's theoretical model with TC and RC as moderators of solidarity as an antecedent to effective decision-making.

This document focuses on Phase 3 of this research focused on a 26-year longitudinal study (1991 to 2016) using concurrent triangulation mixed-mode methodology. This study had three data collecting strands (Strands-1–3) and 5 data analysis standards (Strands-1-5), and an interpretation phase. Two of the data analysis strands (Strands-4–5) triangulated the results of the earlier analysis (Strands-1–3). Strand-1 collected data associated with biweekly minutes of the IESG. Strand-2 collected online data regarding WGs and technologies the IESG standardizes. Strand-3 collected data from a 2017 survey of the IESG. Strand-1 analyzed these minutes using IPA techniques to detect the same behavioral themes found in the 2017 survey. Strand-2's IPA analysis of the IESG minutes coincided with the Strand-1 analysis but examines the themes for IESG decisions. A mixed-mode quantitative method analyzed these theme counts for IESG group behaviors and decisions using descriptive statistics, scale reliability tests, correlation, and HRM modeling (after testing for suitability of these tests). Strand-1 and Strand-2 qualitative analysis examined these data qualitatively based on the theoretical model. Strand-3 had a one-time re-survey of the IESG members that expanded the survey done in Phase 2 by including survey questions on TC and RC plus open-ended questions on task and conflict. The quantitative method in Strand-3 used data gathered from survey responses on IESG group behaviors and perceived effectiveness of group decisions, and the theme counts gathered from Strand-2 IPA analysis on real results. Strand-3's analysis method used the same quantitative tests on these statistics as the IPA analysis (descriptive statistics, scale reliability tests, correlation, and HRM modeling).

Strand-4 and Strand-5 sought to make a reliable conclusion on whether the research hypotheses are true. Strand-4 quantitative compared the IPA analysis and the survey results to determine whether the research hypotheses were true. Strand-5's qualitative analysis compared the qualitative results from the IPA analysis and the survey with the historical data gathered about the progression of standards in the IETF to achieve its mission. Qualitatively effective IESG decisions would help the IETF publish high-quality, relevant technical standards to improve the Internet.

Chapter 2 – Literature Review

This theoretical review reviews theory and experimental research on consensus decision-making, task and conflicts in SMTs, mixed-mode studies on team leadership, and research on the IETF and consensus decision-making in the IESG. The consensus decision-making research includes the definition of consensus decision-making, consensus decision-making processes, consensus-decision-making in teams, and solidarity. The mixed-mode studies on team leadership describe the methods of mixed-mode studies and the application of these methods to the study of team leadership and TMTs. Finally, this theoretical review provides the background for the proposed model of consensus decision-making with solidarity antecedents and the methodology for this research.

Consensus Decision Making

Consensus decision-making is used to make critical decisions in senior management teams (SMTs), cross-functional, virtual multi-national teams, and in teams creating international standards for the Internet (Bradner, 1996). Parker (2006) stated that effective teams used consensus decision-making for “key decisions” (p. 666). Teams use consensus decision-making to increase each member’s commitment to implementing the group decision (Amason, 1996; Korsgaard, Schweiger, & Sapienza, 1995; Parker, 2006). Consensus decisions require cooperative efforts by team members and team leaders to reach an agreement on the “problem statement and the recommended solution” (Parker, 2006, p. 666). Cooperative efforts include leadership behavior, group cohesiveness, cognitive process, emotions reduction of team conflict, solidarity, and organizational citizenship behavior (Cole & Bedeian, 2007; Kotlyar et al., 2011; Rowland & Parry, 2009; Sanders & Schyns, 2006b).

Decision-making by consensus by teams or small groups has been the object of empirical and theoretical studies in leadership for over 50 years (Dess & Origer, 1987; Kotlyar et al., 2011; Rowland & Parry, 2009). In addition, a few researchers (Kotlyar et al., 2011; Sanders & Schyns, 2006a, 2006b) have examined how leadership or team cooperative behaviors interact with team consensus decision-making. This section reviews the relevant literature on how consensus decision making, solidarity (horizontal and vertical), and leadership style impact solidarity.

Consensus Decision-Making in Teams

Early empirical research on consensus decision-making focused on SMTs who used consensus decisions to determine organizational directions (Dess, 1987; Dess & Origer, 1987). Dess and Origer (1987) defined consensus as “agreement of all parties to a group-decision … [that] occurs only after deliberation and discussion of all pros and cons of the issues, and when all (not a majority)” (p. 313) of the team members were in agreement. Current research on consensus decision-making has investigated team consensus decision-making at all levels of an organization (Kotlyar et al., 2011; Sanders & Schyns, 2006a; Tagger & Ellis, 2007; Van Ginkel & van Knippenberg, 2012). Both current and past researchers examined consensus decision-making by examining team relationships, team decision-making processes, the process of forming consensus, and the challenges to consensus.

Team Relationships

Research on consensus decisions shows that group relationships with high levels of cohesiveness, TI, and cooperation have stronger consensus decisions (Dess, 1987; Dess & Origer, 1987; Kotlyar et al., 2011; Sanders & Schyns, 2006b; Zaccro, Ritman, & Marks, 2001). Conversely, groups with affective or relational conflict will have a lower quality of consensus decision-making, lower group satisfaction, and lower effective acceptance of group decisions (Amason, 1996; Jehn, 1994, 1995). Conversely, increasing group cohesion increases group productivity, satisfaction, social influence, and intragroup interactions (Barker, Wahlers, & Watson, 2001).

Group cohesiveness refers to a product of complex forces that “bind members to a group” (Barker et al., 2001, p. 57). The complex forces of group cohesion include socioemotional cohesion based on interpersonal attraction, task cohesion, and situational cohesion (Barker et al., 2001; Cole, Bedeian, & Bruch, 2011). Interpersonal attraction occurs based on visible characteristics, such as age, or invisible characteristics, such as education (Jehn, Chardwick, & Thatcher, 1997). Interpersonal attraction follows the social identity theory of social categorizations, suggesting an individual divides people within an organization into in-groups and out-groups based on visible or invisible characteristics (Hogg, 2001; Hogg & Terry, 2000). This social categorization depersonalizes team members, team leaders, and those outside the team. This depersonalization occurs when individuals create stereotypes for the perceptions of in-group and out-group feelings and behaviors (Hogg et al., 2006; van Knippenberg & Hogg, 2003). Self-categorization causes an individual to internalize the stereotype behaviors as appropriate self-behaviors. Team leaders who have prototypical behavior of the in-group may be perceived as trustworthy, effective, and cooperative, whether the leadership behaviors or schemas demonstrate these qualities (van Knippenberg, 2011). Prototypical team members are perceived as aligning to the group prototype whether or not their behavior aligns.

Group cohesiveness may also be based on task or situational cohesiveness. Task cohesiveness is based on shared work values, expectations, and group norms regarding group tasks (Tagger & Ellis, 2007). Members with higher task cohesiveness put more energy into the group tasks (Prapavessis & Carron, 1997) and continue with the organization longer (Jehn, 1995; Jehn, Greer, Levine, & Szulanski, 2008). An example of task cohesion is the team that developed the iPhone. Teams with high TI within a group require a higher cohesiveness to be effective (Barker et al., 2001). Situational group cohesion occurs when events cause cohesive behavior.

Sanders and Schyns (2006b) defined cooperation as a “contribution of individual effort, time, and resources to interdependent tasks and actions that benefit group or organization” (p. 539). Cooperative behavior is characterized by reciprocity in two directions: horizontal and vertical. Gouldner (1960) defined reciprocity as “a mutually contingent exchange of gratifications” (p. 161). Horizontal reciprocity occurs between team members. Vertical reciprocity can exist between a team leader and a team member. The social exchange theory suggests that, based on childhood and previous experiences, team members have expectations on equitable social exchanges between team members and between leaders and team members (Cole, Schaninger, & Harris, 2002).

Decision-Making Processes

Effective team decision-making requires a process to obtain a quality decision that all team members can commit to cooperatively implementing (Kotlyar et al., 2011). For example, a quality decision for a problem-solving team requires team members to work cooperatively to state the problem and select a solution. Precise and clear communication of a problem statement and the value of each potential solution are necessary components of good decision-making. Clear communication of problems and solutions requires an atmosphere of “openness for sharing opinions” (van Woerkom & Sanders, 2010, p. 139).

Team leaders can either elaborate the problem statement further or seek common ground on consensus on the problem statement. Van Ginkel and van Knippenberg (2012) found that teams followed the team leader by seeking elaboration or common ground during decisions. A leader can create a favorable climate for resolving dissent into consensus by periodically summarizing the common ground on problem statements, potential solutions, and a selected solution (Parker, 2006; Rowland & Parry, 2009). Rowland and Parry (2009) noted that organizational structures with relational leadership and lateral job roles for team members aided the decision-making process.

Decision-making teams have diverse members to gather the information needed for critical decisions (Mitchell & Boyle, 2008). Jehn, Northcraft, and Neale (1999) found that this “informational diversity positively influences group performance” but that its impact was “mediated by task conflict” (p. 741). Cognitive TC via devil’s advocate or dialectical inquiry techniques may provide initial benefits, but these initial benefits diminish quickly (Amason, 1996). Research has also shown that the full potential of diverse information has been under-utilized in these decision-making teams (Bjorklund & Holt, 2011; Van Ginkel & van Knippenberg, 2012). Both sets of research show that to use diverse information, teams must share common mental models for (a) the team’s tasks, (b) how to process information regarding problems and solutions, and (c) how to reach a consensus (Van Ginkel & van Knippenberg, 2012; Zaccro et al., 2001). Developing these shared mental models within a team when a group starts from diverse viewpoints may take time and arouse conflicting ideas and emotions.

Team leaders' expectations can help set team norms for problem-solving activities (Tagger & Ellis, 2007). Team problem-solving activities will include resolving team conflicts, collaboratively working to find problem solutions, strong inter-team communication, goal-setting, planning tasks, and coordinating tasks actions. Team leaders and teams striving for cohesion and cooperative behavior must allow for critical thinking within a team. Adverse environments, such as threats or crises, may cause teams to increase group cohesiveness and rigidity of the response (Harrington, Lemak, & Kendall, 2002).

Forming Consensus

Groups use consensus decision-making when all members need to commit to implementing a group decision (Amason, 1996; Kotlyar et al., 2011). Apparent consensus decisions delayed by team members or cynically presented are really non-consensus decisions (Amason, 1996). In consensus decision-making, each group member must agree that a particular solution is acceptable enough and commit resources to enact it in a timely fashion (Yukl, 2010; Korsgaard et al., 1995). A crucial component in each member's commitment to a consensus decision is the perception of procedural fairness in hearing each member's viewpoint before the decision. Parker (2006) suggested that consensus decision-making should only occur if "no clear answer" (p. 667) exists, no single expert source of opinion exists in the group, and sufficient time remains available.

Challenges That Consensus Decision-Making Must Overcome

Consensus decisions must overcome divisive organizational behaviors, team relational conflict, reciprocity of attacks, and differences perceptions of the task. Organizational behaviors can either contribute toward the consensus decisions of a team or be divisive. Supportive behaviors are those behaviors in which individuals become "good soldiers" who "take the organization's goals as their own, are highly committed, and don't question organizational decisions" (Fields, 2002, p. 235). Divisive behaviors can be free riders (Hechter, 1987), “smooth operators,” or a “saboteurs” (Fields, 2002, p. 235). Hechter (1987) described a free rider as an individual who benefits from a group without putting in personal effort. Smooth operators set their own goals as the top priority, optimizing every action to maximize their benefits. Saboteurs do not optimize their goals or the organization’s goals but “passively resist authority” and “violate work rules” (Fields, 2002, p. 235). Smooth operators and saboteurs are both divisive behaviors in consensus decisions.

Team member diversity arises from visible individual dissimilarities (e.g., age) and invisible dissimilarities (e.g., education). Jehn, Chadwick, and Thatcher (1997) found that visible differences increased RC. However, invisible dissimilarities increased TC. Diversity within team members may also create subgroups within the social network of the group or team. Social subgroups form along multiple alignments of subgroup characteristics, such as social grouping, education, or in-group/out-group distance from a group fault-line (Bezrukuva, Jehn, Zanutto, & Thatcher, 2009; Thatcher, Jehn, & Zanutto, 2003). If only two subgroups form along fault lines or no-fault lines, there is a higher potential for group conflict during consensus decision-making. During conflicts within a group, one attack may cause a reciprocal attack, increasing team conflict.

Solidarity

Hechter (1987) defined solidarity operationally as “the greater the average portion of each member’s private resources contributed to the collective ends, the greater the solidarity of the group” (p. 18). Hechter’s (1987) definition of solidarity comes from his rational choice model of society where individuals as rational actors operate micro-economic tradeoffs between-group obligations and individual personal goals. Institutions and groups operate by a set of rules that constrain individuals and obligate individuals to efforts to create jointly developed goods. Groups can benefit individuals by producing exclusive goods or social support for the members. Hechter (1987) stated that the rational choices model could explain the free-rider problem and differentiated compliance with group obligations or norms. The free-rider problem occurs when individuals enjoy group benefits without contributing individual resources. The group detecting a free rider can use group norms to limit the free rider's access to group-developed (exclusive) goods or social support. The variance in how individuals comply with group obligations depends on how much an individual values group resources. If the individual easily obtains group goods or services and movement between groups occurs without cost, the individual may minimize the group obligations. However, if the individual values the social support and goods from a group, this person may invest more in the group committing more resources (social and material). A high cost of exit or entrance to a group providing unique products may also increase the individual willingness to pay obligations to remain in the group.

Hechter's (1987) operational definition of solidarity aligns with the concepts of cooperative OCB of the *good soldier*. Good soldiers in an organization or a team do things outside their role (ERBs) to promote their group's success. Fields (2002) listed the following four OCB instruments with which to test organizational citizen behaviors: Smith, Organ, and Near's (1983) OCB; Podsakoff and Mackenzie's (1990) OCB; Williams and Anderson's (1991) OCB; and Moorman and Blakely's (1995) OCB. Organ (1997) and Podsakoff et al. (2000) questioned whether organizational citizen behavior defined by these four instruments was distinct from in-role behavior and formed a construct that would influence organizational performance based on clear antecedents. Koster and Sanders (2006) found that OCB did not as accurately predict cooperative behaviors as the combination of HS and VS.

Koster and Sanders (2006) used solidarity to define the cooperative behavior within a group and stated that "solidarity involves at least two people who choose to cooperate or not" (p. 523). HS involves one or more co-workers cooperating by putting in extra private resources toward the collective effort. VS is solidarity between a team leader and the team members (Sanders & Schyns, 2006b). Koster and Sanders (2006) pointed out that the receiving party often reciprocates solidarity when the other party chooses to exhibit solidarity behaviors. Sanders and Schyns (2006b) found that HS and VS did not always occur together.

Both HS and VS arise from team members and leaders' choosing to be reciprocal and reinforce group identity and cohesiveness within social networks.

Horizontal Solidarity

HS behavior depends on an individual's choices. First, the individual must choose self-categorization to a group or team. Second, the individual values the group enough to exhibit extra-role solidarity behaviors. Sanders and Schyns (2006b) tested for these ERBs in an empirical study: (a) helping a co-workers finish tasks, (b) helping "a co-worker when things go wrong that nobody is responsible for," (c) apologizing to a co-worker for mistakes, d) dividing "pleasant and unpleasant" tasks with a co-worker, and (e) "living up to agreements with co-workers" (pp. 542–543). These behaviors each included individual choices, extra-role efforts, and benefits for the co-worker.

Hechter's (1987) rational choice model is built on top of reciprocity in social exchanges. Individual actors choose to give resources for which the group reciprocates by providing access to exclusive goods or services. Social exchange theorists (Cole et al., 2002) regard cooperative social behaviors as based on reciprocity. Koster and Sanders (2006) found that an employee's perceived solidarity with co-workers was the strongest predictor of solidarity. Koster and Sanders also tested for the OCB behavior of altruism and generalized compliance using the Smith et al. (1983) OCB instrument and found that workers who perceived HS with co-workers experienced altruistic behaviors from co-workers.

HS behaviors exist within the complex forces of group cohesiveness and identity. The self-categorization inherent in HS increases in-group behaviors. The self-categorization theory states that an individual who self-categorizes will accent intra-category similarities, increase communication similarities, and increase cohesiveness to the group (Hogg & Terry, 2000). However, HS is different from affinities or emotional ties of in-group behaviors. It operates in individual choices to provide extra personal resources to support the group. Koster, Stokman, Hodson, and Sanders (2007) found that HS was positively related to TI within a group.

Solidarity behaviors have substantial costs, so the "free-rider" behavior may be tempting within a group or team. Forces that work against free-rider behavior are the inherent reciprocity and group/team norms within a social network. Forces that promote free-rider behavior are egoist tendencies in the individual and out-group social networks. Seers (1989) defined team-member exchange (TMX) as the "individual member's perception of his/her exchange within the peer group as a whole" (p. 119). Like leader-member exchange (LMX), the TMX can differ for each member. Seers, Petty, and Cashman (1995) indicated that TMX was a negotiation of role expectations between the individual and the team as part of the team norming processes.

These social ties between the individual and team impact the individual's informal networks (ego networks), the formal relationships to the organization, and the inter-organization relationships (Balkundi & Kildruf, 2006). The intra-level coalition within the social network that the individual participates in may enhance or detract from the voluntary monitoring of other team members to prevent the "free-rider" problem. An individual may not expend effort to monitor social networks if the group's social capital is lower than the coalitions within the individual's informal, organizational, or inter-organization networks. In contrast, an individual who highly values the group's social capital may exhibit solidarity behaviors that include monitoring team members to prevent "free riders."

Vertical Solidarity

VS behaviors originate from team leaders' choices in the leader-follower relationship. Researchers have examined the two leadership styles of transformational leadership (Bass & Riggio, 2006) and LMX (Graen & Uhl-Bien, 1995) to determine if VS links to particular leadership behaviors. A transformational leader uses idealized influence, individual consideration, intellectual stimulation, and inspirational motivation with followers. Followers choose to respond to the transformational leader's positive emotions and actions. Sanders and Schyns (2006b) conducted an empirical test with 193 employees and 35 teams in Dutch firms. Sanders and Schyns (2006b) found that "[the] relationship between group cohesiveness and [group] vertical solidarity is positive if employees perceive their supervisor as high transformational leader, and negative if the employees perceive their supervisor as low" (p. 542). This research showed that the leader’s choice of leadership style in a cohesive group impacted the VS in the group.

The LMX form of leadership similarly offers leaders and followers choices. A leader may offer high quality of exchange to the follower, but a follower can choose not to engage in the relationship (Graen & Uhl-Bien, 1995). Schyns, Kroon, and Sanders (2006) examined the relationship between LMX behaviors of followers and leaders and the VS behaviors between followers and leaders in an empirical study with 360 Dutch employees. The LMX behaviors studied in this research consisted of respect, positive affect, and loyalty between the leader and follower. Schyns et al. (2006) found a positive relationship between a leader’s LMX behavior and VS. These two studies reinforce the concept that VS is based on leaders' and followers' choices and a norm of reciprocal actions.

Empirical researchers (e.g., Koster & Sanders, 2006; Sanders & Schyns, 2006b; Schyns et al., 2006) tested for VS behaviors. All these researchers checked for the following solidarity behaviors from an employee to a supervisor: (a) helping a supervisor finish his/her tasks, (b) helping “a supervisor when things go wrong that nobody is responsible for,” (c) apologizing to the supervisor for mistakes, (d) dividing “pleasant and unpleasant” tasks with a supervisor, and (e) “living up to agreements with [a] supervisor” (Koster & Sanders, 2006, p. 537). The supervisor solidarity behaviors to the employee included (a) “supervisor helps me [employee] finish tasks,” (b) “supervisor is willing to help me [employee] when things went wrong unexpectedly,” (c) “supervisor apologizes to me they have made a mistake,” (d) “supervisor divides pleasant and unpleasant task equally between them and me,” and (e) “supervisor lives up to agreements” (Koster & Sanders, 2006, p. 537).

Sanders and Schyns (2006b) found that group perception of strong transformational leadership links group cohesiveness and VS. Schyns et al. (2006) found a positive relationship between LMX, TMX, solidarity, and team performance (horizontal and vertical). This positive relationship suggests that VS positively affects group cohesiveness and performance. However, additional studies need to examine VS’s impact in other contexts, such as decision making. Schyns et al. suggested repeating this research with non-Dutch firms.

TC and RC in Strategic Management Teams and Top Management Teams

SMTs and TMTs are decision-making teams that solve organizational problems, chart strategic directions, and implement these decisions within the organizations. Consensus on a decision among the TMT or SMT members facilitates the widespread implementation of the TMT decision. Still, some types of conflict among members may decrease the quality and influence of a decision. Yukl (2010) indicated that modern research had found that critical determinants for high-performing teams included effective task actions, team relationships, and group processes. Some behaviors that increase task effectiveness are a commitment to shared objectives (Podsakoff, MacKenzie, & Ahearne, 1997), accurate shared mental models (Edwards, Day, Arthur, & Bell, 2006; Klimoski & Mohammed, 1994; Lim & Klein, 2006), and skills and role clarity of team members (Marks, Zaccaro, & Mathieu, 2000). Some examples of effective relationship behaviors include “mutual trust and cooperation” (Yukl, 2010, p. 342) and collective efficacy and potency. Effective group processes include an ability to coordinate with external groups and an “internal organization” that coordinates tasks while maintaining team relationships (Rico, Sanchez-Manzanares, Gil, & Gibson, 2008). Finally, high-performing decision-making teams have effective processes to pool information, ideas, and resources by using clear communication to discuss a problem and come to a high-quality decision.

Amason and Sapienza (1997) and Jehn et al. (1999) found group processes in decision-making groups were impacted by the following group characteristics: “group size,” “status differentials,” “cohesiveness,” “membership diversity,” “emotional maturity,” “physical environment,” and “communication technology” for virtual teams (Yukl, 2010, pp. 354–356). Jehn (1995, 1997) noted that TC, RC, or process conflicts might negatively affect workgroup performance. However, TC can benefit group cognitive processes when groups tackle non-routine tasks in some circumstances. Jehn and Chatman (2000) found that teams had high performance and satisfaction if task conflict was proportionally higher than affective or process conflict. However, perception of conflict influences how individuals experience conflict. If one member’s perception of conflict is higher than another member's perception, this member will experience more conflict. For example, Jehn and Chatman (2000) found that when members of production teams had high levels of perceptual disagreement regarding the levels of relationship and process conflict, it resulted in "lower performance and more negative attitudes" (p. 69).

Researchers have found that cognitive conflict increases the quality of TMT's decision-making processes, understanding, and organizational performance resulting from those decisions. Factors that increase cognitive conflict may increase affective conflict, which decreases the quality of decision-making and corporate results (Amason, 1996; Amason & Sapienza, 1997). Ensley and Pearce (2001) found that cognitive conflict among TMTs of new ventures increased shared strategic mental models and effective decision-making resulting in organizational growth, profits, and revenue. However, this same cognitive conflict also increased affective (relational) conflict. When this RC reached a certain point, it decreased new venture performance in organizational growth, profits, and revenue. Amason and Sapienza (1997) found that larger sizes of TMTs increased cognitive and affective conflict. Larger TMTs can also increase mutuality and openness levels, decreasing relationship (affective) conflict and allowing cognitive conflict. Amason and Sapienza (1997) defined mutuality as "shared consequences" for decisions where individuals may accommodate others for the "good" (p. 500) of the team. Guinot, Chiva, and Mallen (2015) found that altruism directly facilitated organizational learning and indirectly reduces relational conflict since relational conflict decreases learning. Ensley and Pearce (2001) found that an increase in the sense of belonging decreased cognitive and affective conflict. Amason and Mooney (1999) found that a team's past performance was an antecedent of affective conflict, but past performance is not an antecedent of cognitive conflict. Jehn and Bezrukuva (2010) found that hidden fault-lines that divided a group based on individual differences (e.g., race, nationality, background, or some entitlement) might go from dormant to active, either causing coalitions to form or intragroup conflict. Similarly, Ayub and Jehn (2014) found that two nationalities within a group created a potential for conflict and lower performance. A group with people from various nations or cultural backgrounds has the opposite effect. Solidarity considers a sense of cooperation within a group, but it is unclear how solidarity interacts with different levels of conflict or antecedents of conflict.

The results from these researchers align with a theory from Stewart et al. (1999) that teams go through stages of development. The first stage is the forming stage which occurs during the gathering of individuals into a group. Next, the group enters the storming stage as members interact and conflict over tasks, interrelating, and group processes. Finally, the group enters the third stage, when members develop norms for these facets of daily work. These three development stages are necessary before the group can perform its tasks effectively. In healthy team development, the team develops shared cognitive models, openness, and healthy interpersonal relationships that increase the sense of belonging. If conflicts remain after the storming and norming phases, it impacts team performance. These conflicts can be hidden, perceptual, or actual conflicts about tasks, relationships, or processes.

Loughry and Amason (2014) indicated that debates continued regarding TC's value because people view conflict differently. Theory and solid empirical evidence support the view that TC creates high-quality decisions with greater acceptance, enhancing organizational performance. In contrast, relationship (affective) conflict decreases the quality of decision-making and the acceptance of decisions resulting in decreased organizational performance. Loughry and Amason (2014) reviewed four meta-analyses on different types of conflict by DeDreu and Weingart (2003); deWit, Greer, and Jehn (2012); DeChurch, Mesmer-Magnus, and Doty (2013); and O'Neill, Allen, and Hastings (2013). Based on these meta-analyses, they concluded that the debate regarding TC was based on (a) intervening effects complicate the relationship between conflict and team, and (b) how people perceive conflict differently (Loughry & Amason, 2014). Based on this complex nature of conflict, any research investigating the conflict in a team in an empirical study needs to define the type of conflict precisely. Furthermore, surveys in such an empirical study on conflict should use tested instruments for that type of conflict in that environment. For example, Jehn’s (1994, 1995) ICS was used in four research projects on TMTs (Amason, 1996; Amason & Mooney, 1999; Amason & Sapienza, 1997; Bucholtz, Amason, & Rutherford, 2005; Ensley & Pearce, 2001; Lankau et al., 2007). Some of these studies used the ICS scale with 7 items instead of 9 items (4 for RC and 3 for TC). Due to these studies, Pearson et al. (2002) assessed the ICB scale and found that 6 out of 9 items best captured the ICS scale. Therefore, this research uses the 6-item scale.

Consensus Decision-Making in Information and Computer Technology Standards

Few empirical studies have examined consensus decision-making processes in ICT SDOs. Instead, research on ICT SDOs focuses on the impact of internal and external forces have on the SDO's processes or how the SDO impacts external companies or markets. Research into the internal forces for ICT SDOs includes the following: research into the social construction of SDOs (Ding et al., 2013; Egyedi, 2003), teams in open source communities (Egyedi, Vrancken, & Ubacht, 2012; Fielding, 1999), surveys on improving standards bodies (Spring et al., 1995), legitimizing open standards (Werle & Iverson, 2006), and the ecology of standards bodies (Nickerson & Muehlen, 2006).

Within a few of these internal topics, researchers have examined the group dynamics surrounding consensus decision-making in open standards or open-source projects delivering code or the group dynamics of the IETF. In addition, researchers have studied how external forces and external companies impact ICT standards. These external forces which affect ICT SDOs include competitive product strategies (Besnahan & Yin, 2007), patents (Rysman & Simcoe, 2010), and technological and economic forces (West, 2007). Researchers have also studied how the standards created by ICT SCOs exert forces on external companies and society. The research on the effects of ICT SDOs on external companies includes: how ICT standards influence product development (Rachuri et al., 2008) or how standards technology diffused into an industry impacts companies and market development (Choi, Lee, & Sung, 2011). In addition, Allen and Sriram (2000) studied how the innovations in ICT standards impacts societies.

Egyedi (2003) suggested that social construction theory was the best way to view open standard creation in ICT SDOs or open-source projects. Within this framework, the social actors consist of the SDO's committees or WGs, the SDO, and the actors who influence the standards. These social actors are bound together by various attributes, including social shaping by a paradigm, process, goals or targets, and role. For example, Fielding (1999) found that creating open-source code in the Apache project bound the actors to a paradigm of group decisions on code using a consensus decision-making process to determine code included in a release. O'Mahony and Ferraro (2007) conducted a mixed-mode study to examine how leaders emerged in the Debian open source project over 13 years with the same paradigm of open source code, consensus decision-making process, and the goal of a Debian Linux OS release. O'Mahony and Ferraro found that the Debian open source community expected leaders to build consensus and represent that consensus of the developers in decisions to the technical steering group.

Mixed-Mode Designs in Leadership and Group Research

Mixed-mode research can be defined as a philosophy, methodology, research design paradigm, research program to get a complete view of a problem, and combining two popular techniques. Researchers have adapted mixed-mode methods to specific types of research in leadership, nursing, health, education, and public health fields. As mixed-mode research methodology has evolved, leadership research has used these evolving techniques to examine individuals, groups, and teams, including TMTs. This section reviews the evolution of mixed-mode methodology, different types of mixed-mode methods, and why concurrent triangulation mixed modes improve the reliability of research results. From this beginning, the section shows how concurrent triangulation of qualitative methods with quantitative survey methods mitigates threats to qualitative and quantitative validity and reliability. Next, this section discusses researchers use concurrent triangulation mixed-mode methods to investigate groups, teams, and TMTs due to the complex web of behaviors of the participants. This review of mixed-mode methodology concludes with a detailed review of two articles on mixed-mode methods used in leadership research.

The first article is a meta-analysis by Stentz, Plano-Clark, and Matkin (2012), which reviews 55 published leadership articles with mixed-mode methodology. The second article is Rowland and Parry's (2009) mixed-mode research on antecedents of "consensual commitment" (p. 1) on effective consensus decision-making, defined as those decisions that had post-decision solidarity and commitment to objectives. Rowland and Parry (2009) found that organizational designs that enhanced a leader's "relational leadership style" and "generated later job moves" (p. 1) enhanced team consensual commitment and protected against dysfunctional team dynamics. These two papers provide the background on mixed-mode methods necessary to the literature review on consensus decision-making in the IETF and the IESG as a TMT.

Mixed-Mode Methodology: Evolution and Definition

Early mixed-mode researchers combined multiple data sources such as a quantitative survey with qualitative interviews. In the 1980s and 1990s, researchers using mixed-mode research debated whether qualitative and quantitative methodologies are divergent paradigms that could never be linked together or whether these paradigms could be linked in specific situations or pragmatically by researchers. From 1990 to 2000, mixed-mode researchers developed best practices for combing qualitative and quantitative methods and describing the combined methods in a standard notation. This notation used "QUAN" for a primary quantitative methodology, "QUAL" for a primary qualitative methodology, "quan" for a secondary quantitative methodology, and "qual" for a secondary methodology. Secondary methods embedded in a primary method were placed inside a primary methods box (see Table 3). Concurrent collection methodologies were shown side-by-side with arrows to analytical processes. Sequential parts of the process were indicated with arrows between the two process boxes. Since 2000, mixed-mode research has advanced via expansion of techniques, with application of these techniques to nursing, interpretation of results, and critiques of previous work, such as definitions.

Creswell and Plano-Clark (2011) pointed out that the critiques of mixed-mode researchers have surfaced definitions since its inception. Some researchers ask whether the constructionist philosophy, which asserts “multiple ways of seeing, hearing, and making sense of the social world,” is valid (Creswell & Plano-Clark, 2011, p. 3). Some researchers have begun with one method of philosophy (Greene, Caracelli, & Graham, 1989) and progressed toward the constructivist viewpoint (Greene, 2007). Creswell and Plano-Clark (2011) focused on how mixed-mode methods combined a specific philosophy of inquiry with a research design methodology. Creswell and Plano-Clark (2011) defined mixed-mode research as one in which the researcher (a) “collects and analyzes persuasively and rigorously both qualitative data” to address the research questions, (b) “mixes the two forms” by both embedding the methodologies and combining data for analysis, (c) uses different mixtures in a theme as part of a “multiple phase investigation” (p. 5), and (d) frames these with theoretical lenses and a pragmatic world view. Creswell and Plano-Clark (2011) suggested that the combination of qualitative and quantitative provides a better understanding of research problems than either approach alone.

Creating Mixed-Mode Methodology

Creswell and Plano-Clark (2011) pointed out that mixed methods provide a general framework. Within this framework, the researcher selected specific options and justifies their choices. Researchers since the 1990s have engaged in creative, critical evaluation of specific frameworks to advance the capabilities and benefits of mixed-mode research. Greene et al. (1989) suggested that the benefits of mixed-mode research entailed (a) triangulation of results via complementary results that would clarify, enhance, and illustrate each other; (b) initiate the discovery of new frameworks or paradoxes; and (c) expansion of the breadth of research. Bryman (2006) expanded on these three reasons. Bryman suggested mixed-mode triangulation offsets the weaknesses of either approach, enhances of credibility of results, and allows investigation of unexpected results from one method by using a second method. In the area of complementary, Bryman (2006) suggested mixed-mode methods could aid the discovery of new paradigms or paradoxes by (a) providing completeness of inquiry that includes diverse views, (b) the ability to explain the results of research using one method by the research in using a second method, and (c) ability to answer two research questions each suited toward a different method. For researchers who seek to expand the breadth of existing research, mixed-mode research can use the sequencing of different methods to build on or confirm results, develop instruments, or facilitate different sampling. These experts on mixed-mode methodology suggest a detailed description of the reasons behind a particular mixed-mode methodology (Bryman, 2006; Creswell & Plano-Clark, 2011; Greene et al., 1989).

The guidelines for evaluating the benefits and deficits of a specific mixed-mode study come from the theoretical discussions on evaluating prototypical mixed-mode designs. For example, Greene et al. (1989) assessed prototypical mixed-mode techniques based on the categories of method interactions, the timing of method strands, the relative priority of strands, the mixing strategy of methods and interaction points, and theoretical framework or philosophy. Creswell and Plano-Clark (2011) defined the following six prototypical designs based on these five evaluation criteria: explanatory sequential design, exploratory sequential design, transformative design, embedded design, convergent parallel design, and multiphase. Creswell (2009) included a seventh “concurrent triangulation strategy” (p. 20) prototypical mixed-mode. Table 3 summarizes how researchers (Creswell, 2009; Creswell & Plano-Clark, 2011; Tashakkori & Teddlie, 1998) indicate these seven prototypical designs differ on these five criteria.

For clarity, Table 3 combines method interaction and timing of method strands into one column. The convergent parallel design, embedded design, multiphase, and concurrent triangulation utilize pragmatism as the umbrella theoretical framework and philosophy. All of these prototypical designs, except the transformative design, use a constructivist paradigm for the qualitative strands and a post-positivist paradigm for the qualitative strands. The transformative design uses a transformational paradigm to guide experimental designs. A convergent parallel design differs from the concurrent triangulation strategy in the interaction points. A convergent parallel design interacts at data analysis or interpretation after completing separate qualitative and quantitative data analysis. A convergent triangulation strategy may have interaction points during data collection, during data analysis, and at the interaction points.

Table 3: Types of Mixed-Mode Design

Design	Interaction/ timing	Relative priority of strands	Mixing strategy/ interaction point	Theoretical framework/ philosophy
Explanatory sequential design	Interactive/ sequential, QUAN→ qual	Quantitative higher	Mixed / After quantitative to qualitative results to aid interpretation of quantitative	Explicit / Phase 1: post-positivist Phase 2: constructivist
Exploratory sequential design	Interactive / sequential QUAL→ quan	Qualitative higher	Mixed / QUAL at collection, qualitative data aids interpret of qualitative data	Implicit /Explicit Phase 1: Constructivist Phase 2: Post-positivist
Transformative design	Interactive / Any	Any option based on a theoretical framework	May mix during all phases	Transformative theoretical framework / transformative world view
Embedded design	Interactive, concurrent QUAL or QUAN	Main method enhanced by a secondary method	May mix during design, data collection, data analysis, or results	Implicit or explicit theory or world view in the main method or pragmatist theory if concurrent
Convergent parallel design	Independent, concurrent QUAL + QUAN	Equal	Integrating: at data analysis or interpretation / after separate data analysis	Explicit theory / Pragmatism is the “umbrella” theory (Creswell & Plano-Clark, 2011, p. 73)
Concurrent triangulation	Interactive, concurrent QUAL + QUAN	Equal	May integrate at data collection, data analysis, or interpretation	Explicit theory / Pragmatist umbrella theory
Multiphase design	Interactive, Sequential studies where each study informs the subsequent study	Each phase of study may have a different method	Each phase may be a single type or mix types	Implicit or explicit theory per phase or the world view of phases is pragmatic

These prototypical mixed-mode research designs vary based on the purpose behind particular mixed methods for triangulation; the priority of the strands; and the strategy complementary results of qualitative (QUAL) and quantitative (QUANT), initiation of discoveries or paradoxes, and expansion of the breadth of research. Table 4 summarizes how these four facets of purpose for mixed-mode designs vary within seven prototypical designs.

Table 4: Purpose for Prototypical Designs

Prototype design	Triangulation	Complementary results	Initiation of discoveries	Expansion of existing research
Explanatory sequential design	none	Qualitative results explain quantitative data	none	Form groups based on quantitative for purposeful qualitative exploration
Exploratory sequential design	none	Quantitative results generalize qualitative data	Explore and refine new areas. Develop instruments.	None.
Transformative design	none	none	Identify social injustice	Challenge social injustice
Embedded design	Need validation or follow-up on results	Need for a broader understanding of phenomena provided by qualitative and quantitative datasets	Need exploration before experiments or survey	Expand the scope of research by expanding methods
Multiphase design	Triangulates data between studies	No single study can answer all questions	Supports needs assessment, program development, and program evaluation	Handle an incremental set of research questions
Convergent parallel design Concurrent triangulation design	Methods validate or collaborate with each other.	Collect complementary data	Research allows initiation of discoveries due to synthesizing of qualitative and quantitative	Expands research scope by design

The first prototypical design, explanatory sequential, focused on complementing quantitative research (e.g., a survey) with qualitative research (such as qualitative interviews that explain or discuss the survey). A second example of the explanatory sequential research design is a quantitative survey to help select groups within a population for qualitative exploratory research. The second prototypical mixed-mode design, exploratory research, seeks to gain from qualitative research the knowledge that quantitative research applies to a specific problem. For example, research to create a new survey instrument uses a qualitative study to form a pool of information to help form new instruments and a quantitative survey to confirm the specific instrument. Transformative mixed-mode, the third prototypical design, uses a mixed-mode design to identify new social injustice and expand research. The last four prototypical designs (embedded, multiphase, convergent parallel, and concurrent triangulation) aim to expand existing knowledge. Because these four designs seek to expand existing knowledge, reliability of the expanded knowledge and integration with existing knowledge is fundamental to the design. Due to these goals, triangulation is a critical component of these four designs.

Using Concurrent Triangulation to Improve Validation and Reliability

Researchers of concurrent triangulation mixed-mode design use methods that validate or collaborate with each other to improve the research validity of the qualitative and quantitative portions of the research. Research validity implies that the research results are valid, and reliability refers to the research's consistency and repeatability. Quantitative validity consists of internal validity, external validity, construct validity, statistical conclusion validity, and instrument validity (Cabanda, Fields, & Winston, 2011; Creswell, 2009; Girden & Kabacoff, 2010). Creswell (2009) suggested qualitative validity means the qualitative methodology takes high-quality, in-depth qualitative data from different sources.

Researchers garner high-quality, in-depth qualitative data in field studies by combining prolonged field observation, thick descriptions of events, and cross-checking by participants. Historiometric studies require researchers to obtain high-quality, in-depth data by careful data collection, material creation, and sampling. Coding methods for either type of study require peer debriefing and external auditors to check the coding. This section describes how concurrent triangulation can provide alternate methods for reducing qualitative and quantitative validity threats and bias in mixed mode methodology. This section will consider internal, external, construct, and statistical conclusion validity and bias due to common method variance.

Internal validity in quantitative surveys refers to the ability to draw conclusions on relationships between variables. Surveys that retested whole populations encountered events between the two tests that threaten internal validity. These threats came from factors outside the research, such as historical events, the aging of subjects (maturation), and subjects who died or left the group (mortality).

In addition, internal threats came from events before giving the survey's re-test that impact the results. Examples of this type of internal threat included errors due to diffusion of information regarding the survey instrument, biased encouragement for retesting, and early diffusion of post-survey results. Finally, the internal validity of the survey retest was threatened by differences when one individual had familiarity with the survey and another individual did not. Triangulation was used to quantify the level of internal validity. Triangulation of the first test with the retest (second test) was used to quantify internal validity issues. Triangulation qualitative research into historical records of the group interactions was used to compare the group's interactions before historical events, maturation, and mortality occurred. Table 5 summarizes ways to resolve threats to the internal validity of surveys via quantitative methods and triangulation with qualitative data.

Table 5: Resolution to Quantitative Validity Threats Using Concurrent Triangulation

Type of validity	Threat to quantitative survey research	Quantitative survey methodology resolution to threats	Concurrent triangulation resolution to threats
	Historical events Maturation Mortality	No resolution via survey research design	Triangulate with qualitative analysis of historical records of the group surveyed
	Regression or biased encouragement for retesting	Write a non-biased survey solicitation.	
Internal validity	Differences in test context or instrumentation.	Provide re-test instrument with the same sequence, and add questions at the end	Triangulate post-survey interviews after survey using qualitative methods.
	Differences in pre-test and post-test information	Provide the same pre-test and post-test information to all participants.	
	Interaction of participants selection and treatment (survey)	Restrict the claims regarding the results to the survey group.	Triangulate longitudinal studies within a group to other groups of a similar type.
External validity	Interaction of setting and treatment	Conduct the survey in new settings to see if results differ.	Retest the survey in a new setting.
	Interaction of history and treatment	Replicate survey at a later time.	Triangulate the re-test survey done at a later time with the original survey.

Type of validity	Threat to quantitative survey research	Quantitative survey methodology resolution to threats	Concurrent triangulation resolution to threats
Construct validity	Face validity Content validity	Not easily tested.	Triangulate results with post-survey qualitative interviews.
	Predictive validity, Concurrent validity, Discriminant validity of the instrument	Measure concurrently one or more alternate concepts using existing instruments.	Use mixed mode coding of qualitative documents by using survey instruments questions for predictor(s) and alternate predictor(s). Triangulate the predictor constructs.
Statistical Conclusion validity	Insufficient power for the statistical test	Select sample size to be large enough for the statistical method but small enough to get the effect.	Triangulation with qualitative analysis of parallel data from historical sources or Interviews.
	Violation of statistical assumptions	Run statistical tests on data prior to running statistical tests.	Triangulate survey data with post-survey data from interviews of participants regarding survey results and survey questions.
	Fishing	Run statistical methods only in appropriate circumstances	
	Unreliable survey instrument	Test reliability of survey instruments before using data.	
	Common method variance	Design research to avoid common method bias. Test for common method bias. Potential tests depending on data are: Harman's single factor test, partial correlation procedure, confirmatory factor analysis CFA) of multiple traits, multiple methods model (MTMM).	Triangulate the identification of common method variance with post-survey interviews.

External validity is the degree to which methodology and unique quality limit the ability to generalize the results beyond the initial sample. Factors limiting generalizability are the sample's unique characteristics or unique methods or procedures. Creswell (2009) suggested external validity threats to experiments are treatment interactions with participant selection, setting, and history. To resolve the external validity threat to generalizability due to the survey participants, quantitative survey researchers can only restrict their claims. Survey studies that request longitudinal data on leadership teams generalize the external generalizability to the specific group from within the organization. For example, longitudinal studies of TMTs within a group can generalize the results to the TMT from within the group. Interpretations based on triangulation with other TMT group studies of TMTs may help triangulate these results. Retests in different settings or times can help generalize the results.

Construct validity implies that the variable tested in the study represents the theoretical construct. The types of construct validity tests are face validity, content validity, predictive validity, concurrent validity, discriminant validity, and convergent validity. Girden and Kabacoff (2010) suggested that construct validity is the most difficult to prove. Face validity means that the measure's content in the study reflects the theoretical construct. Content validity implies that the measure's content is linked to the theoretical and actual content that the instrument seeks to measure. Threats to survey face validity are that the participants will not understand what the survey asks about or link it to the actual content. A resolution to these threats is to question some participants in a post-survey interview, and these post-survey interviews can be quantitatively analyzed. Concurrent triangulation of the survey results with the post-survey interviews allows the researcher to identify errors in the survey and factor these errors into the quantitative and qualitative results of the survey. Predictive validity implies that the measure can predict future behavior. Concurrent, discriminant, and convergent validity consider how accurately an instrument measures a construct. Concurrent validity considers how the measure of a construct in a scale is related to other measures of the same construct, and discriminant validity indicates how that measure is different from previous measures. Finally, convergent validity indicates that a new scale is closely related to other scales measuring the same construct. A researcher can address these four construct threats (predictive, concurrent, discriminant, and convergent) through the design of the research process. Researchers can present threats to predictive validity with research designs that use a causal theoretical model to link the predictor's construct and the criterion construct. A research design that concurrently uses an existing instrument and a new instrument can determine if current, discriminant, or convergent validity threats exist. For example, if a new survey instrument for solidarity predicts the criterion better than the existing instrument for OCB-generalized compliance, solidarity has predictive, concurrent, discriminant, and convergent validity.

Statistical conclusion validity in a study means the study used appropriate statistical methods for the sample size, desired statistical power, and assumptions about the data. Threats to statistical conclusion validity arise from statistical tests run on unsuitable data or fishing for results without regard to theory or methods. Unsuitable data has insufficient power for the statistical test, violate the test's statistical assumption, or come from an unreliable instrument or faulty methods (Cabanda et al., 2011; Girden & Kabacoff, 2010; Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). Girden and Kabacoff (2010) indicated the power of a statistical test depended on having a sample size large enough to detect the significant effect, but not too large, which made any effect significant. Survey studies that target small populations, such as members of TMTs, may need to change statistical methods or simplify models to get enough observations for the study.

In addition, statistical methods depend on certain assumptions and circumstances to be valid. Neglecting to check if the research data violates those assumptions or fishing for results randomly using statistical methods without checking theoretical assumptions threatens the validity of the statistical conclusion. For example, multiple regression depends on independent variables without multicollinearity, singularity, and outliers that have a relationship with the dependent variable that is linear, normally distributive, and had homoscedasticity of error residuals (Hair, Black, Babin, & Anderson, 2010; Pallant, 2010). The research analysis method must check for these characteristics in the data to be suitable for the multiple regression analysis. Another example of unsuitable data in a statistical test occurred in the mixed-mode analysis of IPA theme counts. Qualitative IPA analysis used survey encoding methods to provide theme counts for the same constructs as survey questions. Scale reliability tests on these theme counts did not equate to scale reliability tests run on survey responses. The reliability tests for the qualitative IPA encoding considered qualitative measures (theme grids or weighted node diagrams). An unreliable survey instrument or unreliable methods for IPA analysis and generation of theme counts would threaten statistical reliability. Mindful triangulating the quantitative results from survey data with quantitative results from the IPA analysis would provide insights into the reliability of the data. Exploratory researchers should use triangulation between the survey results with interviewee discussions of the reliability of the data and the survey results to determine whether the data are reliable.

Podsakoff et al. (2003) defined common method variance as variance attributed to a “measurement method than the constructs the measures represent” (p. 879). Common method bias is a term used to discuss the potential sources of common method variance in a research methodology. Common method bias can apply to quantitative, qualitative, and mixed-mode. Podsakoff et al. (2003) defined the major categories of common method bias as common rater effects, item characteristic affects, item context effects, and measurement context effects. The common rater bias can occur in surveys where a participant rates both the predictor and the criterion variable. Qualitative IPA analysis can also have common rater bias when a single rater codes most of the data. Common rate bias occurs because raters desire to be consistent, socially acceptable, and aligned with their implicit theories, acquiescent biases, and leniency tendencies. The rater’s transient mood when taking the survey may alter these biases. The form of the survey may also bias the survey participant’s responses. For example, the survey respondent can react to the wording of the question (item characteristic effects), the order of the questions (item context effects), or the appearance of the survey (measurement context). Any of these common rater biases can cause common method variance in the results, which becomes a threat to statistical conclusion validity. The statistical methodology assumed that the predictors significantly influenced the criterion when control variables were constant. The statistical results could have flaws if this assumption was untrue due to common method variance.

Maxwell (1992) suggested that qualitative experts have debated the concept and use of qualitative validity. Maxwell included descriptive validity, interpretive validity, theoretical validity, generalizability, and evaluative validity in qualitative validity. Creswell (2009) suggested qualitative validity meant the following:

obtaining qualitative data from different sources, validating that transcripts and documents are error-free, spending prolonged times in the field to understand context, using detailed thick descriptions, using peer debriefing of coding, and letting participants and an external auditor check coding. These two definitions are similar for descriptive and interpretative validity. Maxwell's (1992) descriptive validity links to Creswell’s (2009) obtaining an accurate account of the phenomena from different sources and checking it for errors. Threats to descriptive qualitative validity are inaccurate accounts by an observer or errors in transcription. Research designs can mitigate these descriptive validity threats by gathering additional information. For example, information gathered from pre-survey explanation sessions or post-survey reviews can help determine if online reports are accurate.

Qualitative interpretative validity is the concept of the first cycle of coding. This cycle starts capturing the phenomena from the participant's view by spending time in the field to understand context, coding with detailed thick descriptions, using peers to debrief the coding methods, and letting participants and an external auditor check the code. Threats to interpretative validity come from the research’s lack of understanding the concepts of the environment or coding the first cycle of coding based on the researcher's bias rather than the participant’s viewpoint. Adopting Creswell’s (2009) suggestion mitigated interpretative validity threats. Creswell (2009) suggested spending time in fieldwork learning the concepts and working with peers, participants, and auditors to check the first coding cycles.

Qualitative theoretical validity, generalizability, and evaluative validity examine how well qualitative research data and theory work to validate each other. Qualitative theoretical validity comes from validating a theory's phenomenon via research data. Threats to qualitative theoretical validity have been found when there is a lack of valid agreed-upon theory accompanied by a method to encode or the researchers fail to report negative results. Researchers mitigated theoretical validity by using agreed-upon theories, describing encoding methodologies in detail, and reporting hypotheses proven or rejected clearly. The generalizability of a theory comes if a theory was applied in different situations by purposefully sampling, so it could be generalized to other groups within the community or outside. Threats to generalizability have occurred when there is a lack of purposeful qualitative sampling or the sample used did not provide sufficient material to understand the phenomena and test ideas about the context of the phenomena. For example, during interviews, it was essential to understand how the nature of the relationship between the interviewer and informant helped the interviewer ask the questions so that it is clear the interviewer captured all of the informant's perspective rather than just a portion.

Evaluative validity occurs when the theory and the work can be placed in an evaluative framework as objects to study. The evaluative framework should support both qualitative and quantitative results. For example, the IETF (2019) has an agreed-upon framework for evaluating the group output of the IESG cohorts as TMTs. Each year's IESG by number of RFCs published, the number of workings groups created, managed, or closed, and the number of IETF actions. Placing the qualitative results of what made an effective IESG consensus decision-making process into that evaluative framework aided the acceptance of the work by the community.

Table 6: Resolution to Qualitative Validity Threats Using Concurrent Triangulation

Type of validity	Threat to qualitative survey research	Qualitative methods to resolve	Concurrent triangulation resolution to threats
Descriptive validity	Inaccurate factual reports of the account (physical and behavioral)	Check account with participants or get approved public documents.	Gather data from survey participants in pre-survey explanation sessions or post-survey interviews on public documents.
	Inaccurate transcription of results	Check account with transcription tape and participants	Triangulate results from transcript-based reports with other qualitative documents and surveys.
Interpretive validity	The researcher lacks an understanding environment's language and concepts	Spend time in the field	In developing the survey, spend time to understand environments, language, and concepts
	Phenomena coded from researcher's bias rather than the perspective of participants	Review first-level coding with auditors, participants, or peers who understand environment and phenomena.	Review questions on first level coding with survey participants
Theoretical validity	Lack of valid theory agreed upon by the community and a method to apply it to data.	Early discussion with peers to develop theory and methods to apply	Triangulation of theory by two methods can help validate the theory, the phenomena, and the methods.
	Lack of presentation of material that is negative to themes	Present both negative and positive material.	Triangulation can help put negative and positive data about theory in a larger context.
Generalizability validity	Lack of purposeful sampling of data to able to understand the phenomena and test ideas about phenomena Interviewers lack of understanding informant's context and their relationship	Careful preparation work before sampling or before interviews to understand phenomena and participants	Triangulation with previous qualitative or quantitative work can provide the researcher with a better understanding of phenomena or the interviewees.
	Lack of generalizability of phenomena	No resolution	Confirm the lack of generalizability by looking at other results.
Evaluative Validity	Lack of an evaluative framework to place theory and reason to attach theory to it.	Develop an evaluative framework with theory.	Triangulate the evaluative framework with qualitative research, which may aid acceptance of the evaluative framework.

Using Concurrent Triangulation to Improve Reliability

The reliability of a study implies repeatability or consistency. Quantitative reliability can be test-retest reliability, internal consistency, or interrater reliability (Cabanda et al., 2011). Creswell (2009) stated that qualitative reliability indicated that validity methods are the same across projects and different researchers. One measure of this validity was interrater scores that compared the rating of multiple raters who could be the primary researchers, researchers' peers, auditors, or participants. Concurrent triangulation mixed-mode design allowed the qualitative and quantitative methods that complemented each other to improve the reliability of the mixed-mode research. This section examines how triangulation resolves threats to quantitative and qualitative reliability.

Quantitative methods and qualitative reliability used statistical analysis to detect a lack of reliability. Test-retest reliability of a quantitative survey indicated that the survey used to retest participants would have the same results. The Pearson product-moment correlation coefficient test results indicated the correlation between the results of the two surveys. Quantitative survey research utilizing research methods with solid validity was likely to have a strong correlation (+1 or -1) between the two scores. Threats to test-rest reliability were poor validity of the survey instrument, unknown changes to the instrument or context, or reactivity of the subject. Any resolution of these threats to test-retest reliability depended on the threat, but careful post-survey interview questions helped the researcher triangulate the reasons for the problem. Internal consistency was measured either by split-half reliability or Cronbach's alpha. The split-half reliability correlated half of the survey results based on half of the items for a construct. A Cronbach's alpha based its score on all items' interterm correlation coefficients. Threats to reliability came from unreliable instruments, so it was crucial to use instruments pre-tested for the population sampled. Triangulation with post-survey interviews helped the researcher understand the problems with the survey.

Quantitative surveys measured interrater reliability by measuring how two raters judged behavior by selecting the same value on a survey. Qualitative surveys used interrater reliability to measure how two researchers coded the reliability.

Cohen's Kappa was the statistical test to detect this unreliability. Post-survey interviews with two raters who took the survey explained why they rated the behaviors differently. Post-qualitative coding analysis discussed the items where there was low interrater reliability

Table 7: Resolution to Qualitative and Quantitative Threats to validity

Type of reliability	Threat to quantitative survey research	Methods to resolve the threat	Concurrent triangulation resolution to threats
Quantitative test-retest reliability	Poor quantitative validity in research Unknown changes to test Unknown participant reactivity	Detect unreliability by checking if test and retest scores are correlated. Resolve the threat of poor validity. Use statistical methods to check for unknown changes.	Use post-survey interview questions to inquire about the survey's methodology, participants' understanding of other latent factors, or problems with survey questions. Triangulate these answers with test reliability scores and seek to understand the causes of the reliability score.
Quantitative internal consistency	Poor reliability of measurement instrument.	Detect by the split-half reliability test or Cronbach's alpha test for scores based on items.	
Quantitative interrater reliability	Two survey participants have a similar behavior	Detect the interrater reliability error can using Cohen's Alpha.	Use careful post-survey interviews to help the researcher understand why the two raters were similar.
Qualitative low interrater reliability of coding	Bias of rater Rater does not understand phenomena or context	Plan for cycles of interrater discussion in methodology (discussion between raters, re-coding of documents, and a retest of interrater reliability.)	Triangulate viewpoints of the rater during the discussion between raters.

Methodology for Qualitative and Quantitative Research Into Groups and Teams

Hollingshead and Poole (2012) emphasized that research into groups requires precise control and measurement of behaviors, an attitude of realism to observe the behavior where it occurs, an ability to generalize the methodology to allow transfer to different populations, and a research method based on theory. Pratt and Kim (2012) suggested that group research investigates phenomena at the interaction of multiple theories so that qualitative methodologies, such as ethnography or netnography, might help discover the undiscovered intersection of these theories. Netnography is the study of groups by observing electronic bulletin boards and chat rooms. Hinds and Cramton (2012) suggested that global groups should use electronic bulletin boards, chat rooms, and computer-aided meetings. These computer-aided mediums for groups allow the research to observe a global team meeting, but it does not remove the requirement to understand the different cultures and biological rhythms of people in a meeting. For example, in a teleconference of the current IESG of the IETF, the members attended from their homes or businesses in Europe, North America, China, Australia, and other locations around the world. Transcriptions of these computer-aided meetings should be sent to global group members and agreed upon prior to considering them valid. Meyers and Seibold (2012) suggested that a qualitative analysis of a series of group meetings should inquire whether the groups argue or share information and how this communication occurs. Researchers should check if group members exchange messages in a group-based symbolic language. Another factor is whether the group is aware of how communication impacts (or fails to impact) group outcomes. TMTs may be very aware of the process of communication. Creating a coding schema for a group meeting requires the standard steps in qualitative research. These steps include examining the material, developing a tentative schema, and refining the schema in exploratory coding of the data. The coding process involves cycles to advance from descriptive validity through interpretive validity, theoretical validity, and generalizability to an evaluative framework.

Kashy and Hagiwara (2012) suggested that analyzing group data using quantitative techniques may require analysis of individual and group scores with multilevel modeling to determine if the variance in scores was between groups, within groups, or mixed between and within. If group members have roles, these roles may form internal dyads or subgroups. Multilevel modeling allows the encoding of individuals, dyads, and leaders in the model that utilizes social relationship concepts. For example, dyads can be coded as actor-partners within a group, while leader-group relationships require a one-with-many encoding. Villa et al. (2003) noted that “leadership researchers have long assumed that effectiveness of leaders is dependent on situational factors” (p. 3). These situational factors interact with leaders of groups or leadership teams, such as TMTs. However, leadership studies have had problems finding these moderators due to decreased statistical power to detect moderator "interactions in small sample sizes" or "unequal group sizes" (Villa et al., 2003, p. 5). In addition, the predictor variables in leadership research studies can interact with situational factors. Villa et al. (2003) suggested restricting moderator tests to theory-based moderators and testing predictor variables for multicollinearities. Villa et al. (2003) noted that some research had detected theoretically based moderators, which were statistically significant, by restricting the regression equation to one leadership behavior and a single moderator (p. 10).

Mixed-Mode Leadership Research (1995 to 2011)

Researchers have adapted mixed-mode methodologies to general research or research in specific fields such as leadership, nursing, health, education, and public health. Stentz et al. (2012) examined the 55 articles denoted as mixed-mode studies in *The Leadership Quarterly* journal from 1995 to 2012 and found only 15 had the following four elements, which Creswell and Plano-Clark (2011) defined as central to the mixed-mode study: "(1) extent of interaction, (2) relative priority, (3) timing, (4) where and how" (p. 1175) the qualitative and quantitative elements are mixed. Stentz et al. (2012) expanded on these four definitions to allow additional classification of these 15 articles. Four of these mixed-mode articles involved leadership theories related to teams (Amabile, Schatzela, Moneta, & Kramer, 2004; Carmeli & Schaubroeck, 2006; Morgeson & DeRue, 2006; Rowland & Parry, 2009). Carmeli and Schaubroeck (2006) used mixed-mode research to examine how group processes impacted the decision-making of TMTs leading companies during periods of organizational decline. This section evaluates the mixed-mode methodology used by these four leadership studies related to teams and leadership.

Stentz et al.'s (2012) expansion on the central components of mixed-mode examine how quantitative and qualitative components are mixed in the research design, data collection, data analysis, and interpretation. This discussion denotes a single qualitative or quantitative component as a "strand" for clarity. Each quantitative or qualitative study strand poses a question, collects data, analyzes data, and interprets the results. Timing describes the order of execution of the qualitative and quantitative strands in the study. The timing of a study can be sequential such as qualitative strand before quantitative strand or concurrent. Timing can also be multiphase timing if the study includes quantitative and qualitative strands in multiple phases of a research program. Stentz et al. (2012) defined mixing as “when and how to integrate or combine the two types of data” (p. 1175). Stentz et al. (2021) defined the point when integration occurred as the “point of interface” (p. 1175). This point of interface can happen during any of the four steps in research: research design, data collection, data analysis, and interpretation. Mixing strands can entail (a) merging the qualitative and quantitative datasets, (b) connecting analysis of one dataset (quantitative or qualitative) to the collection of a second dataset, (c) embedding quantitative or qualitative components in a larger design, and (d) using a theoretical framework to integrate two different types of data. Mixing can be within larger research strands at any interface point or after two strands (qualitative and quantitative) finish, and the theoretical interpretation combines both strands.

Each of the four mixed-mode research studies identified by Stentz et al. (2012) had a unique methodology. Amabile et al. (2004) utilized an exploratory concurrent research design based on the following theories on organizational creativity’s components (Amabile, 1997), leadership, and LMX. This exploratory research was one part of a multiphase research project, but for this part, it collected data from 139 subordinates on 26 teams. The data collected from all participants included two questionnaires: a daily and a monthly questionnaire. The daily questionnaire for subordinates contains a qualitative ($QUAL_1$) daily diary of work activities in teams and a leadership support instrument ($quan_1$).

The monthly questionnaire had a peer creativity scale ($quan_2$). The quantitative data for these two scales were analyzed using descriptive statistics and scale reliability tests ($quan_{A1}$ and $quan_{A2}$). The qualitative data from the daily responses ($QUAL_1$) was subject to the following three qualitative analyses: (a) a content analysis for behaviors of leaders, (b) a reflective reading analysis investigating the links between a leader’s behavior and subordinate’s perceptions, and (c) an in-depth analysis of extreme behaviors on two teams. The first qualitative analysis ($QUAL_1$) uses content analysis to thematically code the open-ended questions on the daily questionnaire for leadership behaviors and frequency of behaviors. The thematic encoding for the leadership behaviors used categories aligned with Yukl, Wall, and Lepsinger's (1990) Multiple-Practices Survey (MPS) used on the quantitative questions on the daily questionnaire. All leadership behaviors detected by this thematic coding were filtered to remove behaviors that did not link to leadership support or had low daily frequency. This qualitative analysis resulted in eight MPS categories of behavior ($QUAL_{A1}$). The researchers triangulated the qualitative result with the quantitative analyses of leadership support ($quan_{A1}$) and peer creativity ($quan_{A2}$). Recasting this triangulation the notation, this triangulation is $QUAL_{A1}$ [$quan_{A1}$] → +[$quan_{A2}$]. The second qualitative analysis ($QUAL_{A2}$) used a method of reflective reading with iterative codings to find patterns and links between the leadership behaviors and the quantitative results for each of 26 teams ($QUAL_{A2}$[$Qual_{A1,}$ $quan_{A1,}$ $quan_{A2}$]). The third quantitative analysis ($QUAL_{A3}$) did an in-depth study of 2 teams with extreme scores (one positive and one negative) on leader-subordinate dynamics. The in-depth study involved collecting additional background material on the team's company, projects, and people for these two teams ($qual_3$). Two researchers analyzed the in-depth data on these two teams based on the theoretical model ($QUAL_{A3}$ [$qual_{A1}$, $qual_3$,$quan_1$, $quan_{2,}$ $QUAL_{A1}$, $QUAL_{A2}$]). The interpretation of results considered the two quantitative analyses [$quan_{A1}$and $quan_{A2}$] and the three qualitative analyses ($QUAL_{A1}$, $QUAL_{A3}$, and $QUAL_{A3}$). The notation shown in Table 8 provides a short-hand for this analysis of the intermix of quantitative data and methods and timing (where "→" indicates a sequential timing). This notation showed the nuances in the mixed-mode method design from Amabile et al. (2004).

Carmeli and Schaubroeck (2006) used an explanatory sequential mixed-mode design based on an explicit theory on TMT behavioral integration relationship to organizational decline. The theoretical model posits that the perceived quality of strategic decisions moderates the impact of TMT behavioral integration on organizational decline. The explanatory design used two sequential phases with quantitative survey research in Phase 1 ($QUAN_1$) combined with a qualitative case study in Phase 2 ($qual_1$). The quantitative study collected data from

116 firms from a variety of industries with TMTs (out of 217 firms queried) and analyzed these data ($QUAN_{A1}$) using descriptive and multivariate statistics (correlations and hierarchical regression). The research team concluded that the perceived effectiveness of strategic decisions did not moderate organizational decline/success based on the quantitative data. The qualitative case study examined four out of the 116 firms who responded to the survey on TMTs. The qualitative case study collected market analysis write-ups and did interviews of TMT members. The case study in-depth analysis examined the qualitative reports regarding the survey questions and other data regarding each TMT within their companies.

Table 8: Use of Mixed Modes With Teams

Study	Prototypical design type/ timing	Extent of interaction Type Interaction Priority Point(s)	Research theory type and topics
Amabile, Schatzela, Moneta, and Kramer (2004)	Exploratory Research Part of multiple phase study with concurrent phases.	Data collection $QUAL_1$ + [$quan_1$] (daily) [$quan_2$] (monthly) [$qual_3$] (in-depth) Data analysis $QUAL_{A1}$ [$quan_{A1}$] → +[$quan_{A2}$]→ $QUAL_{A2}$ [$QUAL_{A1}$,$quan_{A1,}$ $quan_{A2}$]→ $QUAL_{A3}$ [$qual_{A1}$, $qual_3$,$quan_1$, $quan_{2,}$ $QUAL_{A1}$, $QUAL_{A2}$] Data interpretation [$quan_{A1}$],+[$quan_{A2}$] + $QUAL_{A1}$ $QUAL_{A2}$ + $QUAL_{A3}$	Explicit theories: Organizational creativity (componential theory) (Amabile, 1997), leadership behavioral theory, LMX
Carmeli and Schaubroeck (2006)	Explanatory / sequential	(survey) Data collection Data analysis Data interpretation $QUAN_1$ → $QUAN_{A1}$ → $QUAN_{Int1}$ → $qual_2$ (case study) Data collection $qual_{A2}$ Data analysis Data interpretation $qual_{Int2}$ → Triangulate $QUAN_{Int1}$ + $qual_{Int2}$	Explicit theory: TMT behavioral integration for decision-making (Hambrik, Cho, & Chen, 1996)

Study	Prototypical design type/ timing	Extent of interaction Type Interaction Priority Point(s)	Research theory type and topics
Morgeson and DeRue (2006)	Exploratory / Sequential, Part of a multiple phase research program	Interviews→ survey (Interviews [critical events, disruption] $QUAL_1$→ $QUAL_{A1}$→Survey Survey [critical events, criticality, urgency disruption, leader's time] parallel Data collection $quan_1$ (team-members) $quan_2$ (leaders) Data analysis $quan_{A1-2}$, ($quan_{1,}$ $quan_2$) Joint data interpretation $QUAL_{Int1}$ + $quan_{Int}$ ($QUAN_{A1}$)	Explicit theory: Leading semi-autonomous teams (Manz & Sims, 1987)
Rowland and Parry (2009)	Multiphase with /sequential timing Multilevel research	Phase 1: Grounded theory $QUAL_1$ data collection (12 meetings + interviews) $QUAL_{A1}$ data analysis Phase 2: Survey $QUAN_2$ data collection (one-time survey) $QUAN_{A2}$ data analysis Phase 3: Triangulation $QUAL_3$ [$QUAL_{A1,}$ $QUAN_{A2}$] Triangulation of results informs patterning per leader Joint interpretation	Explicit foundation with an implicit investigation Explicit theory Organizational design as a context that impacts: a) effective team decision-making, and b) leader's influence on decision-making via transformational or transactional styles.

Note. $QUAL_1$ – data collection for qualitative Strand-1,
$QUAN_2$ – data collection for quantitative Strand-2,
$QUAL_{A1}$- Analysis for Strand-1,
$QUAN_{Int1}$ – Interpretation for Strand-1

Morgeson and DeRue (2006) used an exploratory sequential design based on an explicit theory that leader intervention in critical events improved the effectiveness of semi-autonomous teams (Manz & Sims, 1987). This study involved 293 team members on 42 teams in 4 organizations. The qualitative data ($QUAL_1$) was collected in one-on-one interviews with the 42 team leaders and analyzed ($QUAL_{A1}$) using content analysis to determine the critical events in 42 teams($QUAL_{Int1}$). After determining the critical events, the researchers administered two quantitative surveys ($quan_1$, $quan_2$). One survey queried all team members per team, asking them to characterize each critical event in criticality, urgency, duration, disruption to their team, and intervention actions ($quan_1$). The second quantitative survey queried each team leader regarding how much each critical event disrupted their team and the time it took the leader to resolve the critical event ($quan_2$). The survey's data was analyzed using scale reliability tests descriptive and multivariate statistics (correlation, Hierarchical linear modeling; $quan_{A1-2}$). The data interpretation considered the results of both quantitative and qualitative data.

Rowland and Parry (2009) used a multiphase mixed-mode research design with sequential timing to explore the influence of meso-level organizational design on micro-level leadership behaviors and macrolevel (team) team decision-making that influences organizational outcomes. The theoretical underpinnings of this research were the situational leadership theory from House's (1971) path-goal theory and Fielder's (1967) LPC contingency model. These theories suggest context impacts leadership, effective decision making, effective team decision-making, and the leader's influence on decision-making via transformational or transactional styles. Their multiphase research design included three sequential phases. Phase 1 was grounded-theory qualitative ($QUAL_1$) data collection on four TMTs. The qualitative strand ($QUAL_1$) collected data on teams via observation of 12 team meetings and one-one formal and semi-structured interviews of team members outside these meetings. These observations and interviews were theoretically encoded, and progressive analytical insights on the patterns were recorded in analytical notes by the researcher. Subsequently, the researchers compiled these analytical memos into (a) a matrix and (b) the structure of a hierarchical model. Limitations of interrater reliability occurred because the researcher only encoded the observations in words and did not encode actions beyond words (e.g., tone or facial expressions) in the meeting transcripts. Due to this finding, only the researchers who observed the meetings encoded the behaviors. Validation of the results was checked by having the team members validate the results of the analysis of the groups. Phase 2 in this research was a quantitative survey ($QUAN_2$). Rowland and Parry's (2009) quantitative strand used a leadership style survey to determine the follower's perception of each team's leader's style (QUAN2). The third phase of this research generated qualitative data ($QUAL_3$) by triangulating the qualitative and quantitative analysis to find patterns of recurrent behaviors exhibited by each team. The triangulation and patterning techniques used an iterative process to find and validate findings. The final step in this research design was to provide joint interpretation.

Rowland and Parry's (2009) triangulation process was of particular interest to this research because it dealt with whether consensus or strategic decision-making was appropriate for modern organizations. This current researcher considered Dooley and Fryxell's (1999) thesis that the consensus decision-making process that allows dissent during consensus but expects consensus and solidarity afterward is outdated. Dooley and Fryxell (1999) suggested that a strategic decision-making process should replace consensus decision-making. Strategic decision-making manages the conflict in groups based on the strategy of its leader. Rowland and Parry (2009) found in the triangulation (Phase 3) that the emergence of a "construct of consensual commitment," which identified a "phenomena of genuine consensus and an intention by team members to follow through" (p. 545) was common in effective TMTs. Organizational design as "reflected in organizational structure and job design exerted a meso-level moderating" influenced this consensual commitment, along with the leader's relational leadership style and support for dissent. Rowland and Parry (2009) also found that a TMT leader's leadership style influenced decision-making behaviors and outcomes of the TMT. Rowland and Parry requested more empirical studies to support these conclusions.

This research answered Rowland and Parry's call. In the next section, this research's Phase 1 and Phase 2 research. The Phase 1 exploratory mixed-mode research in 2012 examined the impact of a leader's style in a TMT utilizing consensus decision-making. The Phase 2 explanatory mixed-mode research considered the antecedents to consensus decision-making that lead to consensual commitment.

Internet Engineering Steering Group Specific Research

The IETF is one ICT SDO studied by several researchers, but few have examined the IETF process and leadership in their research. For example, Gençer (2012) noted that most research on ICT standards examined how ICT standardized technologies diffused into products, but few studied the process. Previous research into the IETF process included studies by Chiao, Lerner, and Tirole (2005), Nickerson and Muehlen (2006), Russell (2006), Rysman and Simcoe (2008, 2010), Simcoe (2007, 2012), and Simcoe and Waguespack (2011) who studied the processes that created IETF standards. Chiao et al. (2005) and Rysman and Simcoe (2008, 2010) examined how patent policy and patents interacted with IETF standards. Nickerson and Muehlen (2006) examined the IETF standards as an ecology where standards were born, grew to maturity, and died. Russell (2006) examined the development and spread of IETF technology as the spread of people and ideas from a discussion group Vint Cerf led at ARPA on IETF technology. Russell (2006) noted that this initial group focused on “running code and rough consensus” (p. 49) instead of the multinational standards in ISO standards. Simcoe (2007) examined how the IETF’s paradigm of “rough consensus and running code” (Section 2.2, para. 3) required a process of consensus decision-making and a target of creating standards that would allow interoperable implementations of Internet standards. Rysman and Simcoe (2008, 2010) considered how patents interacted with IETF standards.

Gençer (2012) and Simcoe (2007) utilized these concepts of the “birth, maturity, and death” of standards to examine the IETF standards creation process within technology groups (denoted as WGs), then approved for publication as a standard. Gençer (2012), Simcoe and Waguespack (2011), Simcoe (2007, 2012) considered specific influences on the standardization process and utilized IETF archival data (mail lists, standards, statistics on creation of standards, and IESG approval statistics). Gençer (2012) examined the speed of the process for original and linked standards in this fashion and linked standards because the linked standard enhances the original standards. Simcoe and Waguespack (2011) examined how "high-status authors received more attention" than low-status authors on "electronic distribution boards" (email; p. 274). Simcoe (2007, 2012) focused on analyzing the archival data on WGs and standard productions to determine how the IETF consensus decision-making process worked to create standards and why the speed of standards creation slowed during 1992 to 2000. Simcoe's theoretical model for IETF standards creation can be used to consider technology, conflict in IETF WGs, patent issues but not the leadership of the IETF (2021a, 2021b).

Simcoe's Measures of Effective Internet Engineering Steering Group Leadership

Simcoe (2007, 2012) developed a theoretical model to estimate the time the IETF took to create a standard based on technology complexity (task issues), WG conflict, and delays in final approval by the IESG management team. Simcoe (2012) expressed this theoretical model as two equations (Simcoe-5 and Simcoe-7). First, the technology complexity varies between IETF standards that propose new protocols or best practices and IETF informational documents that explain these standards. Second, Simcoe (2007, 2012) examined WG conflict as arising from *distributional* conflict, which slowed down the WG consensus for business reasons rather than TC or RC. Simcoe (2012) proposed two measures for distribution conflict: suit-to-beard ratio (business concerns versus academic concerns for technology) and background-IPR. The delays in final approval ($X_{ijt}\theta$) handle the differences in "time-period and technology class" per technology proposal per WG per year (IESG cohort). A single AD handles the technology for a WG for a year, so the delays in final approval ($X_{ijt}\theta$) are influenced by the leadership of an AD (an IESG Member) and an IESG cohort led by an IETF chair. Simcoe's (2012) equations model the delays of the IESG as a TMT reviewing and approving a draft for publication as a constant. Simcoe and Waguespack (2011) based this theoretical model of "standard-setting committees" based on "the stochastic bargaining framework of Merlo and Wilson (1995)" (p. 305). Simcoe's (2012) model predicts delays in standards creation grow longer when distributional conflicts cause standard participants to favor specific technologies due to lead times for complex.

Simcoe (2007) started the research into the process of IETF standard's development, noticing a "significant slowdown" in the pace of standards creations during 1992 to 2000, where the "median time from the first draft to RFC doubled from the 1986 to 1992 rate, "growing from seven to fifteen months" (Introduction, para. 4). Simcoe posited how IETF leadership in the IESG and the WGs (chairs) caused this slowdown. Since 2000, the IETF (2016b) tracked statistics on its productivity regarding the IETF as SDO producing standards. These statistics include the number of standards published per year, the time it takes new work to go from the first draft to a published standard, and the average IESG review period to approving documents. After evaluating Simcoe's theoretical model, this section evaluates Simcoe's (2012) research methods (data collection and analysis) before discussing his results.

Equations in theoretical model. Simcoe's (2012) theoretical model can be expressed as the following equations:

Simcoe-5) "$T_i = I_i \alpha + D_{i j} \beta + D_{i j} I_i \tau + X_{i j t} \theta + \varepsilon_i$" (p. 317).

Simcoe-7) "$T_i = \lambda_k + I_i \alpha_k + D_j \beta_k + D_j I_i \tau_k + X_{i j t} \theta + \varepsilon_i$" (p. 328).

The variables for equation Simcoe-5 are the following:

T_i is the predicted "time-to-consensus" standard *i*,

I_i is the RFC status vector with (standard = 1 or nonstandard = 0) for standard *i*.

α is the "mean difference in time-to-consensus" of standard due to complexity,

$D_{i j}$ is the "distribution conflict" for proposal *i* in Working Group *j*,"

β "measures the correlation between D_{ij} and routine publication delays,"

τ "measures impact of distribution conflict delayed agreements,"

$X_{i j t}$ is a vector to handle differences in time-period and technology class (e.g., year and IETF Area) along with constant q,

θ is constant that adjusts real-time differences $X_{i j t}$, and

ε_{i} - is an error factor. (Simcoe, 2012, p. 317)

Simcoe's (2007) prediction from equation Simcoe-5 is that the time to reach a consensus (T_i) will be the sum of mean-time for standard with same complexity ($I_i\,\alpha$) plus distribution conflict ($D_j\,\beta_k$+$D_j I_i \tau_k$) plus time-delay for technology per time-period ($X_{i\,j\,t}\,\theta$) plus an error factor per standard. Simcoe (2007) modeled the IESG as a fixed delay with bottlenecks that may be caused by "high-level coordination failures or development of bottlenecks in the publication process" (Section 3.3, para. 1). The equation ($X_{i\,j\,t}\,\theta$) indicates this delay of an IESG cohort ($X_{i\,j\,t}$) for proposal i from working j at time t (IESG cohort) times a constant. Simcoe's (2007) conflict equation is complex but only considers the distribution conflict. As authors try to get WG consensus on a proposed standard approved in a WG, distribution conflict is purposeful conflict. This focus does not consider TC or RC in the WG.

Simcoe (2012) stated that the "difference-in-difference framework" in this equation Simcoe-5 produced unbiased results "I_i is exogenous, or τ is constant" (p. 317). Gençer (2012) confirmed that Simcoe's (2012) complexity component α was different for standard documents (Simcoe's $I_i\,\alpha$ component) through this examination of the cross-RFC linkages in IETF standards via a social networking model. However, the IETF assignment of drafts to areas or WGs is not random, so the assignment to a WG can create a bias in Simcoe's (2012) equation. Simcoe (2012) suggested this bias could be caused by: the technical topic ($X_{i\,j\,t}\,\theta$), the authors who steered technical proposals to a particular WG with less distribution conflict ($D_{i\,j}$) or informational standards ($I_i = 0$) rather than protocol standards to reduce the distribution conflict constant (τ). Simcoe (2012) dubbed such authors as "rent-seeking" (para. 4) authors who sought lower distribution conflict in a WG. Simcoe's (2012) results also confirm that the standards/non-standards selection is exogenous. Based on these conclusions, Simcoe revised this equation to the form in equation Simcoe-7 to address the delay components due to the heterogeneity in standard processing and distribution conflict.

Equation Simcoe-7 addresses the heterogeneity concerns by creating vectors of constants for delays for general delay (λ_k), time-to-consensus delay (α_k), distribution-related delays (β_k and τ_k). Equation Simcoe-7 defines these new constants variable as follows

> λ_k is the constant delay per type with for the following types of heterogeneity: 1) origin of mail (dot.org, dot.edu), 2) IETF Technology class, and 3) time-period effects (cohort effect).
>
> α_k is the "mean difference in time-to-consensus" of standard due to the following types of heterogeneity: 1) the origin of mail (dot.org, dot.edu), 2) IETF technology class, and 3) time-period effects (cohort effect).
>
> β_k is the constant that adjusts distribution conflict D_j to routine publication delays
>
> per heterogeneous type
>
> τ_k "measures impact of distribution conflict delayed agreements" per heterogeneous types. (Simcoe, 2012, p. 328)

The equation Simcoe-7 replaces the general delay per IETF type ($I_i \alpha$) with the delay due to the heterogeneity in the mixture of documents ($\lambda_k + I_i\alpha_k$). Simcoe (2012) expanded the distribution delay from $D_{i\,j}\,\beta + D_{i\,j}\,I_i\,\tau$ (Equation Simcoe-5) to $D_j\,\beta_k + D_j\,I_i\,\tau_k$ (Equation Simcoe-7). This modification changes the distribution conflict term from distribution conflict for each Internet-Draft ($D_{i\,j}$) to a distribution conflict per WG (D_j) for proposals in the WG.

Simcoe (2012) examined two ways to measure distributional conflict per WG: (a) business funding of participants and (b) the use of patents by businesses to differentiate products. Simcoe (2012) used a suit-to-beard ratio indicates the entities sponsoring participants in a WG as either business (suits) or beards (academics). Simcoe and Waguespack (2011) classified suit-share as "the percentage of all email domains" from an email address ending in ".com" or ".net" (p. 319) from which participants send email to IETF WG for one year prior to the initial publication of an Internet draft in that WG. Beards-share is the "percentage of all email domains" which does not end in ".com" or ".net" from which participants send emails to the IETF WG in the same period. The concept was that business versus non-business economic focuses would create conflict between WG members. Restating the suit-to-beard ratio and suit-share ratio as equations produces the following:

D_j = suit-to-beard$_j$ = ebiz$_j$/e-acad$_j$

D_j = suit-share $_j$ = ebiz$_j$/e-all$_j$

where:

D_j is the "distribution conflict" for proposals in Working Group *j*,

suit-to-beard$_j$ - is the ratio of business participants to academic participants per working group *j*,

suit-share$_j$ - is the ratio of business participants to academic participants per working group *j*.

ebiz $_j$ is the number of email addresses which came from .com or .net add

e-acad $_j$ is the number of email addresses which from .edu or .org.

e-all $_j$ is the number of email addresses

The measurement of distribution conflict due to patents looks at two facets of intellectual property. The first facet is that the individual who has created the technology in a patent can assign the rights of the patent to a company in recompense for funds (job or monetary payment). An individual can reap additional rewards if their patent is included in an IETF standard, so an individual may have a propensity (PC-ipr$_j$) to include their patented technology in IETF standards. The second facet considers that companies use patents to establish and retain market share, so including a patent in an IETF standard may help the company gain or retain market share. As a result, several companies may compete to add their patented technology to an IETF standard, causing conflict (WG-ipr$_j$). Restating these background-IPR concepts as equation gives the following equations:

Pc-ipr$_j$ = $P_{IPR\text{-}5years*}$ γ [Individual Patent conflict weight]

WGc-ipr$_j$= $\Sigma_{c=1,m}$ ($\Sigma_{j=1,n}$ Pc-ipr$_{j,c)}$ [WG Patent conflict].

D_j = log(WGc-ipr$_j$)

Where:

Pc-ipr $_j$ is an individual's propensity for conflict in WG *j* on a standard.

Pc- ipr$_{j,c}$ is an individual's propensity for conflicts in WG *j* on a standard from company c.

$WG\text{-}ipr_j$ is the WG potential for a patent conflict.

$P_{IPR\text{-}5years}$ is a person's cumulative number of patents in the last five years in the US patent database (USPTO).

γ is an "uncentered correlation between assignee's patent portfolio and the cumulative patent portfolio of all IEF participants" (Simcoe, 2012, p. 41)

j = Working group j.

c = company c for all companies

m = maximum number of companies

n = maximum number of people per company with patents.

Simcoe (2012) proposed that several people pushing their IPR for a WG might increase distribution conflict across all documents. This theoretical model proposes that WG distribution conflict for a company arises from the total sum of each person's propensity for patent conflict ($Pc\text{-}ipr_j$) for all people with patents in a company ($\Sigma_{j=1,n} Pc\text{-}ipr_{j,c}$). A WG's distribution conflict is the log of the sum of the distribution conflict for all companies ($D_{j} = \log [WGc\text{-}ipr_j]$) with a proposal in the WG. Simcoe (2012) used the log() function to reduce the noise due to the skew of some companies with an extensive patent portfolio.

Data collection and analysis. Simcoe (2012) collected data on the cycle of publishing drafts from the IETF public website from January 1993 to June 2008. The researcher first selected Internet-Drafts (an IETF-standard document) submitted to any WG from January 1993 to December 2003. From this initial sample, the researcher detected 3,521 Internet-Drafts submitted to 249 IETF WGs, and 2,601 of these drafts had more than one revision. These 2,601 WG drafts with more than one revision were the focus sample examined in this research. Simcoe (2012) gathered the following things for each of these 2,601 WG drafts: (a) date of initial submission, (b) the drafts final state (state-standard RFC, nonstandard RFC, expired, or censored), (c) the date the draft reached the final state, and (d) a list of authors. The cumulative list of authors from all Internet-Drafts in the sample was used by Simcoe to collect data on patents per author in each WG from the U.S. patent office (USPTO). Next, Simcoe (2012) collected email addresses from each WG's open email lists from January 1992 to December 2003 to determine the participants' email addresses. Then, the email addresses were analyzed to determine if the addresses were from business (e.g., .com or .net) or nonbusiness addresses (e.g., .edu or .org).

Simcoe's (2012) data analysis methods used descriptive statistics to determine the time for publishing IETF RFCs per category of draft and Ordinary Least Squares (OLS) techniques to determine the impact of each type of the distribution conflict on those times. The descriptive statistics calculated the total number of RFCs published per year per category and the mean time between WG initial draft and publication of the RFC. Table 9 lists the descriptive statistics per year on Simcoe's (2012) sample for the drafts completing the cycle per category and the mean across all years. Table 9 provides a comparison with the number of documents the IESG approved for publication according to this research. Simcoe (2012) indicated the meantime for RFC publication in 1993 to 2003 was "774 days (2.1 years)" for standards, "595 days" (1.5 years) for non-standards, and "487 days for an expired proposal" (pp. 15–16).

Simcoe (2012) used OLS techniques to determine the link between distributional conflict and consensus-decision delays using Simcoe-7 as the model to fit the IETF standards data (drafts, standard RFCs, and non-standard RFCs). An increase in distributional conflict from an increase in suit share resulted in a linear increase in time delay, but an increase in IPR related conflict (log[Patents]) provided fluctuations (possibly curvilinear). The average suit-share (1993 to 2003) was 73.34 mean, with a standard deviation of 16.04, and the average IPR conflict was 7.61, with a standard deviation of 2.98. The OLS analysis of model fit for equation Simcoe-07 for the IETF publication data from January 1993 to December 2003 found significant correlations for correlation between dot.org (commercial) authors with high-suit-shares ratio and delay and rent-seeking authors. The correlation between distribution conflict (D_j) for dot.org (commercial) authors with high suit-shares ratio and standard-track delay for standards found additional delays 9.4 days for commercial author (dot.org), with correlations of 3.7 at $p < 0.05$ ($R^2 = 0.22$) and found additional delays of 10.5 days for a commercial author, with correlations of 6.1 ($p < 0.1$, $R^2 = 0.22$). Simcoe's (2012) results show that a marginal increase in distribution conflict results in an increased concession cost for the commercial author.

Simcoe's (2012) results show that the distribution conflict in WGs as IETF migrated from “a quasi-academic organization to a high-stakes forum for technical decisions” for commercial organizations between 1993 to 2003 caused “increased politics and a slowdown of consensus decision-making” (p.330). Simcoe’s (2012) model showed variances between commercial authors and noncommercial authors and rent-seeking of authors that found differences in WGs in different areas and years. However, Simcoe’s examination only resolved 19% to 21% of the variation in time delays for all RFCs and 21% to 22% of the variation in time delays for standards RFCs. Furthermore, this study did not consider delays caused by other types of WG conflict (task, relationship, or process) or the delays due to the IESG review of the Internet-Draft before publication.

Table 9: Simcoe (2012) Sample and Results Versus 10% Sample of This Research

Period	Sample (totals)		Simcoe's (20123) result							Simcoe 2012	10% sample estimate
	Drafts to RFCs	WG with RFCs	Observed types of drafts (totals)				Mean duration - T_i (days)				
			Standards	Non-standards	Expired	Censored	Standards	Non-standards	Expired	Total RFCs	RFC approved
1993	58	35	25	19	14	0	490	221	174	44	120
1994	86	41	43	19	24	0	514	484	419	62	80
1995	128	48	67	21	40	0	738	426	353	88	200
1996	167	61	87	39	41	0	751	535	600	126	340
1997	304	76	119	57	128	0	573	558	472	176	130
1998	246	73	78	55	113	0	448	652	447	133	240
1999	279	82	94	71	113	1	814	652	503	165	260
2000	326	79	143	55	125	3	930	757	469	180	90
2001	379	100	131	76	159	13	1002	602	598	207	120
2002	325	87	139	70	98	18	847	592	423	209	180
2003	303	85	127	71	78	27	703	514	591	198	180
Total	2,601	249	1,053	553	993	62	774[*1]	595[*2]	487[*3]	1606	1930
Range of yearly	-58 to 137	-13 to 21	-41 to 49	-16 to 21	-61 to 87	0 to 10	-225 to 366	-100 to 263	-188 to 307	-43 to 50	-210 to 140

Period	Sample (totals)		Simcoe's (20123) result							Simcoe 2012	10% sample estimate
	Drafts to RFCs	WG with RFCs	Observed types of drafts (totals)				Mean duration - T_i (days)				
			Standards	Non-standards	Expired	Censored	Standards	Non-standards	Expired	Total RFCs	RFC approved
change*6											
OLS increase							49 days	31 days			
Increase in suit-share							A linear increase in suit-share increases time delay	Fluctuations are not linear or simple curvilinear			
Increase in IPR conflict							Positively correlated to delays	Negative correlated to delays			

*1 – Simcoe's (2012) published number for average-time-to-consensus for standard RFC, but the calculated average is 719 days.

*2 – Simcoe's (2012) published number for average-time-to-consensus non-standard RFC, but the calculated average is 553 days.

*3 – Simcoe's (2012) published number for average–time–to–expiration standard, but the calculated time is 464 days.

Problems in applying Simcoe (2012) to an Internet Engineering Steering Group study. Simcoe's (2012) theoretical model does not consider the complexity of the conflict in a WG or the impact of IESG leadership on $X_{i\,j\,t}\,\theta$. These facets make it difficult to apply Simcoe's (2012) data or results to a study of consensus decision-making in the IESG. The conflicts within a WG that cause delays can arise from distributional conflicts between SDOs (e.g., ITU and IETF) or conflicts due to group interactions (TC, RC, or process conflict). Each IESG member serves as a TMT member and AD. An AD's job includes aiding WG groups and WG leaders (WG chairs) to handle all types of conflict in the WG. Simcoe (2012) examined the correlation between authors seeking lower distribution conflict (rent-seeking authors) by moving WGs and the delays caused by these actions based on two-year periods from 1993 to 2003. These 2-year periods also represent the efforts of two IESG cohorts' technical leadership. The technical areas in the IETF during 1993 to 2003 included Application, Transport, Internet, Routing, Security, and Operations and the Sub-IP Area.

Simcoe (2012) found a correlation between authors seeking WGs with low distribution conflict and reduced delays in WGs in only the Application and Transport areas. Simcoe (2012) suggested this correlation occurred because during 1993 to 2003, the lower level protocols (Routing, Internet) had established IETF protocols, but the higher layer protocol standards remained under development. The rent-seeking reduced the standards publication delays by 9.0 days for application WGs with a correlation of 2.3 at $p < 0.001$ and 5 days for Transport WGs ($p < 0.1$, $R^2 = 0.22$). The rent-seeking reduced delays for all IETF publications by 7.5 days for Application (correlation 3.4 at $p < 0.05$), 11.7 days for transport (correlation 3.2 $p < 0.001$, $R^2 = 0.21$). The correlation between rent-seeking authors for standard documents and specific IETF cohorts found significant correlation in reduced publication delays by 4.8 days for cohort (1995, 1996; correlation 2.5, $p < 0.10$, $R^2 = 0.21$), 6.4 days for cohort (1999, 2000; correlation 2.4, $p < 0.01$, $R^2 = 0.21$). The correlation between rent-seeking authors for all IETF documents and yearly cohorts found significant correlation in reduced publication delays in cohort (1999, 2000) by 7.5 days (correlation 3.4, $p < 0.05$, $R^2 = 0.19$) and cohort (2000, 2001) by 12.0 days (correlation 3.4, $p < 0.05$, $R^2 = 0.19$). Part of this variance could be due to leadership in the Transport where two IESG members provided continued leadership for the transport during that time, causing rent-seeking authors to favor WGs in that area.

Simcoe's (2007) understanding of the selection of members for the IESG is flawed, and his focus on the WG causes his research to downplay the impact of leadership in SDOs. Simcoe (2007) suggested the IESG members selection occurs via lottery mechanism rather than IETF community selection via a nominations committee that acts on behalf of the community. Simcoe (2007) was correct that the IETF selected the IETF's nominations committee members by lottery from active members for each year. However, an IETF nominations committee spends 6 to 9 months listening to community input before making decisions on the selection of IESG members This misunderstanding might have caused Simcoe (2007) him to downplay the impact of the leadership of these IESG members and the IESG cohorts as a TMT, which changed in composition each year. Simcoe (2007) also did not consider the impact of the IETF chair's leadership on the IESG.

Internet Engineering Task Force Statistics

The IETF has operated on consensus decision-making since its creation in 1987. Due to the 1992 to 2000 slowdown that Simcoe (2007, 2012) mentioned, the IETF administration had tracked the pace of IETF standards creations, approval, and publication as Request for Comment (RFCs). The IETF administration groups included the IETF secretariat, the RFC Editor (2016a, 2016b, 2020), and the IANA (2020). The RFC Editor (2016a, 2016b, 2020) and the IETF (2020m, 2020n, 2020o) secretariat had published statistics on the publication of IETF RFCs since 1987. Since 2007, IANA (2020) tracked and published monthly statistics on the publication of IETF standards. The IETF community has perceived that changes in the leadership of the IETF strategic management have slowed down or sped up this pace. IETF community discussions regarding this perception (sometimes heated debates) varied on who had caused the slowdowns.

The IETF's (2021b) two SMTs include the IESG (2000) and the Internet Architecture Board (IAB). The members of the IESG and IAB are all volunteers.

Some volunteers rotate off these SMTs each year, and the IETF nomination committee (nomcom) selects new volunteers to take their places. The IETF (2021a) chair leads the IESG, and the IAB chair leads the IAB. The IETF WGs create standards for a specific topic within the following general areas of Internet technology, such as applications, routing, Internet protocols, real-time data, transport layer, management and operations, and security. WG chairs lead each WG. ADs lead and manage the WGs within an IETF area and belong to the IESG. By organizational mandate, all decisions within the IETF are made using consensus decision-making, whether in a WGs, the IESG, or the IAB. The IETF community usually appoints individuals to be the members of the IESG and the IETF chair who will guide the IETF to create standards that keep improving the Internet. The community selects leaders it hopes will keep up the pace of the standards development and then change leaders if the pace of standards development is problematic. For example, a new IETF chair began work in 2013. The community focused on evaluating this IETF chair's progress in 2016 to 2017.

The current IETF (2016b) Datatracker system provides the data on the publication rate of RFCs information for the years 1968 to 2016 on a per-year and per-month basis plus statistics on authors, document life-cycle, document reviews performed by WGs, ADs, and IESGs, publication approvals, and RFC publication time. In addition, the IETF Datatracker database keeps track of document reviews, including dialogues between the authors and reviews on technical content. The IESG reviews also provide indications of vertical TI (WG to AD or AD to IETF chair) and AD to AD horizontal interdependence. These vertical and horizontal TI statistics indicate interactions of IESG members.

The IETF Document Statistics web page (Arkko, 2016a) provides links to IETF document statistics on RFC publication rates (monthly and yearly), document authors and documents per IETF areas via statistic web page (Arkko, 2016b), RFC Editor (2016a, 2016b) publication statistics on publication rates (monthly and yearly). This web page also provides document life-cycle statistics with total processing time, IESG processing time, RFC Editor (2020) processing (Arkko, 2012a), and review cycle statistics for IESG and ADs Review IESG and AD

Review (Arkko, 2012b). The IETF statistics (Arkko, 2016b) also track each author's country, affiliation (private/public), gender, citizenship, and geographical continent for each draft.

Hares Unpublished Research Into the Internet Engineering Steering Group at a Top Management Team

This researcher did two investigations into the IESG as a TMT before this research during Ph.D. projects during 2012 to 2013. The researcher treated these two investigations as Phase 1 and Phase 2 of a multiphase research project. The research in this document was Phase 3. The Phase 1 research in 2012 was a mixed-mode exploratory analysis of the phenomena of the IETF chair's leadership to determine if there was a difference between three individuals who led the IESG as an IETF chair. This exploratory IPA mixed-mode analysis arose from IETF Chair 6's observation in 2012 that not all IETF chairs had provided effective leadership. The IETF chairs examined were IETF Chair 4 (2001 to 2005), IETF Chair 5 (2005 to 2007), and IETF Chair 6 (2007 to 2012).

This section describes the method for exploratory research, the qualitative results, and the organizational theory which explains this result. Appendix F provides the consultant report, the mixed-method research using an IPA, the theme counts, and qualitative results. Appendix F also re-examines the consultant's report based on the data discovered in this research's 10% IPA analysis. Chapter 4 qualitative analysis also provides insights on the IESG under these three chairs.

The Phase 2 exploratory research was a 2013 survey of all IESG members on solidarity, TI, OCB, and perceived results. This section provides an overview of this survey's methodology to show how it helped form the methodology for the 2017 survey. The feedback on conflict from the post-survey interviews encouraged this researcher to add a conflict instrument and open-ended conflict questions to the survey. The researcher believed it was enough to cause the 2017 survey to be a unique one-time survey. However, the 2017 survey acted as a retest of the 2013 survey. Thus, the 2013 survey was another survey dataset for this research project. The 2013 survey had separate forms for the IETF chairs and the IESG members, just as the 2017 survey does. This research handled the 2013 survey response data according to this research's methods, as revised in Chapter 4. Chapter 4 also contains the quantitative results of the 2013 survey and a discussion of the validity risks and potential common method biases for the 2013 survey. Appendix Q contains the full details of the 2013 and 2017 surveys.

Hares Exploratory Interpretative Phenomenological Analysis Research (2012)

This leadership analysis used an IPA to analyze the formal minutes from five biweekly meetings of the IESG during 1 year of the leader's tenure. The middle year of the shorter tenures was chosen rather than beginning years since a leader's relationship with a SMT may take time to be established. The IESG data were collected from five IESG meeting formal minutes from 2003, 2006, and 2011 taken from the public IETF Web site (www.ietf.org/iesg/minutes). The researcher selected the five meetings per year from IESG meetings from March until May.

The qualitative IPA analysis encoded the following themes: (a) the format of minutes, (b) participants, (c) IESG actions, and (d) collaborative actions. Coding the format of the minutes for each minute validated that the minutes contain the same information. Coding the participants determined if the IESG meeting contained the same members or if some members were absent. Coding of the IESG actions determines how effective each IESG Cohort was in its tasks to (a) review and publish documents, (b) create and manage WGs, (c) handle management actions, and (d) handle liaison actions. Finally, coding for collaborative actions determined whether the group interactions were dyadic or multi-person activities. Multi-person interactions were consensus decisions that either multiple people to discuss the topic or no discussion. The assumption from "no discussion" was that the IESG had already reached a group consensus. After encoding all the IESG minutes, the qualitative analysis entailed gathering themes into master themes. MAXQDA (2016; a Computer Assisted Qualitative Data Analysis software) was used to track the themes and to generate theme counts for a mixed-mode analysis. The mixed-mode analysis uses descriptive statistics to compare results for the four categories of encoding (format of minutes, attendees, IESG actions, and group collaboration).

Exploratory interpretative phenomenological analysis research (2012) mixed-mode results. The mixed-mode analysis showed that the samples for each year were roughly equivalent and that each IESG cohort had a unique mixture of collaborative actions. The samples for each year (2003, 2006, and 2011) had the same format for formal minutes for each meeting and similar attendees. The formal minutes indicated the attendees to these meetings included IESG members, non-voting service members, and non-voting liaison members (IAB chair and IAB liaison), as Table 10 shows. The IESG meetings had between one to four IESG members missing. Outside of the IESG secretariat that sequences the meeting, the IESG members were the main actors in the meeting. For most meetings in the period surveyed, either all IESG members were present, or one member was absent. During two to three meetings, some IESG members were absent to attend other meetings. The service members included individuals from the IETF secretariat, ICANN, and the RFC Editor (2020). Starting in 2006, IETF volunteers who had provided narrative minutes also attended these meetings. The IAB liaisons included the IAB chair and the IAB liaison.

Table 10: Attendance in Internet Engineering Steering Group Meetings

Year	IESG members	Absent IESG members	Service members	IAB liaisons
2003	13	0-3	3-4	2
2006	15	1-2	5-9	2
2011	15	0-5	5-6	2

Each of the three IESG cohorts had a unique mixture of the types of decisions and the total number of decisions. The theme analysis of decision types found in each sample included decisions related to document publication, WGs, IETF management (MGT), IETF process decisions, and liaisons to other organizations. Table 11 presents the number of these master themes found in each year. Based on the theme counts in Table 11 for the sample taken, the IESG cohorts in 2003 and 210 attempted roughly the same number of decisions, but the IESG cohort in 2006 attempted only 123 decisions. The total number of decisions in 2006 was ~58% of the total decisions in 2003 and 2001. Fewer actions in 2006 suggested a less effective IESG based on the number of decisions attempted in the sample.

Table 11: Internet Engineering Steering Group Decisions

Year[2]	All decision	Document decisions	WG decisions	Management decisions	Process decisions	Liaison decisions
2003	214	103	39	66	5	1
2006	123	91	7	20	5	0
2011	210	148	18	33	4	7

The master themes related to cooperative interpersonal relationships were: dyadic actions, multi-person actions, zero discussion actions, and management actions. The researcher encoded the themes for management interactions in the IESG minutes for 2003 for situations when group interaction occurred, but there was insufficient information to categorize the group interactions. The IESG minutes in 2006 and 2011 had sufficient data to categorize all discussions as dyadic, multi-person, or zero discussion. Table 12 lists the theme count totals for the types of collaboration for the 5 minutes sampled in 2003, 2006, and 2011. Good consensus decision-making had multiple people discussing the decision before or within the meeting. Therefore, the multi-person or zero discussions indicate a productive discussion prior to a consensus decision. Conversely, a dyadic discussion was less likely to be a productive discussion before a consensus decision. The collaborative interactions (multi-person interactions and zero discussions) were 87% of the total decisions in 2003 and 78% of all decisions in 2011, but only 64% of the total decisions in 2006.

Table 12: Exploratory Research (2012) – Collaborative interactions

Year[3]	Dyad interaction	Multi-person discussion	Zero discussion	Management discussions	All I-P interactions
2003	2	52	39	12	105
2006	21	22	21	0	64
2011	17	32	28	0	77

[2] 2003 meetings: 4/3, 4/17, 4/30, 5/15, 5/29
2006 meetings: 4/13, 4/27, 5/11, 5/25, 6/8
2011 meetings: 4/14, 4/28, 5/12, 5/26, 6/09
[3] Same meetings as above.

Organizational theory and results. Fielder's (1964, 1967) LPC contingency theory explains the observed phenomena. Fielder's (1964, 1967) LPC theory states that leadership actions (low LPC/task-oriented, or high LPC/relationship-oriented) directly impact group performance moderated by situational issues of leader-member relationships, leadership position power, and task structure. This exploratory study concluded that the only significant situational influence that changed was the person who led the IESG SMT. Based on this influence, it indicated the leadership of the IETF Chair 5 in 2006 caused performance problems.

The exploratory study noted that the other potential influence was the interpersonal relationships of the cohorts in 2003, 2006, and 2011 or the tasks each IESG undertook. The membership of IESG changes every year, so the IESG cohorts in 2003, 2006, and 2011 contained different people. With different people on the IESG, personal interactions were different in these three IESG cohorts. In addition, each IESG cohort faced different tasks as new technologies arose for standardization and older standards need augmentation to meet new demands. This exploratory study indicated the value of examining the formal minutes for IESG decisions and the personal interactions between members (HS) and between the IETF chair and members (VS). The researcher gave a consultant report on this initial research to the IETF chair and the current IESG cohort in 2012 (see Appendix D for a copy of this report).

2013 Survey of the Internet Engineering Steering Group Members on Solidarity

This researcher surveyed the IESG members in 2013 to determine if organizational solidarity (S) is an antecedent of effective consensus decision making in the 24 years of the IESG TMT teams (1989 to 2012). Expanding the indication from the 2012 exploratory IPA analysis of IESG minutes that there were differences in the effectiveness of IESG consensus decision-making, the 2013 survey examined how IESG HS and VS relates to IESG effective consensus decisions. HS involved two or more people who cooperate by putting extra effort toward a collective goal. The HS investigated included HS between IESG members and HS between IETF chairs past and present. VS involved the leader or the member putting in extra effort to have a cooperative and productive vertical relationship. The VS investigated was between the IETF leader and each IESG member. The researcher defined the combination of HS and VS of individuals as group solidarity. The survey recorded the perceived results based on the self-reported opinions from the IESG members and IETF chairs to determine results.

Insights gained by the researcher from the 2013 survey guided this research proposal to avoid similar flaws. This section provides additional details on the assumptions of this research, the sampling methodology, and the valid responses received for the survey. The researcher also found flaws in the content and format of the survey instrument format. The assumption behind the 2013 survey was that the mean of individual responses would indicate the group behaviors for an IESG cohort. Stewart et al. (1999) suggested that group dynamics create a group norm that may differ from individual patterns of relating. The 2013 survey questions asked about solidarity and TI in relationship to other IESG members or the IETF chairs. If all members within an IESG cohort responded to the IESG survey, the mean represents an average view of the solidarity and TI of the members. Conversely, if only two out of 15 IESG members in a cohort replied to the survey, then the IESG cohort mean would only indicate the mean of these two surveys. The researcher tried to address this shortcoming in the 2017 survey by using IPA analysis of historical records to triangulate the survey results.

2013 survey sample method and participants. The researcher designed the 2013 survey as a one-time survey sent to all past and present IESG members and IETF chairs still active on the Internet. This research defines active IESG members as IESG members with an active email address. The active IESG members as of 2013 were 76 IESG members and six IETF chairs. The researcher sent the survey on July 1, 2013, and sent follow-up emails to remind survey participation from July 2 to July 30, 2013. The survey received 41 valid and 2 invalid IESG survey responses plus 4 valid and 2 invalid IETF chair responses. The survey was not anonymous and encouraged respondents to send open-ended comments to the researcher via email. The survey participants agreed to take the survey even though it was not anonymous. The survey's failure to be anonymous violated the best practices of human resource tests, although the IESG members agreed to this status before taking the survey. After additional training in the Ph.D. program made the student aware of this flaw, the researcher removed any demographic data from the 2013 survey data kept online. The researcher designed the 2017 survey to be anonymous.

2013 survey instrument. The survey instrument had the following parts: (a) IESG member identification as AD or IETF chair, (b) IESG terms, (c) behaviors section, (d) IESG consensus decision making, and (e) an optional demographics section. The IESG member identification section queries the person for a position (IESG member or IETF chair) and a term of service (years). This IESG member might wish to respond for each year IESG cohorts or bundle some IESG cohorts together. This paper used "cohort response" for an individual's response relating to one IESG cohort. The IESG term section allowed the person to select which years the behavior and consensus decision-making sections applied. The survey allowed the individual to bundle responses, and the post-processing created individual cohort responses from those bundled. The potential total number of cohorts from the 89 active IESG members is 289 cohort responses from 1989 to 2013 and 269 cohort responses from 1991 to 2013. The total number of IESG cohort responses was 129 from 1989 to 2013, with 125 of these cohort responses from 1991 to 2016.

The behaviors section combined an organizational solidarity survey instrument created by Lambooij et al. (2003), MacKenzie et al.'s (1991) OCB test for generalized compliance and altruism, and Van der Vegt et al.'s (1998) tested for task interdependencies. The behaviors section replicates the instrument except for replacing the phrase "co-worker" with "IESG team member" and replacing the phrase "supervisor" with "IETF chair." Hares's (2012) survey replicated Koster and Sanders's (2006) handling of these terms since the terms were made company-specific. All survey items had a 7 point Likert scale (1 = *strongly disagree,* 7 = *strongly agree*). This 2013 survey instrument replicated Koster and Sanders's (2006) survey, including the OCB tests' wording. The reliability tests run on survey response data indicated these scales are reliable (standardized Cronbach's alpha scores of 0.72 to 0.99). Appendix Q in Section 4 contains the details of the reliability tests. VS questions measured the cohort response in two directions (to/from) as viewed by the AD or the IETF chair. HS was also measured per cohort response in two directions as to/from for IESG members. Finally, the OCB and task independence scales focused on the individual's perception per cohort response.

The fourth part of this survey asked the individual to self-report on the effectiveness of the IESG cohort's consensus decision-making and the influence of the IETF chair. The survey asked the individual how effective their IESG cohort was when deciding to publish documents as RFCs or making decisions WG actions or IETF actions. The RFC publication question asked for the respondent's perception of the IESG effectiveness when publishing documents as standards versus non-standards (experimental, informational, and independent stream). The researcher chose to ask for both types based on the previous research from Simcoe (2012) and Gençer (2012). The final question in the fourth part asked the IESG member whether the IETF chair positively impacted their IESG cohort.

Flaws in the 2013 survey. The flaws in the survey content were the non-anonymous survey and the lack of consideration on how conflict impacts group solidarity. As discussed, the non-anonymous survey did not follow best practices for research regarding humans. The original 2013 analysis methods considered these self-reported opinions on results as the actual results, but this method has common method bias since the survey respondents provide both the independent variables (predictors) and the results (criterion). The conclusions for the 2013 survey suggested that any future research investigate using counts obtained from the IESG minutes on documents approved for publication, WGs, and IETF management actions approved. The methodology described in Section 3 addressed these errors by planning to use organizational results adjusted to align with the terms of IESG cohorts (March to March). As Chapter 4 discusses, the organizational results could not be adjusted to create per IESG cohort results, so Strand-2 of the Phase 3 research needs to collect these data from the IPA analysis. Due to time restrictions for Strand-2, the actual results used in the Phase 3 data analysis came from an estimate. The estimate came from the IPA of 10% of the IESG minutes from 1991 to 2016, confirmed by an IPA analysis of the 100% sample of the IESG using the alternate methodology for data collection. The Phase

3 data analysis of the survey responses examined whether the theoretical model fit the behavior data and the results when the results are perceived or estimated real results. Strand-5 qualitative examines why the perceived and the real results varied and whether this research's solidarity model should apply to both types of results.

The 2013 survey methodology included post-survey interviews. During these post-survey interviews, the researcher received comments regarding the format of the survey and results from the IESG members and the IETF chairs. Some survey respondents were concerned with their recollection of facts from their terms on the IESG. Other comments from IESG members and IETF chairs indicated that conflicts due to internal factions within their IESG cohorts or a contentious IESG member caused the IESG cohort to be less effective. The IESG interaction as a change agent in the network and IT industry caused interactions with other SDOs to remain contentious or highly political. These external stressors caused some IESG members to remain contentious as individuals or groups, and these individuals caused group conflict. Group conflict caused some IESG members to be demoralized. These content flaws discussions caused the researcher to add a proven instrument on conflict to the 2017 survey.

Interviews with IETF chairs after the 2013 survey also suggested an alternate reason for the results regarding the 2006 IESG cohort found in the 2012 exploratory survey. The IETF chair suggested the IESG members took on publishing standards documents with external deadlines which caused high levels of TC during the review process. These TCs engendered RCs, which diminished the IESG HS and VS and resulted in reduced group outcomes for the group's consensus decision-making. One future research consideration is whether conflict moderates the effect of the HS and VS antecedents.

Summary of Chapter 2

This literature review reviewed the current research and theories in consensus decision-making, conflict, and solidarity in teams and TMTs that formed the fundamental building blocks of the proposed research model. This chapter also considered the latest research on mixed-mode methodology application to leadership studies regarding groups, global groups, and TMTs. A part of this literature review on mixed-mode methods described how the concurrent triangulation methodology improved the reliability of qualitative and quantitative results by resolving validity threats and common method bias. One such threat to quantitative statistical conclusion validity occurred due to flawed moderator measurement methodology. Villa et al. (2003) recommended that tests should show a moderator relationship between constructs by using a theoretical derived simple model utilizing the moderator modeling techniques of Barron and Kenny (1986). The simple model should have variables for one predictor, one moderator, and one criterion. These three variables could optionally be augmented by optionally augmented by control variables. This consideration and the potentially small number of survey responses led this researcher to use a reduced model to detect the moderator in Strand-3 of the Phase 3 research.

Chapter 2’s literature review also reviewed previously published research on the IETF processes and this researcher’s 2012 exploratory research on the effectiveness of the IESG team consensus decisions. Simcoe (2007, 2012) published empirical results on the components of delays in the publication of IETF RFCs. This chapter provides an overview of Simcoe’s (2007, 2012) work and related work by other researchers. This chapter also summarized the results of this researcher’s Phase 1 exploratory mixed-mode research completed in 2012 and Phase 2 explanatory mixed-mode research in 2013.

Appendix D details the 2012 exploratory IPA analysis, and Appendix Q.4 details the 2013 survey. Chapter 3 describes the original methodology design for concurrent triangulation mixed research used for Phase 3 of this research. Chapter 4 describes developing an alternate methodology after initial triangulation showed the assumptions behind the original methodology were wrong.

Chapter 3 – Method

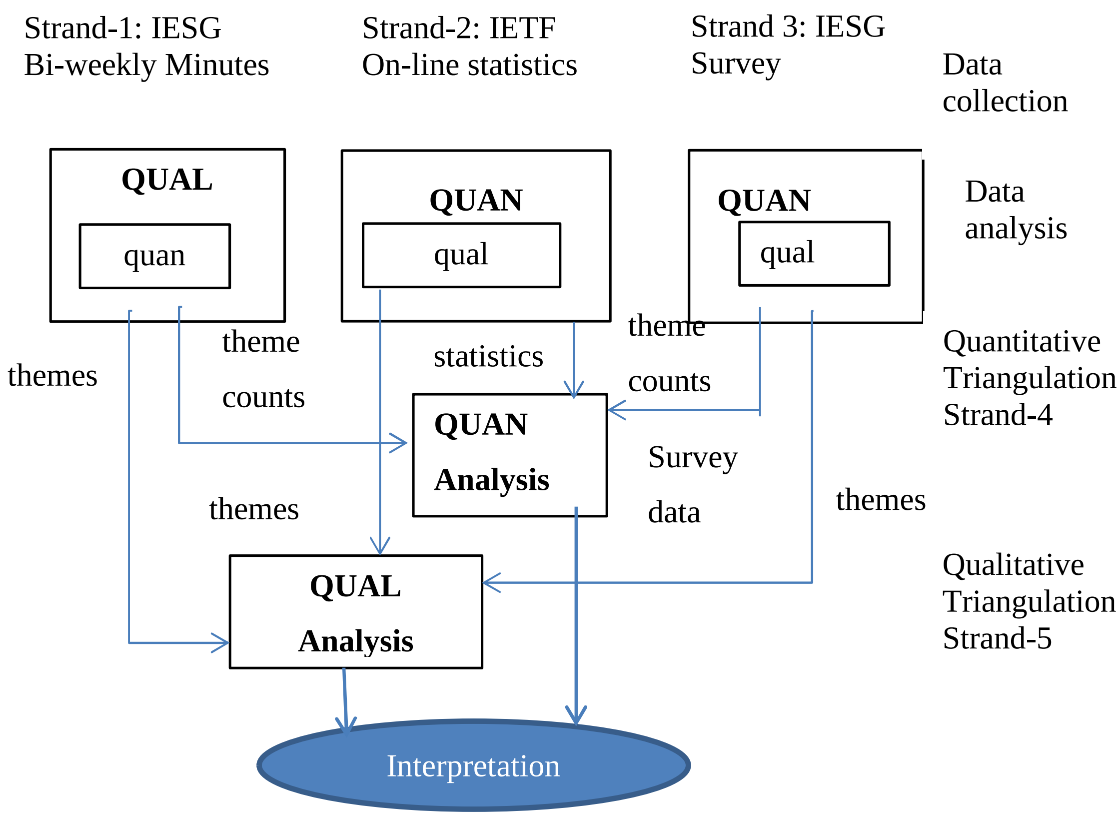
This longitudinal concurrent triangulation mixed-mode study sought to capture the interworkings of 26 different TMTs, which utilized consensus decision-making to create ICT standards in the same organization. Each of these TMTs operated as the IESG for the IETF during one year (March 1991 to March 2016), and each TMT per year (denoted as an IESG cohort). This concurrent triangulation mixed-mode study concurrently collected and analyzed online artifacts of each IESG cohort's interworkings in biweekly minutes and statistics on the results of effective decision-making and a survey each TMT member for their perceptions of the IESG cohort's interworkings and effective decision-making. This longitudinal concurrent triangulation study was the third phase of a multiple phase research project into interworkings of IESG in the IETF, and its methodology was to address problems found in Phase 1's exploratory research and Phase 2's explanatory research.

The first phase of this research was an exploratory qualitative IPA of a small subset (five out of 25) of the biweekly minutes of three IESG cohorts to determine if the leader impacted the IESG the IESG's effective consensus decision-making. As discussed, the results of Phase 1 showed that during the tenure of one of the leaders, the IESG cohort was less effective in their decision-making. In open-ended interviews regarding the results of Phase 1, some members of these three cohorts and the chair indicated the examination had flaws due to the limited sample of the minutes and the focus on collaboration rather than TMT member-to-member interactions and TMT member-to-leader interactions.

Phase 2 of this research was an explanatory mixed-mode study that combined a survey with follow-up interviews using open-ended questions. The Phase 2 survey queries each member of each IESG cohort to determine the member's perception of HS, VS, and effective consensus decision-making. The survey also queried each respondent for the individual's perceptions on OCB Generalized Compliance and Altruism, TI, and demographic information. The analysis of these survey responses in 2013 used descriptive and multivariate statistics (correlation and HRM) to determine if HS and VS were antecedents to the perception of effective consensus decision-making in the IESG cohorts. This analysis suggested that a member who perceived HS exhibited by other TMT members and VS exhibited by the IETF chair also perceived effective consensus decision-making. The interviews regarding the results indicated that TMT members acknowledge that TI was necessary to feel HS, but TC or RC between the IESG members could moderate the impact of horizontal or VS on decision-making. The IETF chairs interviewed also indicated that when IESG undertook large reorganization tasks, these tasks could strongly impact (positively or negatively) the interworkings of the IESG, and the IESG minutes would indicate these challenges.

This Phase 3 concurrent triangulation mixed-mode study had three strands of data collection and initial analysis: (a) qualitative exploratory mixed-mode analysis of the biweekly minutes (formal and informal) of each IESG cohort, (b) quantitative explanatory mixed-mode analysis of the IETF online statistics, and (c) quantitative exploratory mixed-mode study using a one-time survey of the IESG members with two open-ended questions. The design of Strands-1–3 focused on minimizing threats to validity, reliability, and common method biases. The mixed-mode analysis in Strands-1–3 generated quantitative and qualitative results. Strand-4's quantitative data analysis method quantitatively triangulated the quantitative results from the data analysis in Strands-1–3 and required a write up of a qualitative summary on the analysis. Finally, Strand-5 qualitative data methods triangulated the qualitative data from Strands-1–3 to determine if the patterns of TMT behavior aligned with the theoretical model. The triangulation in Strand-4 and Strand-5 operated concurrently. The qualitative summary note from Strand-4 aided the joint interpretation of the data. Figure 5 shows a diagram of the strands and interaction points in mixed mode methodology notation.

Figure 5: Mixed-mode design.



This methodology section describes this longitudinal mixed-mode study's original overall design and purpose. After describing the overall design, this section describes the data collection and analysis for Strands-1–3, expected sample sizes for data collection in Strands-1–3, and potential validity risks and bias collection. From this vantage point, this section will describe the concurrent triangulation design in Strand-4 methods and qualitative methods in Strand-5. Finally, this section contains the purpose of each facet of the mixed-mode methodology, the variables, and the points of interaction between strands, data collection methodology, data analysis plans, ethical concerns, schedule, and budget. This research involves a one-time re-survey of the IESG, which requires approval from the Institutional Review Board (IRB). This original methodology was revised in order to achieve accurate results. Chapter 4 describes the journey to establish an alternate methodology.

Longitudinal Concurrent Triangulation Methodology

This longitudinal concurrent triangulation mixed-mode methodology aimed to synthesize qualitative and quantitative data to discover how relationship and TC moderate HS and VS as antecedents of effective consensus decision-making in 26 different cohorts of the IESG, a TMT of the IETF doing open ICT standardization. The goal of Strand-1 examining the biweekly minutes and Strand-2 examining the IETF statistics was to determine how historical artifacts showed a particular TMT operated with HS and VS, conflict (task and relationship), TI, and effective consensus decision-making. The purpose of Strand-3 was to examine the perception of members of each of the IESG cohorts regarding the cohort. Due to the maturation of the members of many early TMTs, members could have dropped out of active participation in the IETF, forgotten details of the experience, or the experience might have taken on a “halo” effect. Due to this gap between perception and the actual experience that remained when considering artifacts and perception, it was necessary to triangulate behavioral patterns to determine the actual interaction of these 26 TMTs (1991 to 2016) operating to fulfill the same job in a volunteer organization. Therefore, each of the strands collected data and did a primary analysis (QUAL or QUAN) followed by a secondary analysis (quan or qual) to provide data for the triangulation process.

The triangulation phase of this methodology followed the initial analysis, just as Rowland and Parry’s (2009) triangulation research design had the triangulation phase follow the quantitative and qualitative research design. The quantitative triangulation phase was used to compare the quantitative data found in historical artifacts versus the perceptions of IESG members of the IESG cohort as a group. The qualitative triangulation was used to perform the same investigation from the qualitative data. The triangulation aimed to discover quantitative and qualitative data patterns that would confirm the proposed model from two viewpoints. These patterns might have arisen from the theoretically proposed influences to the IETF decision-making processes, such as technology complexity, business influences (participation or intellectual property rights [IPR]), or the IETF chair’s leadership of the IESG.

Variables in Full Model

This methodology utilized the following types in each of the five strands: independent variables, dependent variables, control variables, and discovered thematic variables. The definitions for independent, dependent, and control variables came from the theories discussed in Chapter 2. The discovered thematic variables were major themes unrelated to the theory-based variables discovered during the qualitative IPA evaluation of biweekly IESG minutes in Strand-1 or IPA evaluation of the open-end questions from the one-time survey in Strand-3. The researcher encoded each theory-based variable using the survey instrument's questions as the common ground between qualitative and quantitative methods in Strands-1–3 to facilitate the concurrent quantitative and qualitative data triangulation (Strands-4–5). The triangulation in Strand-4 used all the quantitative data from Strands-1, 2, and 3 to test the theoretical model and discover any new model that better fit those data. Similarly, the triangulation in Strand-5 used the qualitative data from Strands-1–3 to discover new themes. Additional patterns were found and annotated by coding memos based on the initial data combinations in these triangulation strands. These patterns were tested in Strand-4 by computed variables or encoding the patterns as themes in the qualitative triangulation in Strand-5. This section defines the theory-based independent, dependent, and control variables and the criteria for discovering non-theory variables.

This research into the group-level behaviors of the IESG as a TMT operated on composite variables at the group level. Each of the independent, dependent, and control variables in each of the four hypotheses of the full theoretical model was a composite variable measured at the individual level but compiled into TMT levels for a single year's IESG cohort as a quantitative mean score or as average quality for qualitative scores. Table 13 shows all theoretical variables with the type and level of each variable and the theoretical definition.

Table 13: Independent and Dependent Variables per Hypothesis

Variable	Hypothesis	Type	Theoretical definition
Horizontal solidarity	H1, H3, H5	Independent, composite at the group (team) level	Horizontal solidarity involves a co-worker cooperating with another co-worker by putting extra private resources toward the cooperative effort.
Vertical solidarity	H2, H4, H6	Independent, composite at the group (team) level	Vertical solidarity involves a follower cooperating by putting extra private resources toward the cooperative effort.
Relationship conflict	H3, H4	Independent, composite at the group (team) level	Relationship conflict is a composite of questions on Jehn's Intra-Group conflict (ICS) scale regarding relationship conflict items.
Task conflict	H3, H4	Independent, composite at the group (team) level	Task conflict is a composite of questions on Jehn's Intra-Group conflict (ICS) scale regarding relationship task items.
Effective consensus decision-making	H1–H4	Dependent, composite at group (team) level	Effective consensus decision-making involves making consensus decisions that all group members can support and enact in time fashion. The IESG's effective decision-making results involve publication of RFC, management of WGs, and management of the IETF.

Similarly, the independent, dependent, and control variables in the two hypotheses in the reduced model were composite group variables measured at the individual level but compiled into the group level for an IESG cohort. The measures of the theoretical variables in the IESG minutes (Strand-1) and the IETF Statistics (Strand-2) indicated the recorded actions, but the measures of the variables in the survey (Strand-3) indicated the perception of these same variables by the IESG members. The theme-based variables discovered in the IPA analysis of the IESG minutes (Strand-1) or the IETF statistics (Strand-2) were variables based on the recorded actions, but the theme-based variables discovered in the IPA analysis of the open-ended questions of the survey arose from perceptions of an IESG member. The researcher discovered these theme-based variables in IPA analysis in Strands-1–3 at the individual level in the historical record, but theme totals summarized this variable to a group-level composite variable.

Independent and Dependent Variables

The independent variable in Hypothesis 1 was HS, and the independent variable in Hypothesis 2 was VS. The independent variables in Hypothesis 3 were HS, the RC moderator, moderator factor (HSxRC), and the theory-based control variable of TI. The independent variables in Hypothesis 4 were VS, the RC moderator, the moderator factor (VSxRC), and the theory-based control variable of TI. The independent variables in Hypothesis 5 were HS, the TC moderator, the moderator factor (HSxTC), and the theory-based control variable of TI. The independent variables for Hypothesis 6 were VS, the TC moderator, the moderator factor (VSxTC), and the theory-based control variable of TI. The dependent variable for the six hypotheses was effective consensus decision-making which summarizes the information about each variable in terms of type, use in hypotheses, and definition. This researcher measured each of these variables on an individual basis. The group-level variable was a composite of the individual scores.

Survey research used instruments for each of the four independent behavioral variables to test an individual's perception. This researcher used the variable definitions and the survey instrument questions to test an individual's perception and as the theoretical basis for theme encoding. Other researchers used quantitative survey research to test five behaviors of HS between co-workers via self-reporting and five behaviors of VS between leader and followers (Koster & Sanders, 2006; Sanders & Schyns, 2006a; Schyns et al., 2006). The five behaviors are: (1) helping finish work, (2) helping when something goes wrong no one is responsible for, (3) apologizing if the person makes a mistake, (4) dividing pleasant and unpleasant tasks equally, and (5) living up to agreements. Strand-1 of this research used these questions regarding HS and VS behaviors as the IPA analysis's themes in the IESG minutes. Strand-3 used an adaptation of Koster and Sanders's (2006) survey instrument for HS and VS used in the Phase 2 survey sent to the IESG in 2013, which had good reliability (standardized Cronbach's alpha of between 0.71 to .99).

Similarly, Jehn's (1995) ICS measured an individual's perception of relationship and group conflict. Pearson et al. (2002) recommended the six items shown in Table 14 instead of the original 9 items on Jehn's (1995) ICS for relationship and TC. Pearson et al.(2002) used these 6 items in six survey experiments and found 6-item ICCS instrument was reliable with Cronbach alpha scores between 0.72 to 0.91. Strand-1 used these three RC questions as subthemes for the RC and the three TC questions as subthemes for the TC variable. Strand-3 used these questions in the survey of the IESG members to detect the perception of conflict.

Table 14: Jehn (1995) Intragroup Conflict Scale Items Used

Category	Jehn (1995) ICS items as revised by Pearson et al. (2002)
Relationship conflict	1. How much anger was there among the members of the group? 2. How much personal friction was there in the group during decisions? 3. How much tension was there in the group during decisions?
Task conflict	1. How many disagreements over different ideas were there? 2. How many differences about the content of decisions did the group have to work through? 3. How many differences of opinion were there within the group?

The dependent variable effective consensus decision-making indicated how effective the IESG operated as a TMT. Theoretically, effectiveness in consensus decision-making was to make decisions that all group members can support and enact quickly. Effective decision-making was a composite variable measured at the team level for each TMT cohort in the IESG minutes (Strand-1), the IETF statistics (Strand-2), and the survey form (Strand-3). This researcher calculated IETF effectiveness based on three qualities: RFCs published, WG actions, and IETF management. The WG processing involved the IESG's review of establishment, recharter, and closing of a WG. The IETF management items included items discussed in the IESG biweekly meetings, IESG statements, liaison reports, IESG retreats, and IESG sessions at IETF (public and IESG only).

Table 15: Internet Engineering Steering Group Work Outcomes

Model	Composite value calculation for model
RFC only	The number of RFCs created during the period of an IESG cohort
Working Group (WG) Actions	The number of WGs established, rechartered, or closed during the period of an IESG cohort
IETF Management	Total Management items discussed at all IESG biweekly meetings for the year

Control and Discovered Variables

This researcher measured the following five control variables per cohort year: cohort year, demographic score, TI, OCB generalized compliance, and OCB altruism. The cohort year control variable allowed grouping of the historical data per TMT. The demographics were collected to determine if TMTs were similar or different in composition. The TI was a control variable in Hypotheses 5 and 6. Finally, the OCB measures allowed the researcher to test for the concurrent and discriminant validity between OCB and solidarity measures at individual and team levels. This researcher used all the control variables in the full model, but the reduced theoretical model combined OCB-GC and OCB-A into a single score (OCB). This single score for OCB provided an alternate construct to solidarity.

This researcher collected control variable information in all three strands at the individual level per IESG member. Group variables were composites of the individual variables. The dating of online materials determined the TMT (IESG) cohort year in the IESG minutes in Strand-1 and IETF statistics in Strand-2. The survey's term answers on the survey forms provided the year of the cohort for individual responses for Strand-3. The IESG minutes indicated the IESG members in attendance in each meeting of the IESG. Strand-1's IPA analysis of the meetings saved a list of who attended or did not attend each meeting. Strand-2's research augmented the attendance information with online biographical information on the Internet. Strand-3's one-time survey optionally queried for the same information to respect ethical considerations.

The IETF community eschews mandatory declaration of age, gender, and education so that few IESG may fill in these demographics. The demographic score was a composite score of average age, average gender (1 = *male*, 2 = *female*, and 3 = *others*), average education level (1 = *no high school diploma* to 7 = *postdoctorate work*). The average education level was the highest level reached by an individual. This researcher kept demographic data relating to an individual and any results of demographic analysis in a separate offline database to increase the confidentiality of this sensitive data. Only group scores, which were ratios, were uploaded to the SPSS database for processing.

Table 16: Control Variables

Variable	Hypothesis	Type	Theoretical definition
Task interdependence	H5, H6	Control variable, composite at the group (team) level	Task interdependence is the extent to which members rely on each other to complete their jobs
Year of IESG cohort	H1-H6	Control sets TMT cohort	The year for the IESG term starts at the IETF's Spring meeting (March) and goes to the following year's Spring IETF meeting (March).
OCB – Generalize Compliance	H1-H6	Control variable of an alternate construct, composite at the group (team) level	OCB Generalized compliance is "impersonal helpful behavior" (Fields, 2002, p. 245), such as being on time or not wasting efforts in meetings. The research uses this variable to determine concurrent and discriminant validity
OCB -Altruism	H1-H6	Control variable of an alternate construct, composite at the group (team) level	Fields (2002) defines OCB Altruism as extra-role efforts to help "co-workers personally" (Fields, 2002, p. 245). The research uses this variable to determine concurrent and discriminant validity.
Demographic score	H1-H6	Optional control variable used to determine if IESG cohorts have similar backgrounds. *1	Demographic score based on the average age, average gender (scored 1-male, 2-female, 3-other), and the average highest level of education (scored 1-7, no high school degree to postdoctorate research).

*1 – Alternate methodology did not use this score.

TI, OCB generalized compliance (OCB-GC), and OCB altruism (OCB-A) were measured individually and summarized into a composite at the TMT cohort level. Koster and Sanders (2006) definitions of generalized compliance and altruism from OCB research (Wayne & Cordeiro, 2003) and TI from Van der Vegt et al. (1998). Koster and Sanders (2006) defined generalized compliance as "what a good employee ought to do" (p. 525). The IETF equivalent is "what a good volunteer should do." One definition of OCB altruism (OCB-A) entailed "helping others" (Koster & Sanders, 2006, p. 525) by supporting and orientating new team members. Koster and Sanders (2006) defined TI as "the extent to which members rely on each other to complete their jobs" (p. 525). The Phase 2 survey used these definitions but substituted the term "IESG member" for co-worker" and "IETF chair" in the survey instrument questions.

Table 17 shows the statements used for TI, OCB-GC, and OCB-A. Strand-1 used these statements in the IPA to determine if these behaviors existed in the IESG minutes. Strand-2 determined if artifacts of the control behaviors existed in the online statistics based on these survey instrument statements. Finally, Strand-3 used these statements on the survey instrument to query each IESG member for their perceptions of these group behaviors.

Table 17: Control Variable Survey Statements

Variable	Survey statements
Task interdependence	1. To perform my tasks, I need information from my other team members.
	2. I depend on my co-workers on the IESG to be able to do my job well
	3. To perform my tasks, I have to work together with other team members on the IESG
OCB generalized compliance	1. I fulfill the obligations as stated in the IESG/IETF chair job description
	2. I fulfill all the formal responsibilities that come with my job
	3. I am satisfied with my job performance
Altruism	1. I will help someone who is very busy
	2. I will help doing tasks for others when they are sick or absent

The method for discovering non-theory variables used the IPA methodology to discover themes and then summarized these themes into major themes. The IPA methodology for the theoretical variable was to conduct an exploratory IPA analysis of 5% of the formal minutes from July 1991 to March 2016 and 5% of the informal minutes (December 2005 to March 2017). If new significant major themes were found during this exploratory analysis, these themes were coded for in the IPA evaluation of the IESG minutes and the open-ended questions. Similarly, in

Strand-3, if the IPA analysis of the two open questions detected new significant themes, these themes were examined as a discovered variable. When the qualitative evaluation of the IESG statistical analyses discovered a set of qualitative patterns, the researcher examined the IESG minutes and the survey's open questions for these discovered qualitative patterns. The researcher documented the conclusions of this theme discovery process in Strand-1's IPA analysis and Strand-3's open-ended questions in code memos.

Variables in Reduced Model for Moderation

Chapter 1 described the reduced theoretical model to test for conflict moderation. In this model, solidarity is the predictor, conflict is the moderator, and effective consensus decision-making is the dependent variable. Solidarity is a composite score that combines the HS and VS scores to create a single score. Conflict is a composite score that adds TC and RC to create a single composite score. Finally, effective consensus decision-making remains the composite score of the yearly statistics on RFC published, WG management tasks, and IETF management tasks. The individual respondent's perceived IESG effective consensus decision-making score remains the mean of each respondent's answers on survey questions in Part 4, Questions 1a, 1b, 2, and 3. The dependent variable on score allowed the concurrent triangulation in Strand-4 to compare the real effectiveness of an IESG TMT with the perceived effectiveness. The composite group score for an IESG cohort is the mean of the perceived results, which is the mean of the individual scores.

The control variables for the reduced model includes TI, cohort year, and OCB. The TI and cohort year variables have the exact definition as in the full theoretical model. Hypothesis 2 of the reduced model includes TI as a control variable. The cohort year variable defined the yearly IESG TMT as a group. The control variable OCB was created as the combination of the OCB-GC and OCB-A variables. The composite OCB was generated as the sum of theme counts for OCB-GC and OCB-A and the average survey scale scores for OCB-GC and OCB-A. Although there were other components of OCB, this definition of OCB allowed the reduced model research to have a single variable to contrast with solidarity to determine if solidarity was a discriminant construct validity when this construct was measured concurrently with solidarity. The researcher did not plan to use the demographics ratio in the reduced data model. With limited responses, the differences in demographics was considered too risky.

Table 18: Independent and Dependent Variables per Hypothesis

Variable	Hypothesis	Type	Theoretical definition
Solidarity	H1, H2	Independent, composite at the group (team) level	Composite of Horizontal solidarity and vertical solidarity described in full model above. Horizontal solidarity involves a co-worker cooperating with another co-worker by putting extra private resources toward the cooperative effort.
Conflict	H2	Independent, composite at the group (team) level	The conflict variable is the Intra-Group conflict score from relationship conflict and task conflict.
Effective consensus decision-making results	H1-H2	Dependent, composite at the group (team) level	Effective consensus decision-making involves making consensus decisions that all group members can support and enact in time fashion. For example, the IESG effective consensus decisions results include IETF RFCs published, WG actions (e.g., WG creation), and IETF management actions (e.g., IETF meeting schedule setting).

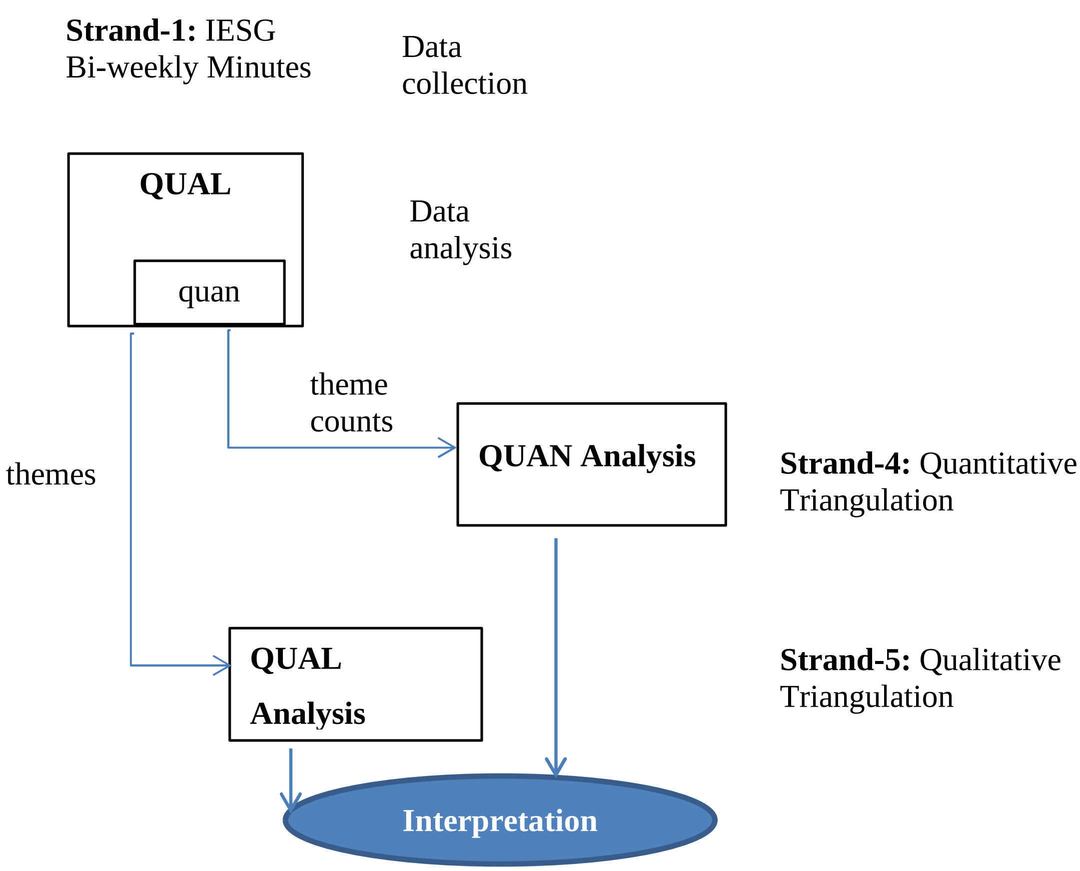
Data Collection Methods

The researcher collected data for Strands-1–3 in parallel. The researcher collected the data for Strand-1 by downloading the IESG formal and informal minutes available at http://www.ietf.org/iesg/minutes.html. The researcher obtained the performance data for Strand-2 by searching for the correct data on the following websites: www.ietf.org, the online IETF data tracker (https://datatracker.ietf.org), the RFC Editor’s (2020) list of RFC publications, and other online sources for biographical (company websites, LinkedIn, academic paper biographical notes). The Strand-3 data came from a second survey of the IESG members of these 25 TMTs (1991 to 2015). Strand-3 data contained both quantitative scores from instruments and open-ended questions. Strands-1–3 collected both quantitative and qualitative data. The research triangulated the quantitative results of the analysis from Strands-1–3 in Strand-4's quantitative triangulation and the qualitative results from Strands-1–4 in Strand-5's qualitative triangulation process. This section provides an overview of the data collection process for each strand, sample size considerations, assumptions on data linearity due to Hares (2013), potential common method biases, and resolution.

Data Collection in Stand-1

The first step of data collection was to download the formal and narrative minutes from the IETF online source (http://www.ietf.org/iesg/minutes.html) into the MAXQDA (2016) database. There were 598 formal minutes from the 1991 to 2016 IESG cohorts and 246 narrative minutes from the 2005 to 2016 IESG cohorts. The formal minutes were 2 to 6 pages, and the narrative minutes were 6 to 300 pages in length. All biweekly minutes were downloaded from the IETF website and uploaded into the MAXQDA for processing.

Figure 5: Strand-1 mixed-mode design.



The themes count totals for the major themes, themes per survey question from the survey, and discovered variables were planned to be transferred from the MAXQDA (2016) program to the SPSS program using the MAXQDA feature to output SPSS files. This MAXQDA to SPSS transfer ability operates per file, so the theme counts totals were planned to be uploaded per biweekly IESG minutes.[4] The researcher named the files uploaded from MAXQDA with the IESG cohort date. The first step in SPSS processing these data was to encode the date of the IESG minutes and IESG cohort into the uploaded files based on the file's name. The Strand-1 quantitative analysis using the SPSS program created the yearly cohort sums from these meeting sums.

The sample size for quantitative analysis of the theme counts depended on the number of decisions per IESG cohort year. Based on the Phase 1 exploratory IPA analysis, the decisions per meeting for a 100% sample were between 25 to 43 decisions or between 565 to 950 decisions per year with 23 meetings per year. Group behavior theme counts per IESG decision were sums of the individual behaviors detected for that decision. During a particular IESG decision, there might be zero individual behaviors for a particular behavior per decision, which would mean the group behavior theme count for that decision would be zero. For example, a decision could have zero individual behaviors of VS behaviors, so the group behavior theme count would be zero. Therefore, the total theme counts for a behavior per meeting is the sum of the per decisions theme counts total for all decisions. For example, suppose the RC theme has three individual themes encoded in one decision, and no theme counts in the other 40 decisions. Then, the meeting theme count for the RC theme would be 3.

<u>Data Collection in Strand-2</u>

The researcher collected Strand-2 quantitative and qualitative data from the IETF's website (www.ietf.org) and the IETF's online database (datatracker.ietf.org), the RFC Editor's website (www.rfc-editor.org), the IANA

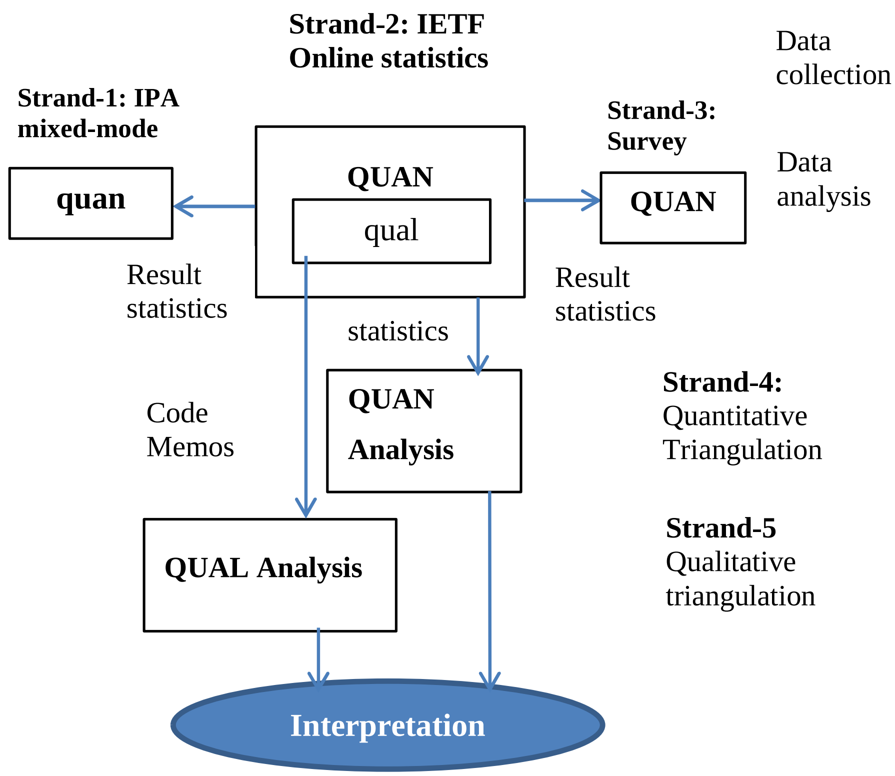
---

[4] This method of theme count total collection could not be used in the revised methodology. A hand-merge of the formal and narrative minutes was needed to determine create the actual list of IESG decisions and behaviors during those decisions.

web page (www.iana.org), and professional websites (e.g., linked-in). The quantitative data collected in Strand-2 included the statistics for each IESG cohort's results from consensus decision-making, evidence of control variables (TI, OCB-GC, and OCB-A), and demographics for the attendees detected in Strand-1's IPA analysis. Appendix E has the codebook that describes storing these quantitative variables based on the survey instrument. The statistics on effective consensus decision-making results included the rate of RFC publication, WG management, and IETF management items. If these statistics on results came from multiple sources, the researcher triangulated the sources during the analysis. For example, the rate of publication of RFCs is provided from three sources: IETF website, RFC Editor's website, and IETF Datatracker (an online database). The researcher also collected data on RFC publication that allowed the Strand-2 analysis to look at Gençer's (2012) theory of clustering and Simcoe's (2007) theory of technology complexity and WG conflict.

The researcher also collected demographic data on IESG meeting attendees from IETF websites, RFC publications, and professional and educational websites. This sensitive demographic data was stored offline to respect the privacy of each attendee, and no demographic data were present per individual. Although the researcher collected the ancillary demographic data from public websites—the best practice for retaining the privacy of information recommends this practice of keeping the data in an offline storage

Figure 6: Strand-2 mixed-mode design.



Before forwarding the quantitative statistics to Strand-3 for use in the data model, the researcher made three checks on the validity of the data. First, the researcher checked the data outliers or non-linear distributions. These errors pointed to erroneous statistics collected from the IETF website. Next, the researcher cross-checked the data against the errors found in Phase 2 research. Third, the researcher triangulated the result statistics per cohort year against the decisions with results found in the IPA analysis of Strand-1. These two measures of the results (statistics and IPA theme counts) should correlate since the IESG approves RFCs for publications. If all three tests were valid, the statistical data collected from IETF online sources was considered valid.

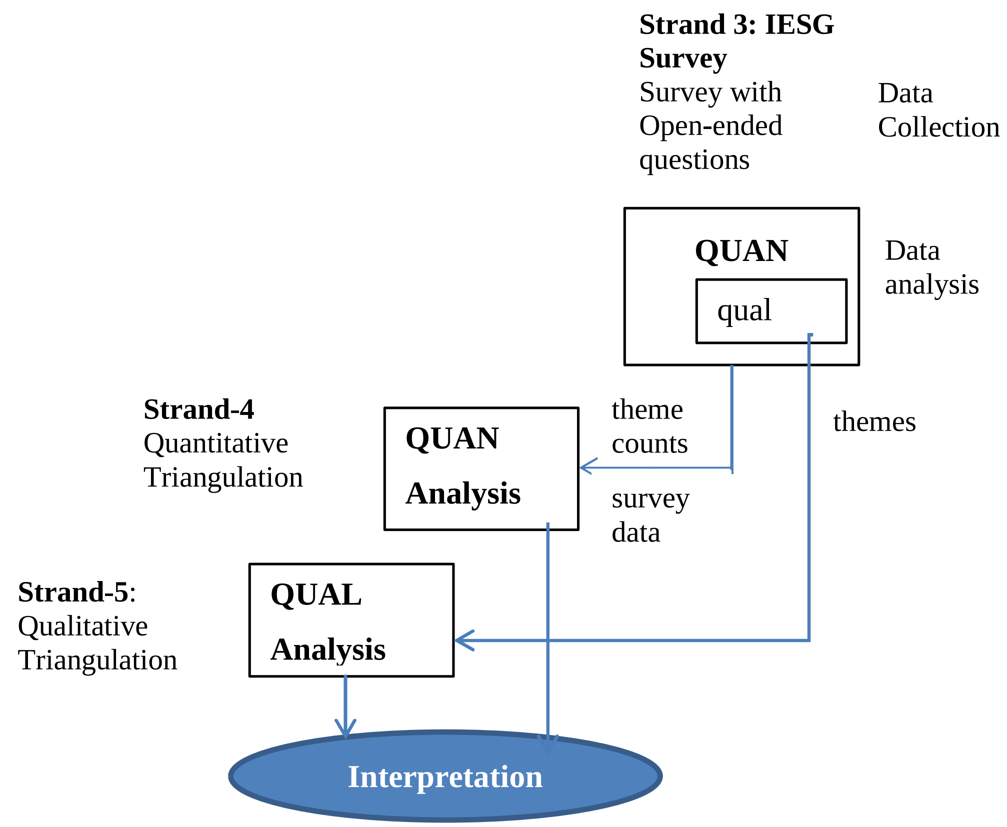
Strand-2 collected qualitative data on the progression of technology standards in WGs per area and changes in focus per IETF chair. The researcher collected data on the WG charter and the technology standards created by each WG in the IETF's history. This WG data included the type of technology standards the WG was to create, the number and type of standards published, the duration of the WG, and ADs (IESG members) who managed the WG. The WGs the researcher collected data on are listed in Appendix R. The qualitative analysis for Strand-2 examined the technology to determine the progression of Internet technology during the IESG cohorts per IETF chair. The qualitative data collected per IETF chair came from the IETF Meeting Proceedings that indicated the focus of the IETF chair, external forces impacting the IETF, and organizational changes made by the IETF chair. During the data collection, code memos were used to document the validity of Gençer's (2012) or Simcoe's (2007) theoretical models. These code memos were input into the qualitative analysis for Strand-2 and Strand-4.

Data Collection in Strand-3

This quantitative survey researcher used a one-time online survey of all active IESG members and IETF chairs to obtain self-reported perceptions per IESG term. Active IESG members were those people who had served on the IESG and had active email addresses at the time of the survey. Active IETF chairs were those individuals who had served as an IETF chair and had active email addresses. The survey asked each respondent for their perceptions on the independent variables (HS and VS, RC, and TC), dependent variable (effectiveness of IESG consensus decision making), the control variables (demographics, TI, OCB-GC, OCB-A). This researcher resurveyed IESG members with this expanded survey as a one-time event.

This section discusses the sample size and the participants, the survey instrument, and compares the data collected in the survey with the data collected from the historical record found in the IESG minutes. This 2017 survey expanded the survey used in Phase 2 to include the 6 items from Jehn's (1994, 1995, 1999) ICS and two open-ended questions on conflict. Pearson et al.'s (2002) analysis of Jehn's (1999) 9 items suggested that 6 items (three for RC and three for TC) "best-captured relationship and task conflict" (p. 110). The 2017 survey used the online resources at surveymonkey.com.

Figure 7: Strand-3 mixed-mode design.



Sample size and participants in survey. The IETF (2016a) website lists 97 IESG members and seven[5] IETF chairs for IESG in the 28 years from July 1989 to March 2017. The IAB and IETF chair created the IESG in July of 1989, and each IESG as a volunteer TMT begins its year of service with the first IETF meeting of the year, which usually occurs in March. Four IETF chairs also served as IESG members before becoming IETF chairs. The total population of IESG members and IETF chairs is small, but most individuals served on multiple IESG cohorts. Theoretically, the 97 IESG members served in 28 IESG cohorts and could provide 355 responses about group behaviors during a particular IESG cohort. Similarly, theoretically, the seven IETF chairs could provide 28 years of responses. For this study, the researcher defined the yearly response for both IESG members and IETF

[5] One IETF chair (Michael Corrigan) indicated on www.iein email during Phase 2 survey in 2013 that he should not be included as other IETF chairs.

chairs as IESG cohort responses. However, not all 97 IESG members and seven chairs were active and available via email to take the survey. The researcher defined an active IESG member as one with an active email address and active IETF chairs as IETF chairs with an active email list. The actual sample for this survey was 81 active IESG members (1989 to 2016) and five active IETF chairs (1989 to 2016).

The Phase 2 survey in 2013 had a total population of IESG 89 members and 7 IETF chairs from the IESG cohorts from 1989 to 2016. The number of active IESG members in 2013 was 76, and the number of active IETF chairs was six. For the 2013 survey, the potential number of responses from the IESG members in 2013 was 289 cohort responses, and the potential number of IETF chair responses was 23 cohort responses. The 2013 survey received 43 IESG responses (57% active IESG members) and six IETF chair responses (100% active IETF chairs). The valid IESG member responses in 2013 were 41 IESG members (54% of active IESG members), covering 129 IESG cohorts (45% of active IESG cohort slots). The IETF chair returned four valid responses (67% of active IETF chairs), covering 14 IESG cohort slots (61% of active cohort slots). Although a 50% response rate was suitable for general surveys, the small number of active members meant the statistical analysis needed to consider the reduced theoretical model.

The researcher obtained the email addresses for the active IESG members from the IETF (2016b) Datatracker, email lists, and online public websites. Before the 2017 survey, the researcher collected and verified the email address. After collecting email addresses, the researcher announced the survey via email. The announcement requested the IESG members and IETF chairs to participate in an online survey hosted at SurveyMonkey and gave the date for the survey.

Survey instrument. The survey instrument’s four parts (IESG Term, IESG team behaviors, IESG consensus decision-making, and demographics) had two formats: the IESG member form (Appendix A) or an IETF chair form (Appendix B). Each of these formats had six parts: IETF Term, IESG Behaviors, IESG attitudes toward the job, IESG consensus decision-making, demographics, and open-ended questions. The IESG term selection asked for the person's role (IETF chair/IESG member), the year(s) this response applied to, and for an informed consent. The online survey presented the appropriate form based on the individual's role (IETF chair/IESG). The survey form allowed the respondent to bundle responses over multiple years. In interviews relating to the Phase 1 consulting project (see Appendix D) and in Phase 2 post-result interviews, the IESG members indicated IESG behaviors changed slowly, so bundling responses was appropriate. The 2017 survey form design handled bundling of cohort responses.

The IESG team behaviors section of the survey contained 10 HS items and 10 VS questions from Koster and Sanders's (2006) survey, plus six relationship and TC questions from Jehn's (1999) survey. These 26 items used a 7- point Likert-like scale from 1 = *strongly disagree* to 7 = *strongly agree*. This survey replicated Koster and Sanders' (2006) survey, replacing their term "co-worker" with "IESG team member" and the phrase "supervisors" with "IETF chair." Since Koster and Sanders modified the general questionnaire to fit the nine different organizations, this modification did not change Koster and Sanders's (2006) methodology. Koster and Sanders (2006) and previous research projects in Holland (e.g., Lambooij, Flache, & Siegers, 2009; Lambooij, Sanders, Koster, & Zwiers, 2006; Sanders & Schyns, 2006b; Sanders & Van Emmerik, 2004) found the Organizational Solidarity (OS) scales reliable, with Cronbach alphas of solidarity toward supervisor (0.75 to 0.85), solidarity from supervisor (0.89), solidarity toward co-workers (0.77 to 0.89), and solidarity from co-workers (0.92). The 2013 survey of this research found the reliability of these scales used for the IESG with 45 IESG members was reliable with Cronbach's alpha of 0.77 to 0.82 (HS toward IESG member; a = 0.82), HS from an IESG member (a = 0.88), VS exhibited toward IETF chair (a = 0.77), VS from an IETF chair (a = 0.82). Pearson et al. (2002) found the reliability of the 6-item version of the Jehn (1995) Scale in five surveys had a Cronbach's alpha of between 0.85 to 0.87 for three RC items and a Cronbach's alpha of between 0.72 to 0.91 for the three TC items. The team behaviors section also asked the two open-ended questions shown in Table 19. This section was last and separated from Jehn's (1999) questions on TC and RC.

Table 19: Open-Ended Questions for Part 6

Focus	Question
Task conflict	In some years, the IESG undertook any significant tasks that caused conflict. If this happened during your term(s), could you explain/ describe the [relationship] between the task and the conflict?”
Relationship conflict	In some years, IESG members experience more conflict in interpersonal relationships than the members expected. If this happened during your term(s), could you describe how this occurred and how it impacted you or your work?

The third section on “IESG attitudes toward job” contained three items on OCB generalized compliance, two on OCB altruism, and three on TI. In addition, this section provided control variables for the theoretical models (full and reduced). These eight items used a 7-point Likert-like scale from 1 = *strongly disagree* to 7 = *strongly agree*. This section utilized the format of the OCB generalized compliance and OCB altruism questions from MacKenzie et al. (1991) in the form that Koster and Sanders (2006) provided them. Koster and Sanders (2006) found the reliability for scales were Cronbach alpha for OCB general compliance of 0.70 and OCB altruism of 0.70. The Phase 2 survey reliability test found the scales were reliable with the IESG with a standardized Cronbach of 0.74 for OCB generalized compliance and 0.76 for OCB Altruism. The three items for TI come from a reliable instrument from Van der Vegt et al. (1998). Koster and Sanders (2006) used this scale in survey research and found the TI scale was reliable with a Cronbach alpha of 0.81. The Phase 2 research used this TI scale with the IESG members and found the scale reliable with a standardized Cronbach of 0.84. These tested instruments with previous experience with the IESG provide a set of control variables.

The fourth section of this survey asked the person to self-report on the effectiveness of the consensus decision-making process and the impact of the IETF chair on the IESG’s effectiveness. This section of the survey measured effectiveness of the IESG consensus decision-making process by making statements about RFC processing, WGs, and IETF process management, and asking the respondent to *agree* or *disagree* based on a 7-point Likert scale, with 1 =

strongly disagree to 7 = *strongly agree*. First, the survey queried the effectiveness of RFC decisions by asking the IESG respondent to rate the effectiveness by subtypes of RFCs as standards and non-standards. The IETF (2021b) defined standards as RFCs for proposed and full IETF standards while defining non-standards RFCs as informational, experimental, and independent stream editor (ISE) RFCs. Although this section’s statements regarding WGs and IETF process management are brief, the IETF uses these expanded definitions. For example, WG actions include BOF actions (proposed, approved, and held) and WG actions (creation [chartering], rechartering, management, and closure). The IETF process management includes IESG processes and meetings, IETF meeting schedules, liaison statements on technology, and resolving IPR issues. The final statement in Section 4 inquired whether the IETF chair positively impacted the IESG’s effective consensus decision-making for the IESG cohort. The survey question asked the respondent to give their perceptions using a 7-point Likert scale (1 = *disagree strongly* to 7 = *agree strongly*).

The fifth section of this survey asked for optional demographic data and survey feedback. The demographic data included age, gender, and education. The age question queried the survey respondent for the person’s age. The gender scale used a scale that included three categories (male, female, and other). The education used an 8-point Likert scale, with 0 = *no high school education* to 7 = p*ostdoctoral studies*).

Data collection of the online survey responses occurred anonymously on an online website (SurveyMonkey.com). After the researcher downloaded the responses from the website, the researcher translated the multiyear responses to multiple single-year responses for both the IETF chair and the IESG members. First, the researcher split the data from each single-year response into the following three parts: (a) quantitative survey answers from Sections 1 to 4, (b) open-ended question responses (Section 2), and (c) the demographic data. After splitting the data, the researcher uploaded the quantitative survey data (single-year responses) to the SPSS program. After checking the open-ended question responses for anonymity, the researcher uploaded the single-year responses to the MAXQDA

(2016) program. Finally, the researcher uploaded the demographic data to the SPSS program and then stored the original data and the SPSS file in offline storage.

Table 20: Data Collection Variables in Strand-1 and Strand-3 for Full Model

Variables	Type of data	Strand 1: Minutes (Reality)	Strand-3: Survey (Perception)
	QUAL	Qualitative IPA analysis encoded the themes based on the 10 survey questions for vertical solidarity as subthemes under vertical solidarity (hierarchically stored in MAXQDA, 2016).	None
Vertical solidarity (independent)	QUAN	The researcher used MAXQDA (2016) mixed-mode tools to generate the total theme counts for vertical solidarity themes per IESG minutes per type of IESG minutes.	The survey used Koster and Sanders's (2006) solidarity instrument tuned for the IESG. The 2017 survey had 5 questions regarding receiving vertical solidarity and 5 questions about transmitting vertical solidarity. An individual score for vertical solidarity was the mean score for all 10 questions.
		An IESG cohort theme count total for vertical solidarity was the summation of the meeting theme counts totals for the cohort year.	An IESG cohort mean was calculated as the mean of all individual IESG cohort responses for that cohort year.
	QUAL	Qualitative IPA analysis encoded the themes based on the 10 survey questions for horizontal solidarity as subthemes under horizontal solidarity (hierarchically stored in MAXQDA, 2016).	None
Horizontal solidarity (independent)	QUAN	The researcher used MAXQDA (2016) mixed-mode tools to generate the total theme counts for horizontal solidarity themes per IESG minutes per type of IESG minutes.	The survey used Koster and Sanders's (2006) solidarity instrument, with 5 questions about receiving horizontal solidarity and 5 questions about giving horizontal solidary to another person. An individual response (IESG member or IETF chair) is the mean score for all 10 questions.
		An IESG cohort theme count total for horizontal solidarity was the summation of the meeting theme counts totals for the cohort year.	The IESG cohort score was the mean of individual scores per cohort year.

Variables	Type of data	Strand 1: Minutes (Reality)	Strand-3: Survey (Perception)
Interpersonal relationship conflict (independent)	QUAL	The researcher encoded the interpersonal relationship conflict themes in the IPA based on the three questions in the survey. These themes were summarized into the relationship conflict master theme via hierarchical structure in MAXQDA-11.	Qualitative IPA analysis of the two open questions encoded interpersonal conflict in the same manner as the IPA.
	QUAN	The researcher used MAXQDA (2016) mixed-mode tools to generate the total theme counts for relationship solidarity themes per IESG minutes per type of IESG minutes.	The individual IESG member score was a mean of the survey’s 3 items on intragroup interpersonal relationship conflict from the 6-item version of Jehn’s (1999) ICS scale.
		An IESG cohort theme count total for relationship conflict (RC) was the summation of the meeting theme counts totals for RC for the cohort year.	IESG cohort mean was the mean of all individual scores for relationship conflict per cohort TMT.
Task conflict (independent)	QUAL	The researcher encoded the task conflict themes in the IPA based on the three questions in the survey. Then, these themes were summarized into the task conflict (TC) master theme via hierarchical structure in MAXQDA-11.	Qualitative IPA analysis of the two open questions encoded task conflict similar to the IPA.
	QUAN	The researcher generated via MAXQDA the theme count totals for task conflict themes are summarized per recorded minutes (formal and informal).	The individual IESG member score was a mean of the survey’s 3 items on intragroup task conflict from the 6-item version of Jehn’s (1999) ICS scale.
		An IESG cohort theme count total for task conflict (RC) was the summation of the meeting theme counts totals for RC for the cohort year.	IESG cohort mean was the mean of all individual scores for task conflict per IESG cohort.
Effective consensus decision-making results (ECD results or perceived results (PR)	QUAL	During the IPA analysis, the researcher encoded a decision with a result for RFC publication, WG actions, and IETF management.	None
	QUAN	The researcher generated via MAXQDA (2016) the theme count totals for ECD results per recorded minutes (formal and informal).	The individual IESG member's score for perceived results (PR) was a mean of the items on RFCs questions (standard and non-standard), the WG item, and the IETF management item.
		An IESG cohort theme count total for ECD results was the	IESG cohort mean was the mean of all individual scores

Variables	Type of data	Strand 1: Minutes (Reality)	Strand-3: Survey (Perception)
		summation of the meeting theme counts totals for ECD results for the cohort year.	for perceived ECD results (PR) per IESG cohort.
Task independence (control)	QUAL	During the IPA analysis of formal and informal minutes, the researcher encoded subthemes based on the 3 items on the instrument from Van der Vegt et al. (1998).	None
	QUAN	The researcher generated via MAXQDA (2016) the theme count totals for TI per recorded minutes (formal and informal).	The individual member's TI score was the mean of the scores on the three items from Van der Vegt et al. (1998) instrument.
		An IESG cohort theme count total for TI was the summation of the meeting theme counts totals for ECD results for the cohort year.	IESG cohort mean was the mean of all individual scores for perceived TI per IESG cohort.
Cohort year	QUAL	The file name contained the date and the IESG cohort.	The researcher transferred the date from the term question on the survey with the open-ended questions as the IESG cohort.
	QUAN	After transferring all the theme counts from MAXQDA per IESG meeting, the researcher added the IESG cohort year to the file.	The survey term section queried for the IESG cohort years. The researcher's transfer of multiyear responses to single-year responses included adding the IESG cohort year.
OCB-GC	QUAL	During the IPA analysis of formal and informal minutes, the researcher encoded subthemes based on the 3 items on the OCB-GC instrument. The IPA encoding included these three OCB-GC themes under the OCB-GC master theme via hierarchical structure in MAXQDA-11.	None
	QUAN	An IESG cohort theme count total for TI was the summation of the meeting theme counts totals for ECD results for the cohort year.	The individual member's OCB-GC score was the mean of the scores on the 3 items on OCB-GC from the survey.
		An IESG cohort theme count total for OCB-GC was the summation of the meeting theme counts totals for OCB-GC for the cohort year.	IESG cohort mean was the mean of all individual scores for perceived OCB-GC per IESG cohort.
OCB-A	QUAL	During the IPA analysis of formal and informal minutes, the researcher encoded subthemes based on the two	None

Variables	Type of data	Strand 1: Minutes (Reality)	Strand-3: Survey (Perception)
	QUAN	items on the OCB-A instrument. These themes are included under the OCB-A master theme via hierarchical structure in MAXQDA-11. An IESG cohort theme count total for OCB-A was the summation of the meeting theme counts totals for OCB-A results for the cohort year. An IESG cohort theme count total for OCB-A was the summation of the meeting theme counts totals for OCB-A for the cohort year.	The individual member's OCB-A score was the mean of the scores on the two items on OCB-A from the survey. IESG cohort mean was the mean of all individual scores for perceived OCB-A per IESG cohort.

<u>*Data Collection in Strand-4 and Strand-5*</u>

The data collection Strand-4 took the quantitative results from Strands-1–3. The quantitative results from Strand-1 included the results of the descriptive and multivariate statistics (correlation and HRM) that used the theme counts as the base. The standard multivariate suitability tests included scale reliability, and multivariate suitability (normality, linearity, homoscedasticity, freedom from correlated errors), and these tests had different norms since these tests analyzed theme count data. The group behaviors in Strand-1 were theme count totals of individual behaviors; thus, the group's overall behavior was a percentage of the total possible individual behaviors. For example, if 3,000 individual showed HS behaviors detected out of 9,000 individual behaviors in a cohort year over 900 decisions, then the group exhibited HS in 30% of the total behaviors. Strand-2 uploaded the analytical results of the descriptive statistics on effective consensus decision-making, TI, and OCB-GC to Strand-4.[6] Strand-4 received the analytical results from the Strand-3 analysis of the survey questions for the IESG cohort responses and the IETF chair responses. Strand-4 also took the input from the mixed-mode analysis of the survey's IPA analysis of open-ended questions.

---

[6] The revised methodology reduced the statistics collected to effective consensus decision-making results.

Strand-5 collected the results of qualitative analysis from Strands-1–3 and a qualitative summary note from the mixed-mode analysis in Strands-1–4. Strand-1's qualitative IPA produced theme grids and weighted theme diagrams for the themes in the IESG minutes. The researcher's Strand-2's qualitative analysis created a map of the IETF areas from 1986 to 2016, a map of technology progression in WGs per IETF area, and the analytical data regarding the IETF chair's impact on the IESG. The Strand-2's analysis entailed creating a summary note per IETF chair with IETF chair's technology and organizational focus, the IETF chair's accomplishments, and the IETF environment as a SWOT analysis (strengths, weakness, opportunities, and threats). Strand-3's IPA qualitative analysis provided a theme chart with examples of TC and RC occurring during each IETF chair's leadership. Strand-4's qualitative summary note considered if the reduced model fit the quantitative data (solidarity as the predictor, conflict as the moderator, and effective consensus decision-making as the criterion). Strand-5 considered these notes and ancillary data collected from the IETF website on the IETF mission.

Multivariate Analysis Considerations: Sample Size, Suitability, and Method Bias

Multivariate statistics require a sufficient sample size for the statistical tests and data with suitable characteristics for the test. For example, data suitable for multiple regression tests must have normality, homoscedasticity, linearity of data predictors, and uncorrelated errors (Creswell, 2009). The IPA analysis covered IESG minutes for the IESG cohorts from 1991 to 2016 (26 years) with 560 to 900 group decisions per year. The Phase 3 survey response rate was equal to or lower than the 50% rate of the Phase 2 survey (2013). With a 50% response rate, the survey would have 43 participants (40 out of the 81 active IESG members and three out of the five IETF chairs). This section considers the sample size that requires the researcher to switch to the reduced theoretical model from the full theoretical model. The second section considers the potential of the predictors (independent variables) and the criterion (dependent variable) to be suitable for multiple regression statistical tests.

One source of correlated errors in the data relationships is common method biases. This methodology avoided common method biases through the design of the data collection and triangulation of multiple data sources. For example, the survey used the independent variables from self-reported behaviors and the dependent variable for results from online historical records. The survey also queried each respondent (IESG member and IETF chair) for their perception of group effectiveness. The methods for data collection and analysis differentiated between perceptions and historical records. The third section reviews the data collection methodology for potential sources of uncorrelated errors from common method biases such as common rater bias, item characteristics bias, item context effects, and method context.

Sample size. The full and reduced theoretical models involved moderators, so multiple regression statistical tests were used to determine if hypotheses for these models were supported. Pallant (2010) recommended that cases should be 50 plus 8 times the number of independent variables ("N > 50+8m") or "15 participants per predictor" (p. 150) for social science or 40 cases in the sample per independent variable for stepwise regression. Table 21 provides the variable summarization that helps calculate the minimum sample size for the full and reduced theoretical model variables.

The full theoretical model had eight independent variables and tests six hypotheses. The independent variables included two antecedents (HS and VS), two moderator variables (RC and TC), and four moderator factor variables. The three control variables in the model hypotheses were the cohort year, TI, and demographic score. The full model required collecting two additional control variables (OCB-GC and OCB-A) for an alternate OCB model to test the construct validity of the solidarity constructs. These variables replaced the solidarity variables (HS and VS), so the variables did not add to the size requirements.

Table 21: Minimum Sample Size Calculation

	Full model	Reduced model
Summary of hypotheses	6 hypotheses in which 4 hypotheses contain moderator variables	2 hypothesis in which 1 hypothesis contains a moderator
Summary of variables	8 independent variables, 3 control variables, 1 dependent variable	3 independent variables, 1 control variables, 1 dependent variable
Independent variables antecedents	horizontal solidarity (HS) vertical solidarity (VS)	Solidarity (S) = HS + VS
Moderator independent variable	relationship conflict (RC), task conflict (TC)	Conflict (C) = RC + TC
Moderator multipliers	HS x RC, VS x RC, HS x TC, VS x TC	S x C
Dependent variable	Effective consensus decision-making (ECDM results)	Effective consensus decision-making (ECDM results)
Control Variables	Task interdependent (TI) Year of IESG cohort (cohort), Demographics	Task interdependent (TI) Year of IESG cohort
Alternate models Control variables	OCB generalize compliance (OCB-GC) OCB altruism (OCB-A)	OCB = OCB-GC + OCB-A
Multiple regression minimum sample size test-1: > 50+ 8*n-IV	IPA: 115 decisions/year Survey: 115 cohort slots	IPA: 75 decisions/year Survey: 75 cohort slots
test-2: > 15*IV	IPA: 121 decisions/year Survey: 121 cohort slots	IPA: 46 decisions/year Survey: 46 cohort slots
Step-wise regression test3: 40*n-IV	IPA: 320 decisions Survey: 320 cohort slots	IPA: 120 decisions/year Survey: 120 cohort slots

Note. N-IV = number of independent variables.

The full theoretical model based on the eight independent variables required a sample size of 121 IESG decisions per year in the IPA analysis and 121 IESG cohort slots in the survey. If the two control variables (TI), the full model required 136 IESG decisions per year and 136 IESG cohort slots in the survey.

The reduced model had three independent variables, one control variable, and one dependent variable. The reduced model combined the two solidarity antecedents into a single variable (S), an antecedent for the dependent variable (effective consensus decision-making). The conflict (C) variable combined relationship and TC variables into a single moderator independent variable. The moderator factor variable for the equation was SxC, the third independent variable. The control variables were TI and cohort year. With three independent variables, the minimum sample size was 75 IESG decisions per year in the IPA analysis and

75 IESG cohort slots. With the addition of the one control variable (TI), the minimum size was 82 IESG decisions and 82 cohort slots.

The sample size for each strand depended on the maximum number of observations available per data source. Strand-1's cases identified the individual group decisions regarding the unique case for behavioral themes in the formal and narrative minutes. The Phase 1 exploratory IPA results indicated 25 to 40 group decisions per meeting the formal minutes, with 50% of these decisions requiring group collaboration. These results suggested that IESG cohorts made 500 to 900 group decisions per year during ~23 meetings recorded in the formal minutes from 1991 to 2016. The IESG (1992) minutes in 1991 only recorded the biweekly meetings for July through December, so there were only 14 minutes in this year. The 1991 IESG cohort might only have 250 group decisions, but this finding should be enough cases for group behaviors in 1991. Based on Pallant's (2010) recommendation, the IPA and survey must collect a minimum of 115 to 120 cases of group behaviors for the full model or 46 to 75 cases for the reduced theoretical model. The IPA analysis contained sufficient cases of group behavior per cohort year for statistical analysis. The formal minutes existed for 598 IESG meetings, and the narrative minutes existed for 246 formal meetings. The summary of the theme counts for group behaviors to an IESG meeting also contained enough cases for hierarchical regression analysis of the full model.

Strand-2 obtains statistics from four IETF-related websites related to IETF (IETF, RFC Editor, and IANA), the IETF (2016b) Datatracker, email threads, and professional websites (Linked-In). Strand-2 provided an alternate source for the dependent variable's yearly results and confirmed values for control variables (TI and OCB-GC), so Strand-1 triangulation validated these control variables. In addition, the researcher collected demographic data from various websites based on attendance records from Strand-1. Therefore, Strand-2's demographics did not directly impact the data model size.

Strand-3's data came from a resurvey of the IESG members. Suppose Strand-3's resurvey participation rate was 50% of the active IESG cohort members (81 IESG members and five IETF chairs), with 50% of the IESG cohort slots. The researcher defined an IESG cohort slot as the case for the survey because an IESG member could respond once for each IESG cohort. The participation at 50% response rate would be 43 people (40 IESG members and three IETF chairs), 160 IESG cohort slots, and 14 years of IETF chair response. The number of cohort slots at the 50% participation rate supported the full theoretical model, but lower participation rates would fall below the minimum values for the sample size for the full model. The triangulation in Strand-4's quantitative analysis required the same theoretical data model for Strands-1 and Strand-3. Since the 2017 survey required the reduced theoretical model, Strand-4 used the reduced model for triangulation. Appendix F provides additional detail on the sample size requirements for data collection from Strands-1–3.

Linearity and suitability of data predictors. The Phase 2 survey in 2013 found that the self-reported predictors (HS and VS) had a linear relationship with the criterion (effective consensus decision-making results) per IESG cohort year. The Phase 2 research collected the criterion (dependent variable) from online statistics per year, but the researcher found errors in the public statistics. The self-reported predictors (OCB-GC and OCB-A) also had a linear relationship with the criterion variable. The sample size in the 2013 survey was 129 IESG cohort slots and 46 people (41 IESG members and five IETF chairs), which was sufficient to use the full theoretical model for analysis by cohort slots. The researcher conducted statistics tests for multicollinearity or singularity on the Phase 2 survey data for the predictors in the theoretical model, and these tests showed that the predictor did not have multicollinearity or singularity. The predictors and criterion variable data from the 2013 survey had linearity, normality, and homoscedasticity. The Phase 2 research used Pallant's (2010) methodology for statistical tests using SPSS for linear regression, correlations, collinearity diagnostics, outlier detection (beyond three standard deviations), normality and normalized probability plots, and error residuals (P-Plots). Based on these results, this methodology assumed that the statistical analysis could use descriptive statistics and multivariate analysis (correlation and HRM).

The researcher's experience with online statistics for the dependent variable, effective decision-making results (results), raised concerns about errors in the public online statistics. The statistics for these results was the summation of the number of published IETF documents (aka RFCs), the number of tasks to manage WGs, and the number of general IETF management tasks. The original calculation of these results included online statistics per calendar year, but some of these statistics had errors. However, the Phase 2 HRM tests found solidarity behaviors (HS and VS) explained 100% of the variance in these erroneous results. The researcher did not anticipate 100% fit of the model because the yearly statistics did not align with the IESG cohort year. Due to the 100% fit, the researcher investigated the online statistics components of the results and found errors in the data. The researcher also tested the "perceived results" even though these results have common rater bias. Strand-2 of this Phase 3 research contained extra data collection to triangulate the online statistics. Strand-3's survey collected the perceived results scores as part of the check on the online statistics.

The Phase 2 research in 2013 tests comparing the solidarity-based models against the OCB models indicated construct validity of HS and VS for face, content, predictive, concurrent, convergent, and discriminant validity. The Strand-1 and Strand-3 data collections measured the OCB-GC and OCB-A concurrently with HS. The Phase 3 data collection ran the same validity tests as in Phase 2 on the quantitative theme and subtheme counts found in the IPA analysis of the formal and narrative minutes and the survey results.

Potential common method biases. The survey form design avoided the following common method biases: common rater bias, item characteristics bias, item context effects, and method context. Strand-1's data collection avoided common rater bias by using different sources of data rating for the predictors and control variables versus the criterion variable (results of consensus decision-making). The criterion variable based on online statistics avoided common method biases from rater issues and item characteristics effects. There was one exception to this avoidance of common rater bias. The survey responses reported an IESG member's perception of each IESG cohort's effectiveness in making consensus decisions. These scores on the criterion variable had a common rater bias with predictors. The section also had common method biases due to the item characteristics by using a common Likert-like scale with 7 items, item context priming due to inclusion in the same survey, and measurement context effects based on being measured in the same medium simultaneously. As discussed, bias in the IESG perceptions in the Phase 2 survey provided valuable insights when triangulated with the result from online statistics.

The survey avoided item characteristics source bias by creating a survey consisting of tested instruments with good reliability, and the instruments for these variables were on different web pages of an online survey. Published results on these instruments indicated good reliability with Cronbach's alpha of 0.70 or greater. Using statistics for the criterion avoided common method biases from rater issues and item characteristics effects. The survey design avoided item context of test instruments by sequencing the four instruments measuring perceptions to use the following order: solidarity, OCB, TI, and effective decision-making perceptions. The survey used this sequence to avoid priming effects on the predictors and separate the predictors from the criterion.

The Phase 2 research in 2013 raised questions on why the IESG perceptions of the per IESG cohort did not match the online statistical reports. The Phase 2 research did not perform common bias resolution techniques on either dataset (Harmon's single-factor test, partial correlation, or confirmatory factor analysis). However, triangulation between historical sources and the survey did uncover the errors in the online data. The historical sources included the IESG minutes for the year, IETF (2016b) Datatracker details behind the online sources, the IETF online statistics, and the RFC Editor (2016a) statistics. The incident brought up the following questions:

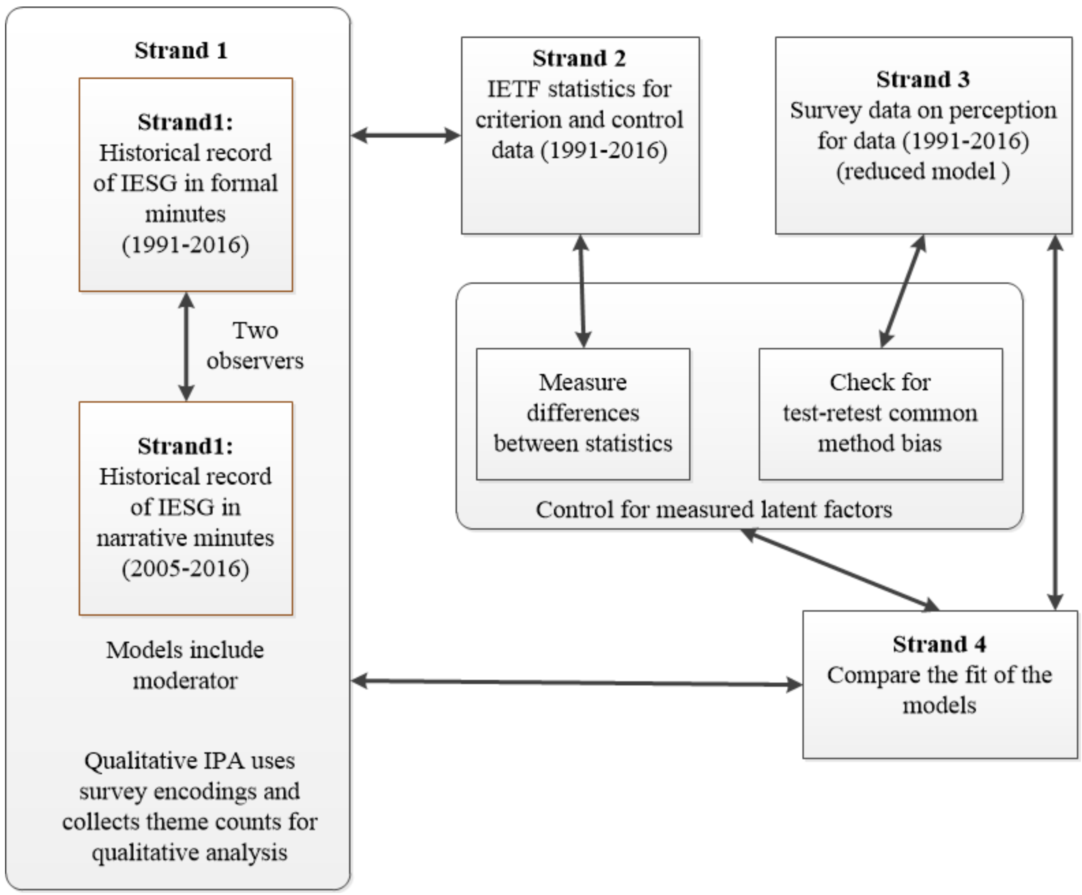
1. Did the survey's common rater bias cause the difference between the IESG cohort's perception of reality and reality?
2. Did the erroneous statistics change the IESG member's perception of reality?

3. Is the difference between perception and reality due to a latent construct such as task or relationship conflict moderated group behaviors?

The online statistics might have changed the IESG perception of reality as post-survey interviews in Phase 2 suggested IESG members relied on online data as a reality check. The interviews with IESG members also suggested conflict regarding tasks and relationships during some IESG cohorts. The researcher considered these three methodology issues that could have caused common method variance when designing the study.

Resolution of common method bias issues. Each strand of the data collection had the potential of unreliable data caused by common method bias or other unmeasured factors. The researcher included triangulation within and between strands to resolve potential bias or errors in each strand. Concurrent mixed-method data collection (qualitative and quantitative) of Strands-1–3 allowed for cross-strand checking of data collected for each strand. The latent construct of conflict was measured and triangulated in Strands-1–3. Strand-4 triangulated the unmeasured methods in Strands-1–3 to examine the data for latent bias or errors. Figure 9 shows the points of triangulating the data within each strand of data and between Strands-1–3. This section reviews how triangulating the data during collection resolves many bias issues and improves reliability.

Figure 8: Triangulation.



Strand-1 had the potential of a single rater observer bias and a single rater bias in the IPA analysis. Two different observers recorded the events into the IESG formal and narrative minutes at two different levels of detail, and the IESG approved both these minutes as accurate. Since most formal minutes from 2007 to 2016 were recorded by one individual, and one individual recorded the narrative minutes from 2008 to 2016, there could be an observer bias in the recorder.

The review by the IESG would usually occur within 2 weeks, but the period could stretch to 8 weeks. Strand-1's IPA analysis was 5% coded in an exploratory process, 95% coded by the researcher, and 10% audited by a peer. During the exploratory process, the researcher coded 5% of the narrative and formal meetings with a purposeful sampling that included formal and narrative minutes from one

IESG biweekly meeting per year, using the survey questions as a descriptive coding guide creating themes based on the survey's scores. From this coding, the researcher developed a codebook.

During the process, peers audited the ratings for 5% of the minutes by coding the IPA analysis. The peers auditing the IPA used the researcher's developed codebook to do an IPA of 5% of the minutes. The peers brought up any concerns or differences in encodings with the research. The three raters discussed differences until the rating have 100% agreement.

After the exploratory phase, the primary researcher coded the remaining 95% of the formal and narrative meetings. After completing the IPA analysis, auditors reviewed 10% of this work. Even with this method to prevent researcher coding bias, coding bias could still occur. These were combined into an "unmeasured latent methods factor" (Podsakoff et al., 2003, p. 891) and the researcher planned to address these latent factors in a confirmatory factor analysis.

Strand-1's mixed-mode results were used to triangulate the data collected within Strand-1 and between Strand-1 and Strand-2. Because Strand-1 collected two parallel records of the biweekly IESG meetings from 2007 to 2016, the theme count data detected should remain the same. Therefore, the researcher compared the theme counts during the IPA analysis. The initial 5% encoding of the analysis included a qualitative comparison of the formats and the content via theme grids, weighted node diagrams, and a quantitative comparison of the theme counts. Triangulation of these minutes aided in creating a reliable view of the historical data in the minutes on the IESG group behaviors (predictor variables) and results of effective consensus decision-making (criterion variable). Two other raters reviewed the researcher's 5% exploratory IPA analysis; thus, these quantitative data checks detected differences in data sources. The researcher checked the data collected in Strand-1 on the criterion variable (results of effective IESG consensus decision-making) by type (RFC, WG, or IETF management) against statistics collected in Strand-2 to determine if both sources are valid.

Strand-2 collected data from the statistics on the IETF website and other associated websites. The data collected from the IETF website included yearly statistics, IETF (2016b) Datatracker records on RFC publications and WG creations, WG webpages, and pages on IETF management (liaison reports, IPR issue, and others). The researcher adjusted these online IETF statistics to IESG cohort years instead of the calendar years. Strand-2 also collected qualitative data on the progression of technology through WG actions (Proposed BOFs, BOFs held, and WG creation, management, and closure) and the IETF chair data. The Strand-2 quantitative data collected were compared to data between different web pages. For example, the researcher compared the yearly statistics on RFCs on the IETF website with the detailed information in the IETF Datatracker or the RFC Editor's website's detailed data. The qualitative data validated the quantitative data. The researcher compared Strand-2 data collected from online statistics to the original historical record in Strand-1 and the perception from the survey in Strand-3.

Strand-3 did not have a common method bias because the statistics collected in Strand-2 provide the criterion variable (results of IESG effective consensus decision-making) and predictors came from the self-reported group behaviors from the survey. However, there was an item context bias in the addition of conflict questions. The triangulation of the Strand-2 statistics with the theme counts from Strand-1 IPA tested the reliability of Strand-2 statistics. Strand-3 used the test-retest comparison between the 2013 survey and 2017 survey answers to compare each IESG cohort's opinion on IESG effectiveness. Strand-3 had a potential for common methods bias due to the item context effects of adding a section for conflict and the measurement context effect of surveying the perceptions of effective consensus decision-making (a criterion comparison value) at the same time as the perceptions of the predictor's variable. Strand-3 used the test-retest comparison between the 2013 survey and 2017 survey to perceptions of all behaviors except for conflict. Conflict scores were compared qualitatively with the IPA of the open-ended questions and quantitatively with the theme counts from the Strand-1's IPA.

A concurrent triangulation methodology in data collection included Strand-1's IPA to obtain theme counts, Strand-2's descriptive statistics to create per cohort statistics, and Strand-3's score calculations for test-retest comparisons. This type of triangulation was used to detect concurrent errors in quantitative data validity, reliability, and common method variance from historical records and the survey.

Strand-4's collected data from the statistical analysis of the quantitative data of Strands-1–3 on the predictors (behaviors) and the criterion (results). The Strand-2 quantitative analysis provided consistent historical data per IESG cohort for the criterion variable (results of IESG consensus decision-making) for Strand-1 and Strand-3 statistical analyses. The researcher also compared this Strand-2 criterion variable against two alternate views of the criterion collected in Strand-1 and Strand-3 to check its reliability. The triangulation of the criterion data from Strands-1–3 in Strand-4 compared three historical sources (IESG formal minutes, IESG narrative minutes, and Strand-2 data) and Strand-3 perceptions to form a composite view of the criterion variable. This composite view helped the researcher to find errors and nuances in the data. The reliability of the criterion variable was vital because the statistical methods in Strand-1 and Strand-3 used the same set of descriptive and multivariate (correlation and HRM) tests to determine whether the research hypotheses were true. Strand-4's triangulation also compared the descriptive statistics on the predictor variables to detect threats to quantitative or qualitative validity or bias. If the researcher found a threat to reliability and validity in Strand-4, the researcher employed the methods described in chapter 2 to address these threats. However, if the researcher could not address this threat via methodological changes, the alternative was to load it as an error factor for analysis.

Data Analysis

The data analysis considered the historical records in the IESG minutes and online sources and compared this data with the perceptions of the IESG members using a series of mixed-mode analyses. The quantitative analyses in Strands-1–4 shared the same data variables based on survey instruments questions. Appendix E contains the codebook for all variables. Strand-1 and Strand-3 used the same set of descriptive and multivariate statistics so that Strand-4 quantitative analysis could compare results. The qualitative analysis in Strands-1–5 analyzed whether the full or reduced data model encompassed leadership behaviors or other latent forces within the TMT interactions. Finally, the mixed-mode analysis reduced sources of common method bias within Strands-1–3 and during the triangulation in Strand-4 and Strand-5.

Strand-1's qualitative analysis used theme grids and weighted node diagrams to verify the validity of the IPA encoding. The researcher generated theme grids and weighted node diagrams generated for peer-checked 5% code-setting IPA and 10% validation IPA. These theme grids and weighted node diagrams for the partial IPA correlated to theme grids and weighted node diagrams for complete IPA (100% for 1991 to 2016). Because the IPA qualitative analysis of the formal and narrative minutes used a mixed-mode encoding scheme with the survey instrument questions as encoding guides, the sum of the theme counts for all questions per scale equated to an instrument scale total.

Appendix G contains the codebook for the IPA analysis. The theme counts from IPA for behaviors found at the individual task level per IESG decisions were summarized to create group behavior values per decision.[7] The data collection in Strand-1 transferred the theme counts totals per meeting for all behavioral variables (independent and control variables), the number of decisions detected, and the results (criterion variable). From Strand-1's theme count data per meeting, the researcher used the SPSS statistical package to generate IESG cohort yearly totals for behavioral variables, decisions detected, and the results.

Strand-1's quantitative analysis used descriptive statistics and multivariate statistics (correlation and hierarchical multiple regression models) after checking for suitability of the data for multiple regressions to determine if the research hypotheses are true. First, the researcher used descriptive statistics to compare theme counts for all behavioral variables, the decision detected, and the results found in the IPA per meeting and IESG cohort year. The descriptive statistics allowed a quantitative comparison of the behavior scales (independent and control

[7] The alternate methodology in Chapter 4 saved the behavioral theme counts per group decision for the 100% analysis of 2015 to 2016 for each type of minutes. The alternate methodology had three types: formal, narrative, and hand-merged (combined) minutes. The 10% analysis of the IESG minutes from 1991 to 2016 uses the per meeting summarization of scales.

variables) and results (dependent variable) between the formal minutes and narrative minutes per meeting and per IESG cohort (per year). The researcher used descriptive statistics to compare the IPA generated theme count totals for the types of results per IESG cohort year with Strand-2's statistics on results. Finally, the researcher ran hierarchical multiple regression modeling tests to determine if the data if the hypotheses of the full model and partial model are supported. This section summarizes the steps in this Strand-1 quantitative analysis of the themes included, and Appendix H contains the codebook for the steps in this quantitative analysis.

The quantitative analysis in Strand-2 examined the variance in data collected from online sources for the dependent variable and the three control variables (OCB-GC, TI, and demographic-score). The online statistical data were used for the dependent variable in HRM tests of the theoretical models in Strand-1 and Strand-3 to avoid common method bias. Phase 2 explanatory research found errors in the online sources, so additional triangulation between the online sources for statistics and the theme counts generated from the historical sources was needed to create reliable data. Strand-2 gathers the online statistical data into categories examined by the survey questions for results (standard RFCs, non-standard RFCs, WG actions, and IETF management actions) and the three control variables per IESG cohort year. The researcher began Strand-2's analysis by using descriptive statistics to compare these statistics between different online sources for each type of data (e.g., standard RFCs). Secondly, the researcher used descriptive statistics to compare these online statistics with the theme count totals collected in Strand-1 to determine any variance or difference per IESG cohort year. A section below summarizes the Strand-2 analysis process, and Appendix I contains a codebook for Strand-2's analysis.

Strand-2's qualitative analysis had two components. The first component was to look at the qualitative data on WGs to determine if IESG as IETF leaders had created a steady progression of technology standards that fulfill the IETF's mission. The content analysis required writing summary notes per area of IETF work and generating a conclusion. The second component was to determine if the

IESG under each IETF chair's management had accomplished the planned advancements in IETF technology and changes in the IETF management. The researcher used content analysis to consider the IETF chair reports in the IETF meetings and IETF organizational changes to determine the success of the IETF chair's plans. The researcher used the qualitative summary of each IETF chair's planned technology and organizational focus in qualitative analysis of the responses from the survey's open-ended conflict questions. The Phase 2 explanatory research indicated that shifts in technology caused conflict and organizational changes cause conflict in the IESG and the IETF. Appendix I contains the codebook for the qualitative analysis.

The Strand-3 quantitative data analysis process had two parts. Part 1 was the quantitative analysis of survey questions. Part 2 of Strand-3's quantitative analysis ran descriptive statistics on the survey's theme counts detected open-ended conflict questions. The steps in Part 2 of the quantitative analysis were with the qualitative data analysis for Strand-3. Survey responses from a single individual may encompass multiple years, so the Strand-3 data collection process translated these multiple-year to single-year responses per IESG member per cohort year and anonymized. The research denoted the single-year survey response from an IESG member as an IESG cohort response and the single-year survey response from an IETF chair as an IETF chair cohort response.

The first step in the analytical process was to generate scale scores for each instrument on the survey that measured perceived values. The perceived value instruments included HS, VS, RC, TC, OCB-GC, OCB-A, TI, and effectiveness of IETF consensus decision-making (PR). The perceived IETF chair effectiveness (chair) was a single value. The second step was to examine the individual scales for outliers and missing data. The third step was to generate the group score for an IESG cohort year by taking the individual scores per year and running descriptive statistics on the individual and group scores. The fourth step was to test the scales from each behavioral instrument for reliability based on the individual responses for behavioral scales (HS, VS, RC, TC, OCB-GC, OCB-A, TI, and PR). The fifth step was to create the reduced model scales (S, C, OCB) and test these reduced theoretical model behavioral scales for reliability. The sixth step was to check the reduced model predictors (S, C), the control variable (TI), and the criterion (results) for suitability for multivariate analysis (normality, linearity, homoscedasticity, independence of error terms, and normality of error distribution). The sixth step also checked the two alternate models (OCB replacement for solidarity and perceived results (PR) replacing results) for suitability for multivariate analysis. Next, the researcher ran correlation tests on the reduced theoretical behaviors, control, and criterion variables used in the reduced model and the alternate models suitable for multivariate analysis. Finally, the researcher ran HRM tests on the reduced and alternate models. The analysis in this research used Pallant's (2010) methodology of using SPSS for each of these steps. Appendix J provides a cookbook for this analysis.

The qualitative analysis for Strand-3's open-ended questions used IPA with survey encodings and discovered themes. The survey encoding used the method for encoding TC and RC instrument questions from the Strand-1 IPA analysis. First, the researcher used the mixed-mode analysis on MAXQDA to generate theme counts per question. Next, the researcher uploaded this theme count data per open-ended question to SPSS and generated scale totals per response. Next, the researcher ran descriptive statistics to generate theme count totals for all responses, each IESG cohort, and the IESG cohorts under one IETF chair. The researcher included these descriptive statistics on the open-ended conflict themes in the quantitative results of the Strand-4 analysis. The encoding of discovered themes in the open-ended questions responses sought to discover additional latent constructs that impact the conflict. This encoding examined the open-ended responses grouped by IESG cohorts under an IETF chair to determine if the conflict exhibited relates to the planned technology focus and organizational changes. Finally, the researcher summarized any discovered themes in Strand-3 qualitative analysis code notes.

Strand 4 combined the quantitative results from Strands-1–3 into an SPSS file for analysis.[8] The quantitative analysis for Strand-1 and Strand-3 used the same descriptive and multivariate statistics methods and included the same statistical data for Strand-2. These same statistical methods analyzed theme-count for behaviors and control variables in Strand-1 and Likert-7 scale values in Strand-3. Similarly, the criterion (results) used statistical counts from Strand-2, and the alternate criterion (perceived results) used a Likert scale value. After appropriate checks for suitability for correlation, the researcher ran correlation tests between the theme count and the Likert scale values for each variable in the reduced theoretical mode per IESG cohort. Additional analysis in Strand-4 was done based on discovered variables in Strands-1–3. Strand-4 qualitative analysis entailed writing up a qualitative summary on the quantitative results

Strand-5 qualitative analysis synthesized the themes, code memos, cohort year summary documents, summary documents per data source (formal minutes, narrative minutes, statistical data, and survey data), and strand analysis summary documents into a combined understanding of the latent forces that influence the behaviors in the IESG TMT per year. Before Strand-5 qualitative analysis, the researcher loaded all themes, code memos, cohort year summary documents, summary documents per source, and strand analysis documents into MAXQDA. Then, the researcher used IPA to discover master themes and create a summary theme charter.

Strand-1 Interpretative Phenomenological Analysis of Minutes

Strand-1's IPA of the IESG minutes progressed through the following stages of analysis: exploratory 5% IPA to create the codebook, IPA of the remainder of formal and informal minutes, mixed-mode generation of counts for themes and subthemes, and validation check for 10% of IPA. The exploratory IPA analysis encoded the themes for theoretical variables according to survey questions plus the Phase 1 variable on collaborative actions and discovered themes in the

---

[8] The alternate methodology described in chapter 4.1 did not allow for a single file analysis due to the reduced scope of the project. Instead. a side-by-side comparison of the results of the statistical analysis from Strand-1 and Strand-3. Since Strand-1 and Strand-3 use the same descriptive statistics and multivariate techniques these results can qualitatively be examined.

text. The encoding allowed the survey question on the positive impact of the IETF chair on the effectiveness of IESG consensus decision-making to be triangulated against the Phase 1 IESG collaborative action. After the researcher established the codebook for encodings, the research encoded the remaining 95% of the sample. Next, the researcher's two peers validated 10% of the encoding. Table 22 shows the original schedule for Strand-1's analysis. This section describes the details of the analysis in each of these phases.

Table 22: Schedule of Strand-1 Analysis

Strand-1	Sample as a percentage of total minutes encoded	Percentage of sample by primary rating	Percentage of sample by two additional raters	Planned Time frame in 2016	*Actual timeframe *1*
Exploratory	5% 42 minutes: 30 formal and 12 narrative	100%	100%	12/1/2016 to 12/5/2016	
All minutes	95% 802 minutes: 568 formal and 234 narrative	100%	0%	12/5/2016 1/30/2017	
Validation	10% 60 formal and 24 narrative.	10%	10%	3/1/2017	

Note. *1- IPA analysis methodology was based on incorrect assumptions. See section 4.1 for a discussion of the research to find a new methodology.

Exploratory analysis (5%). The researcher, during the exploratory IPA analysis in Strand-1, selected 5% of the formal minutes from the 1991 to 2016 IESG cohorts (30 formal minutes) and 5% of the total minutes per year for the narrative minutes from the 2006 to 2016 IESG cohorts (12 narrative minutes). The 30 formal minutes selected by the researcher contained one formal minute from each IESG cohort from 1991 to 2016, plus four additional formal minutes from each IESG cohort from 1992 to 1995. The 5% of the narrative minutes selected by the researcher included one narrative minute from the IESG cohorts for 2005 to 2016.

After reading and rereading the entire group, the primary researcher encoded the formal minutes and the narrative minutes with theoretical themes for each independent, dependent, and control variable using mixed-mode coding based on the survey questions. The full model theoretical independent variables included HS, VS, TC, and RC. The full theoretical model control variables included TI and alternate theoretical construct variables for OCB (OCB-GC and OCB-A). Next, the researcher encoded the dependent variable by tracking: (a) each decision-making action and (b) each decision that resulted in a measurable result (result). The encoding of each decision-making action and result used themes that matched the survey's result types: publishing standards (standard and non-standard RFCs), managing WGs, or managing the IETF administrative functions. The researcher also encoded a control variable for collaborative actions during decision-making based, and this control variable allowed comparison of the Phase 1 IPA with the Phase 3 IPA of the IESG minutes. After encoding all the individual behavioral variables, the research summed individual behaviors into a group theme count per behavioral variable. The primary researcher recorded any comments on trends or themes during this process in code memos.

The last step in the exploratory analysis was to discover themes within the text outside themes the survey encoding defines. The code memos that suggested trends became ideas for discovered themes. The repeating themes the researcher discovered during the IPA process became variables encoded as ancillary control variables for quantitative analysis for the full theoretical model. The researcher summarized the discovered themes into two master themes that could replace solidarity in the theoretical model and a single theme for the reduced model. The encoding process using the MAXQDA (2016) allowed the researcher to store survey-question themes, variable themes, discovered themes, and the discovered master themes so that qualitative analysis could process these data into theme grids and weighted node diagrams. During the 5% exploratory analysis, the researcher created theme grids and weighted node diagrams to determine the validity of the work. Finally, the researcher stored any additional observations on theoretical or discovered variables, oddities, and trends code memos in MAXQDA.

After encoding 5% of the analysis, the primary researcher created a codebook for the rest of the formal and informal minutes encoding. Appendix N contains the researcher’s codebook for IPA analysis encoding. Based on this codebook and 5 hours of training, the two other raters performed the IPA analysis of 5% of IESG minutes (40 formal and 26 narrative minutes). Finally, the three raters met to discuss any differences until reaching acceptable interrater reliability (90% to 100%) on the analysis and the codebook. The researcher saved any observations during the rating discussions as code memos attached to the point in the discussion.

The final step in the exploratory IPA analysis was to generate theme counts for quantitative analysis for behavioral, control, and dependent variables. The researcher used the MAXQDA (2016) functionality to generate theme count totals per IESG minutes (narrative and formal). The preliminary investigations from the Phase 1 exploratory research indicated that the IESG made 25 to 43 decisions per meeting or 560 to 990 in 23 meetings. For the full theoretical model with 10 independent variables (two predictors, two moderators, and four moderation factors) two control variables (TI and yearly cohort), and one dependent variable, the IESG cohort sample size needed to be 120 decisions per year for multiple regression and 400 decisions for step-wise regression. If fewer than 120 decisions existed in 1 year, the researcher switched to the reduced theoretical model. For example, minutes from 1991 contained 9 months, so the total number of decisions recorded in the formal minutes was lower. Appendix F contains a complete list of the number of minutes per year available on the IETF website.

All minutes interpretative phenomenological analysis and validation. The original research plan was to complete the IPA of the remaining 95% formal and narrative minutes using the codebook developed in the exploratory analysis. The researcher generated the theme count totals for all behavioral, control, and discovered variables, plus the subtypes of the dependent variable per IESG minutes. The primary researcher transferred the theme counts from the MAXQDA program to the SPSS program and analyzed these theme counts using the steps described below. The researcher saved additional observations regarding trends as code-memos. Finally, the researcher summarized these code memos in the qualitative summary code memo for Strand-1.

The two additional raters analyzed 10% of the total minutes (60 minutes) to determine if the primary researcher was accurate in her encoding. The researcher selected 60 formal minutes (1991 to 2016) by choosing 2 minutes per year (52 and 8 minutes from 2000 to 2016). The researcher selected 24 narrative minutes by selecting 2 minutes per year from 2005 to 2016. The minutes selected should be gathered spaced by several months. For example, if the 5% analysis selected an IESG minutes in May, the 10% analysis that includes the 5% should choose May and December. The additional raters reviewed the analysis with the primary rater. The primary researcher recorded the additional insights in code memos during these rating discussions. After the primary rating and the 10% checking were complete, the researcher generated (a) themes and subtheme charts, (b) theme grid, and (c) weighted node diagrams for themes. Based on these charts, theme grids, and weighted node diagrams, the researcher wrote up a qualitative analysis on the reliability of the IPA encoding and qualitative analysis of the theoretical models.

Strand-1 Quantitative Analysis of Theme Counts

The Strand-1 quantitative analysis used descriptive and multivariate statistics (correlation and hierarchical multiple regression modeling) to determine if the theme counts found in the historical records in the formal and narrative minutes proved the hypotheses of the theoretical model. As part of the IPA analysis, the researcher transferred the behavior theme counts per meeting for the behavior, control, and discovered variables, plus the subtypes of the dependent variable, into SPSS for analysis. In addition, the researcher uploaded the Strand-2 data to SPSS on yearly cohort statistics per subtype of the dependent variable.

The researcher’s SPSS analysis began by checking each variable for outliers, missing data, and outliers. After checking each meeting’s data, the SPSS analysis used descriptive statistics to create theme totals for all variables per IESG cohort year and years (1991 to 2016). The theme totals per IESG cohort variable were the sum of all meeting behaviors in the cohort year. Finally, the researcher compared the theme totals for the subtype of the dependent variable against the Strand-2 statistics. If there were differences, the researcher investigated the differences.

The researcher tested the behavioral, control, and discovered variables discovered in the historical record to determine if the variables were suitable for multivariate analysis. The researcher used Strand-2 statistics for the dependent variable after validating this variable against the theme count for the dependent variable per year. This check included (a) tests for multicollinearity or singularity of predictor variables, (b) normality tests on the theme counts variables, and (c) tests on relationship between the independent variables and dependent variable to determine if it was linear and had homoscedasticity of error terms, independence of error terms, and normality of error distribution. Additionally, because the behaviors used survey encodings from questions from existing instruments, the researcher conducted scale reliability tests on the theme counts. However, the researcher did not use the scale reliability tests to determine scale reliability because the qualitative tests proved the scale reliability (theme grids and weighted node diagrams). Instead, the reliability tests formed another comparison measure between the theme counts from the formal and narrative counts. After completing the checks on the theme counts per IESG cohort year, the researcher wrote a qualitative memo describing quantitative results and the qualitative observations per year. The qualitative observations for the year were the industry environment, the key players on standardization during the year, and the IESG interactions.

The final step in the quantitative analysis was the multivariate analysis that included correlation tests and hierarchical multiple regression modeling tests (HRM). The researcher conducted the multivariate data analysis separately on the formal and narrative minutes. Based on the suitability tests, the researcher first ran correlation tests between all variables (behaviors, control, discovered, and dependent). Next, the researcher conducted four series of HRM modeling tests to test hypotheses in the two models (full and reduced) on data from the formal and narrative minutes. The first set of HRM tests examined whether the theme counts for behaviors in the formal IESG minutes supported the full model hypotheses or the reduced model hypotheses for 1991 to 2016. The second series of HRM tests examined the theme counts from the narrative IESG minutes for the 2005 to 2016 IESG cohorts. The third series of HRM tests investigated whether the alternate models (from OCB constructs or discovered variable constructs) supported the hypothesis for 1991 to 2016 based on the theme counts from the formal IESG minutes (1991 to 2016) theme counts. The fourth series of HRM tests examined the alternate models based on the theme counts from the narrative minutes from 2005 to 2016.

Strand-2 Analysis of Online Statistics

Strand-2 analyzed data collected from online sources to calculate scales for IESG effective consensus-decision-making actions, task-interdependence, OCB-generalized compliance (OCB-GC), and IETF chair effectiveness per year to determine if the data were valid and error-free. In addition, Strand-2 created an IETF chair effectiveness score and a demographic score. The purpose behind Strand-2's analysis was to provide an alternate source for the dependent variable and two control variables (TI and OCB generalized compliance) to remove common method bias from the full model based on the historical data from the IESG minutes. For example, if the full model regression analysis used the predictors, control variables, and criterion from the theme counts from the IPA analysis, there could have been a common method bias between the predictors and the criterion variable.

After retrieving data from different sources of online statistics for the components of the dependent variable (IESG effective consensus decision-making), the researcher compared the sources using descriptive statistics to determine any variance in the data. The tasks tracked for the effective consensus decision-making per year were the number of RFCs produced per year, the number of WG actions, and the IETF management actions. These data were collected per IESG cohort year and uploaded to the SPSS analysis for Strands-1–3.

Similarly, the researcher planned to retrieved online statistics on TI and OCB-GC.[9] TI behavior was demonstrated within the IESG when two or more IESG members interacted to review documents for RFC publication, write documents for publication, manage WGs, or perform IETF management tasks together. Strand-2's

---

[9] The revised methodology described in chapter 4 had to reduce the statistical data collected in strand 2. The step to query online statistics was dropped in the revised methodology.

quantitative analysis determined what percentage of the total activity per cohort year required TI. Appendix E provides the necessary information to calculate the scales equivalent to the TI scales in the survey for the IESG on an individual IESG member per IESG cohort year. After collecting the data and before analysis, the researcher planned to analyzed the data to determine the outliers and missing data. The descriptive statistics could be used to calculate an average TI and OCB-GC score for a typical IESG member per IESG cohort and all IESG cohorts. The researcher planned to use these average statistics to compare against the IPA analysis theme counts and the Likert-scale used on the survey.

Strand-2 data analysis also created an IETF leadership score and a demographic score per IESG cohort after analyzing online statistical data. The researcher in Strand-2 created an IETF leadership effectiveness score as the ratio of multiple-person interactions with two-person interactions. The research triangulated this Strand-2 leadership effectiveness score against Strand-1's collaborative ratio (single-dyad discussions/multiple-person discussions) and the survey's Likert-7 score on the positive impact of the IETF chair. Strand-2 created the IETF leadership score based on online statistics on collaborative decisions versus dyadic decisions per IESG cohort. The Chair Effectiveness table (Table 98) in Appendix E provides the data variables necessary to calculate the IETF leadership score from statistical data. If there were multiple data sources, the researcher determined why the two sources varied. After selecting the best data, Strand-2 was used to calculate the IETF chair leadership scores on a cohort year basis.

During Strand-2 the researcher planned to create a demographic scale for each year based on the attendance records from Strand-1, plus online data that finds age, gender, and education. The researcher planned to create the demographic scale in an offline process described in the following three steps.[10] First, the researcher combined the attendance data per IESG meeting from Strand-1 with publicly available data on individuals attending an IESG meeting. There were two groups of

---

[10] The alternate methodology removed the demographic score calculation. The reduced scope of the alternate methodology made any demographic scores unwise to use.

people attending an IESG meeting: IESG members and non-IESG members. Second, for each IESG cohort, the demographic score calculated the average age, education, and gender. For all people attending an IESG meeting, the researcher calculated the average age, education, and. Third, the meeting demographic score was calculated as the following:

$$\text{Score} = \frac{\text{IESG (Average age + average education + average gender)}}{\text{Meeting's (average age + average education + average gender)}}$$

All publicly available data on a single person used in this process was considered confidential and kept in a secure offline data store. The researcher checked the demographic data (meeting attendance and personal data) for outliers, missing data, and normality before creating this demographic score per meeting and an average score per IESG Cohort. The early tests showed this ratio less than 0.02 between all years (0.98 to 1.01) for a few meetings in 2015 to 2016. Due to privacy concerns for any demographic data for individuals, even those found in public online results, the researcher only presented this ratio in the results uploaded to SPSS. Since the planned ratio had little variance, this ratio was not considered in the final analysis.

The last step in Strand-2's analysis was a qualitative analysis of the online data and the results of the quantitative analysis. The researcher considered if the online data collected, the analytical results, and the data from Strand-1's historical records matched. The researcher also considered if the quantitative data matched the qualitative data collected regarding WGs and the focus on technology advancement and organizational changes during each IETF chair. The Strand-2 qualitative analysis resulted in the following: technology progression analysis, IETF chair leadership analysis, and validity of online statistics.

Strand-3 Survey Quantitative

The survey data on the IESG member's perception had six parts. Parts 1 through 5 of the survey contained a set of instruments that query for information on the independent, dependent, and control variables. The quantitative analysis of the survey data utilized the 15 steps shown in Table 23 and the codebook for this analysis is in Appendix J.

Table 23: Steps in Strand-3's Quantitative Analysis of Survey Responses

#	Description of step
1	Transfer data from surveymonkey.com to SPSS from the IESG member survey and IETF chair survey.
2	Transform multiple-year responses into single-year responses
3	Remove withdrawn responses
4	Seek missing data or outliers in the IESG member and IETF chair data.
5	Do a reliability check on all instruments using a per IESG response score
6	Calculate per behavior scale values per IESG response (HS, VS, TC, RC, OCB-GC, OCB-A, TI)
7	Calculate average scores per IESG cohort year per scale.
8	Do a correlation matrix per behavior scale to determine the correlation between perceptions of different behaviors.
9	Do a factor analysis on the independent and control variables to determine if the scales have a common factor.
10	Run a multiple regression model to test hypotheses 1 and 2
11	Create the reduced model variables (solidarity, conflict, OCB) from the existing behavior scales per response level.
12	Create IESG cohort average values for the reduced values
13	Check predictor variables for multicollinearity or singularity and the relationship between the predictor and criterion variables for linearity, homoscedasticity (constant variance of error terms), independence of error terms, and normality of error term distribution.
14	Run correlation matrix on the reduced theoretical model variables.
15	Run a hierarchical regression model on reduced mode to determine if hypotheses 1 and 2 are correct.

The IPA analysis of the open-ended conflict questions used a two-pass methodology as Strand-1's methodology. The first IPA encoding pass used the mixed-mode encoding for Jehn's (1995) Conflict Scale. The second coding sought to discover any additional themes in the responses to the open-ended questions. The theme counts from the IPA analysis were analyzed to create alternative scales for TC and RC. Table 24 contains the steps qualitative and quantitative analysis of open-ended questions on conflict, and a codebook for this IPA process is in Appendix K.

Table 24: Steps in Analysis for Open-Ended Questions

#	Description of step
1	Transfer open-end survey from surveymonkey.com to MAXQDA for IPA analysis and do the IPA Analysis. Remove any withdrawn responses.
2	Encode survey-based themes for TC and RC. Seek new themes for discovered variables. Generate theme counts for conflict questions and discovered themes.
3	Transfer theme count totals to SPSS for quantitative analysis.
4	Create the scales per IESG cohort response for TC and RC scales.
5	Compare the scales for task conflict and relationship conflict from open-question analysis) with the survey response answers.
6	Treat discovered variables in the open-ended questions as a new scale, and create a scale value per IESG member. Then, correlate between the discovered scale and the task conflict and relationship conflict scales.
7	Transfer all the quantitative data create to Strand-4.
8	Create node diagrams for the themes discovered in the IPA analysis of the open-ended question, and forward these to Strand-5's qualitative analysis. Write a summary memo per cohort year for the open-ended question analysis,
9	Write up a summary memo for all years for the open-ended question analysis. Include outside data on why the IESG conflict might occur.

The survey instruments for IESG members and IETF chairs used four established scales, a scale on IESG effectiveness in consensus decision-making actions, and demographic questions. The four established survey instruments had proven content validity, construct validity, and predictive validity (Koster and Sanders' (2006) HS and VS scale, Jehn's (1999) ICS scale, MacKenzie et al.'s (1991) OCB scales generalized compliance and altruism, Van der Vegt et al.'s (1998) TI scale.

Furthermore, the Phase 2 survey research in 2013 indicated these instruments had face validity, content validity, construct validity, and predictive validity when used with IESG members. In addition, the Phase 2 survey research found that HS and VS had concurrent validity and discriminant validity as a different construct than the OCB constructs of generalized compliance and altruism. Finally, the survey questions that queried the IESG regarding their IESG cohort's effectiveness in consensus decision-making actions formed an IESG specific scale on the IESG perception. The quantitative analysis of Phase 2 survey questions found these had face validity, content, construct, and predictive validity. In addition, the IESG preception scale results aligned with the online yearly statistics for some years. These five scales (4 established and 1 IESG scale) seemed valid for the context of the IESG survey in 2017, but the researcher planned to conducte a factor analysis on each scale.

Strand-4 Quantitative Triangulation

The researcher in Strand-4's compared the quantitative data gathered from Strands-1–3 for all the independent, control, discovered, and dependent variables. After transferring all the quantitative data into an SPSS program, the researcher correlated the data collected for each independent, control, discovered, and dependent variable. After correlating the data, the researcher planned to conduct seven hierarchical regression analysis cycles using different models to determine if the hypothesis in the full data model and the reduced data model were correct. Chapter 4 describes how the suitability of the data modified the researcher's plans. This section describes the seven planned HRM modeling tests and Appendix L provides a code book the Strand-4 Quantitative Triangulation.

All seven planned HRM models use the online statistical data for the dependent variable score. The first four cycles of HRM modeling use the full theoretical model to test the six hypotheses, and the second three HRM modeling cycles use the reduced theoretical model to test hypotheses. The first HRM test uses theme counts from the formal minutes for the independent variables, control, and discovered variables (1991 to 2016) per meeting. The second HRM test uses the theme counts from the formal minutes for the same variables summarized per IESG cohort year. The third HRM test planned to take the independent, control, and discovered variables from the theme counts from the narrative minutes (2005 to 2016) summarizes per IESG meeting, and the fourth HRM planned to take the theme counts from the narrative minutes for the same variables summarized per IESG cohort. The first four HRM modeling sequences planned to provided answers based on the historical data for the full theoretical model.

The researcher planned to use the final three HRM modeling tests to determine the fit of the reduced theoretical data model variables (solidarity, conflict, TI, OCB, perceived effective consensus decision-making [PR]) for data from the IESG minutes and the survey. The researcher planned to gather the dependent variable from online statistics. The fifth HRM test takes the data for the independent and control variables (per individual response) survey responses summarized per IESG cohort year. Next, the sixth HRM test takes the theoretical reduced model-independent and control variables (solidarity, conflict, TI, and OCB) from the theme counts for the IESG formal minutes (1991 to 2016) summarized per IESG meeting and IESG cohort. Finally, the seventh HRM test draws the data for the independent and control variables from the theme counts from the IESG narrative minutes (2005 to 2016), summarized IESG meeting, and IESG cohort.

Table 25: Steps in Strand-4's Triangulation Process

Context	#	Description of step
Checks	1	Transfer all data from Strands-1–3 to an SPSS file for analysis
	2	Run correlation on the different sources of data per variable
HRM Data models with the full theoretical data model	3	Model 1: Run hierarchical multiple regression (HRM) analysis using the full theoretical data model taking data for the independent and control variables from the theme counts from the formal IESG minutes (1991 to 2016) per IESG meeting, and the data for the dependent variable from the online statistics. Prior to running the hierarchical regression, the following steps are done: a) Test independent variables for multicollinearity, b) Test relationship between independent variables and dependent variables for linearity, homoscedasticity(constant error variance), independence of error terms, and normality of error term distribution,
	4	Model 2: Run HRM analysis using the full theoretical data model taking data for the independent and control variables from the theme counts from the formal IESG minutes (1991 to 2016) averaged per cohort year and the dependent variable from the online statistics.
	5	Model 3: Run HRM analysis using the full theoretical data model taking data for the independent and control variables from the theme counts from the narrative IESG minutes (2005 to 2016) per IESG meeting and the data for the dependent variable from the online statistics.
	6	Model 4: Run HRM analysis using the full theoretical data model taking data for the independent and control variables from the theme counts from the narrative IESG minutes (2005 to 2016) averaged per cohort year. The data for the dependent variable from the online statistics.
Reduced theoretical data model	7	Model 5: Run HRM analysis using the reduced theoretical model taking data for the independent and control variables from the theme counts from survey instrument scores per IESG cohort response and data for the dependent variable from the online statistics.
	8	Model 6: Run HRM analysis using the reduced theoretical model taking data for the independent and control variables from the theme counts from the formal IESG minutes (1991 to 2016) per IESG meeting, and the data for the dependent variable from the online statistics.

Context	#	Description of step
	9	Model 7: Run HRM analysis using the reduced theoretical model taking data for the independent and control variables from the theme counts from the formal IESG minutes (1991 to 2016) per IESG cohort, and the data for the dependent variable from the online statistics.
	10	The researcher compares the survey's perception of the IETF chair's effectiveness and the IETF effective ratio generated from the collaborative theme counts per year via descriptive statistics and correlation (if the data is suitable).

Strand-5 Qualitative Triangulation

The qualitative triangulation considered the themes and subthemes for each variable (independent, dependent, control, and discovered) found in Strand-1's examination of the IESG minutes, Strand-3's examination of the two open-ended questions on TC and RC, and code memos Strands-1–3. In addition, the researcher considered Strand-2 data on technology and organizational changes in IETF areas and the IETF chair's actions. The Strand-5 qualitative triangulation took the constructivist approach. The researcher treated the IESG leadership of the IETF as a TMT within a single year as a phenomenon one observes qualitatively and the 28 years of IESG leadership as a phenomenon with 28 observable phenomena within it. The qualitative triangulation sought to understand the 28 individual phenomena and the 28-year phenomena. Table 26 contains the steps in Strand-5's triangulation process.

Table 26: Steps in Strand-5's Triangulation Process

#	Strand-5 triangulation steps
1	Copy Strand-1 MAXQDA project (data with code and code memos) into a new MAXQDA project. Import Strand-3 analysis of open-ended questions (data with code and code-memos) and Strand-2 code-memos into the new project.
2	Utilize the theme/sub-theme chart utility and theme/sub-theme diagram utility within MAXQDA to find theme charts and diagrams per IESG cohort year.
3	Observe how these themes/sub-theme charts and diagrams change over the years and write up code memos regarding these changes.
4	Compare the themes/sub-themes changes with the changes in the complexity of technologies WGs standardize and the interconnection levels of standards.
5	Compare the themes/sub-themes changes with the IETF management tasks the IESG undertook.
6	Determine if the proposed theoretical models (full and reduced) or some other leadership theory explains the 28 years of yearly phenomena and the 28-year phenomena.

Ethical Concerns

Strand-1 and Strand-2 examined data on public websites and approved for distribution by the IETF. The data from Strands-1-2 were not subject to confidentiality. Strand-3's survey process must disclose the intent of the survey work and indicate the confidential nature of the material. Strand-4 and Strand-5 triangulation of data used Strand-3's data and data from Strands-1-2. Therefore, any publication of the results of this research must take care to maintain the level of confidentiality appropriate for Strand-3 on any combined data. Strand-2 collected publicly available demographic data from public websites to link with Strand-1 attendance data. This demographic data had privacy concerns as it captured demographic information about individuals. This demographic data was treated as private and confidential, even though the data comes from public sources. The researcher kept all data with privacy concerns offline in a private repository.

The survey disclosed the primary and second intent of the research, voluntary nature of participation, the confidential nature of data, and the "first-review" right of the participants in the introduction (see Appendices A and B). The primary use of this research was to aid the IESG, IAB, and IETF nominations committee (NOMCOM). The secondary use provided general research insights on consensus decision-making in leadership teams in voluntary standards communities (e.g., the IETF). Each participant volunteered to participate by checking the box on the survey in Part 1. The participant could have revoked this permission by accessing the survey form code given by the survey to remove the permission. Any IETF chair and IESG member had the "right of first review" on the survey results.

Limitations of this Research Methodology

Limitations on this concurrent research methodology came from the researcher selecting a theoretical model, data collection methods, the data collected, or the data analysis methods. The theoretical data ignored the impact of the TMT member's (IESG member's) leadership on the volunteer leaders of the WGs the TMT member directs. Simcoe (2012) indicated the time to come to a consensus decision on a clear, concise, and technically correct standard by a WG could be impacted by technology complexity (type of RFC) and conflict delays within a WG. The WG leader(s) and the IESG member responsible may reduce distribution conflicts within the WG. The researcher did not include these factors in the research. Due to excluding these data from examination, this survey ignored how quickly WGs produced standards for the IESG to review. The researcher examined the number of standards the IESG reviews in Strands-1 and 2, but the IESG review of a clear, concise standard took less time than a poorly written document.

Each strand had limitations on the data collected, the data collection methodology, and the data analysis methodology. Strand-1 did not collect demographics other than attendance at a meeting, and the demographic score could only be done based on the statistics gathered in Strand-2. Strand-1 data collection method retrieved online minutes with strong historical quality because these records were written at the time of the meeting and approved by all participants within 1 to 4 weeks of the initial recording. The formal minutes provided a list of actions. The narrative IESG minutes described verbal interactions between IESG members, but the narrative minutes were not a transcription. No historical recording provided the perceptions of each member of the IESG as a TMT member. Therefore, the survey was recording perceptions of each member for events that occurred between 1 to 25 years in the past. Therefore, the limitations on IPA that Winston, Fields, and Cabanda (2011) noted regarding the role of language, the suitability of the accounts, and the concern that accounts supply descriptions applied to survey responses. Strand-1 also had a risk in the qualitative analysis because the primary researcher did 85% of the analysis without interrater validation, so the themes and master themes per year might have been in error. This research methodology mitigated this limitation by having three researchers agree on the codebook set after 5% IPA and having a 10% of the remaining 95% checked by two raters besides the principal researcher.

Strand-2 collected statistics regarding IESG consensus decision-making results and demographics. This secondary data source for the dependent variable helped avoid common method bias for the survey and aided triangulation of the effective consensus decision results between the IESG minutes, online statistics, and the survey. However, Strand-2 data collection might not have found all IETF management actions in the IESG minutes. Phase 2 research also found flaws in the online statistics for RFC publications, WG activities, and some management items. The IETF web support team indicated that the best dataset for RFC publication and WG actions was in the IETF (2016b) Datatracker database. Therefore, the researcher needed to search other IETF-related websites for IETF management or demographic data. In addition, Strand-2's data analysis methods were limited statistics on effective consensus decision-making and control variables (TI, OCB-GC, and demographics). Finally, the researcher's Strand-4's qualitative triangulation was used to examine the historical data from Strands-1-2 and the perceptions from Strand-3.

Strand-3 data collection captured survey responses from IESG members and IETF chairs regarding IESG members for the independent, control, and dependent variables. The perceptions of the IESG members had risks to internal validity because of a set of historical events that occurred 1 to 25 years ago. Memories fade because of an individual's maturation or removal from active volunteer status (death or retirement). As memories fade, the perceptions of group behaviors of the IESG fade, especially in memories that were 5 to 25 years old. Strand-3 could also have errors due to the researcher's transformation of multiyear responses to multiple single-year responses. The Strand-3 data analysis used scale reliability tests to determine scale validity and descriptive statistics to provide comparative results per IESG cohort year. If the survey data was suitable for multivariate analysis, the correlation and HRM modeling tests would determine if the theoretical models explain the data variance in the results per IESG cohort year. Data suitable for multivariate regression must have a normal distribution, homoscedasticity, linearity, and an absence of correlated errors. Strand-3's IPA of the open-ended questions on conflict used the same two-pass methodology as Strand-1. Pass-1 for the analysis encoded survey questions on TC and RC. Pass-2 looked for new themes to discover. The responses to Strand-3's open-ended conflict question might not have contained data that with task and relationship survey question themes or new themes to discover.

Strand-4's analysis methods triangulated the quantitative results from the historical data in recorded data in IESG minutes and IETF statistics with the perceived data recorded in the survey responses. The researcher used descriptive statistics, scale reliability tests, and multivariate statistical tests to triangulate the IESG group behaviors and results viewed from historical records and perceptions of the IESG members and IETF chairs. Any errors within Strands-1–3 and Strand-4 data collection methodology could have impacted this triangulation. For example, if the theme counts from the IESG minutes lacked a normal distribution, homoscedasticity, and linearity, the researcher could not use multivariate statistics (correlation and HRM modeling tests) in the triangulation analysis.

Strand-5 analyzed the qualitative results Strands-1–3 and the qualitative summary from Strand-4. The primary researcher examined the qualitative data from Strands-1–4 via IPA to find master themes and cycles of content analysis to triangulate the data. Only the primary researcher did the Strand-5 analysis. One weakness in the Strand-5 analytical methodology was the lack of validation checks from inter-raters.

The researcher asked past IESG members and IETF chairs to review the results from Strands-1–5 and the final interpretation during the interpretation phase. This broad review provided feedback on the triangulation in Strands-4–5 and the interpretation of the results. One limitation of the interpretation phase was that the researcher did not consider IETF data on WG conflict (e.g., mail list reviews) in the interpretation. Interpretation of Strand-4 quantitative results and Strand-5 qualitative results considered past research from Simcoe's (2007, 2012), Gençer's (2012) analysis, and other research into the IETF and IT SDOs.

Table 27: Limitations of Research Methodology Specific to the Strands or Interpretation

Strand	Unexamined model data	Data collection methodology	Data Analysis Methodology
Strand-1 – IESG minutes	Strand-1 does not examine the demographics other than attendance	The recording of narrative minutes only covers only three biweekly sessions in 2005 and most of the	IPA analysis is limited by the role of language, suitability of the account, and accounts that explain

Strand	Unexamined model data	Data collection methodology	Data Analysis Methodology
		IESG biweekly sessions in 2006 to 2016.	rather than describe perceptions. Primary research does 90% of the analysis without interrater validation.
Strand-2 – statistics	Strand-1 only looks at demographic statistics and consensus-decision-making statistics. Some IETF management actions are found only in the IESG minutes.	The Phase 2 research found flaws in the online statistics of RFC publication, WG actions, and IETF management actions.	Analysis limited to demographics and effective consensus decision-making components.
Strand-3 – survey	Strand three examines only the perception of the IESG members. Maturation	Mortality of IESG members means some IESG are dead or no longer responding to IETF email. Maturation of the participants means the perceptions of how an IESG TMT interacted in a past IESG (1-25 years ago) is foggy in the recollections of some IESG members, so the member may not respond. The researcher transfers survey data from the survey website (surveymonkey.com)	Strand-3's quantitative analysis may suffer from missing or errors data. Strand-3's IPA analysis may have accounts that explain perceptions rather than describe perceptions. The primary research does 90% of the analysis without interrater validation.
Strand-4 – Quantitative triangulation	None. All variables considered.	The researcher uploads quantitative thematic counts to SPSS for comparison.	Multivariate statistical analysis requires data with normal distribution, linearity, homoscedasticity, and uncorrelated errors.
Strand-5 – Qualitative triangulation	None. All themes related and code-memos related to variables and discovered themes are analyzed.	The researcher loads qualitative code memos from Strands-1–3 into MAXQDA. Master theme analysis uses theme encoding.	The primary researcher does all analysis, so no interrater validation of Strand-5 occurs.
Interpretation	The researcher did not consider any IETF data on WG conflict in interpretation.	Additional IETF statistics on the RFC review rates and WG creation rates of specific IESG members will be collected and considered.	The researcher considers Simcoe's (2007, 2012) data and Gençer's (2012) and quantitative and qualitative results.

Planned Schedule and Budget

The planned schedule for this research began on December 1, 2016, and ended on July 17, 2017, as shown in Table 28. The data collection stage began with the online data sources on December 1, 2016, for Strand-1 and continued with the survey research beginning on March 3, 2017, continuing through June 3, 2017. The researcher delayed the start of the survey to present the survey project to the 2016 to 2017 IESG. Therefore, the analysis in Strands-1–3 operated concurrently from December 1, 2016, to June 30, 2017. Strand-4 and Strand-5's triangulation took place from July 1, 2017, to July 15, 2017. The interpretation took place from July 15, 2017, to August 5, 2017. During the IETF physical meeting on July 13 to 19, sessions were scheduled with the current IESG and past IESG members to review the final results.

Planned Timeline/Budget

The timeframe for the data collection and analysis was 12/1/2016 to 7/15/2017, and the interpretation phase lasted from 7/15/2017 to 8/5/2017. The long interpretation phase allowed for a staged review of the data by IESG members who were respondents, then the current IETF TMTs (IESG and IAB), and then a broader audience in the IETF. The total budget for this research was $5,820, but only the $5,500 cost to fund inter-raters was expended by the primary researcher.

Table 28: Original Schedule and Budget for the Project (2016 to 2017)

Strand	Data collection	Analysis	Personnel	Equipment	Budget
Strand-1 – minutes	12/1/2016 to 4/20/2017 *2	12/1/2016 to 6/30/2017	Primary research + 2 additional coders	MAXQDA-11	$245 *2 $5000 inter-raters
Strand-2 - statistics	2/1 to 2/28	1/1/2017 to 6/30/2017	Primary researcher	SPSS	$105*2 Already acquired
Strand-3 – survey	4/3/2017 to 6/3/2017	6/4 to 6/6/2017	Primary researcher + 2 additional coders	Survey-monkey site, SPSS, email, MAXQDA-11,	$85/month for survey-monkey*2, $15/mail*2, $500 raters
Strand-4 – Quantitative triangulation	7/1	7/1/2017 to 7/15/2017	Primary researcher	SPSS, MAXQDA-11	Already required
Strand-5 – Qualitative triangulation	7/1	7/1/2017 to 7/15/2017	Primary researcher	SPSS, MAXQDA-11	Already required
Interpretation		7/15/2017 to 8/5/2017	Primary researcher	SPSS, MAXQDA-11	Already acquired
Totals	12/1/2016 to 6/30/2017	12/1/2016 to 7/15/2017	Primary research + 2 additional IPA coders	Email, web-site, survey-monkey site, SPSS, MAXQDA-11	$5820

*1 – Final 2016 IESG minutes were not approved and loaded on website until 4/20/2017.
*2 - Acquired by the primary research prior to research.

Conclusion

This research supported past organizational research from Dess and Origer (1987), Koster and Sanders (2006), Jehn (1999), and Van der Vegt et al. (1998) to connect their insights on TMT processes on consensus decision-making. The researcher linked the insights from these leadership theories on TMT to the volunteer leadership of the IETF as an ICT standards body. The stakes were high for the global ICT industry if the volunteer leadership slowed down the pace of consensus decision making, causing delays in Internet standards publication and causing delays in introducing new ICT technologies by firms. The minor delays in TMT volunteer consensus decision-making were like the loss of a horse’s shoe in the tale of how the loss of a horse's shoe caused the horse to stumble and the rider to fall in battle, which caused the war to end and be lost. This researcher sought to understand antecedents of consensus decision-making that would cause delays in volunteer organizations so that these insights could help TMTs in other volunteer organizations (churches, political organizations, and relief organizations) make consensus decisions.

Chapter 4 – Results

Surprises during a research project are challenging. This researcher was surprised by the reality that IESG (2000), the TMT of the IETF (2021b), operated in a virtual environment, making continuous consensus decisions in an open environment while sharing many details in online databases. The reports on biweekly meetings or statistics only summarized a portion of these online databases. Thus, the researcher made substantial methodology changes to Strand-1 and Strand-2 of the research methodology, and these strands caused some methodological changes in Strands-3–5.

The beginning of Chapter 4 reviews the discovery of the need for methodology change, the research to determine needed changes, and an overview of the alternate methodology. Appendices A to C and E to N contain the revised codebooks for the alternate methodology. The alternate methodology used historiometric best practices suggested by Ligon et al. (2012) to create quantitative data from IPA analysis of the historical data in the IESG minutes (Strand-1 and Strand-2) and the open-ended questions in the survey (Strand-3). No single online source provided statistics for Strand-2 based on a validated list of IESG actions and results per IESG cohort; thus, the alternate methodology created the validated list of results and actions per IESG cohort before creating the Strand-2 counts of actions and results found in Table 53 below. The researcher validated each action and result using online sources, such as IETF (2016b) Datatracker, storing the confirmation in code notes. The researcher used mixed-mode IPA analysis on IESG minutes plus code notes to detect behavior themes (Strand-1) and validated group actions and results (Strand-2).

Due to the extra effort required by data collection altered methodology, the researcher reduced the scope of Strand-1 research to a 10% sample of the IESG minutes from 1991 to 2016 (78 minutes) and complete examination of the IESG minutes from 2015 and 2016 (55 minutes). The Strand-2 actions and results were estimated from counts obtained from the IPA analysis of the 10% sample of IESG minutes and cross-checked for 2015 to 2016 counts based on the 100% analysis of the IESG minutes from 2015 to 2016. Appendix O contains the list of validated actions and results for the 10% analysis, and Appendix P contains the list of validated actions and results for the 100% analysis of 2015 to 2016. Chapter 4 also contains a summary of the results for Strands 1-4. Historiometric mixed-mode practices require detailing both qualitative and quantitative results so researchers can evaluate both. Details on the IPA mixed-mode analysis are in Appendix O for the 10% sample, Appendix P for the 100% sample, and Appendix Q.2 for the IPA mixed-mode analysis of the 2017 survey's open-ended questions.

The Strand-3 survey received only 29 responses (25 IESG members and 4 IETF Chairs). The cause of this limited response appears to be the aging of IESG members and a misunderstanding regarding the 2017 survey. For example, some IESG members believed that if they had taken the survey in 2013, the survey in 2017 was unnecessary. These two issues seemed to have caused a limited number of IESG members to respond to the 2017 survey.

Chapter 4 also discusses the survey data from the 2017 survey and the 2013 survey sample size, descriptive statistics, scale reliability, statistical analysis (correlation and HRM modeling). Chapter 4 provides the results from Strand-4's quantitative analysis of the quantitative data from the historical IESG minutes (10% sample from 1991 to 2016 and 100% sample from 2015 to 2016) and the quantitative data from two surveys (2017 survey and 2013 survey), and the open-ended conflict questions (2017 survey). Due to the number of survey responses, the researcher could only use the reduced theoretical model in the Strand-3 survey analysis and the Strand-4 triangulation. The Strand-4 analysis showed that Hypothesis 1 of the reduced model was supported, but Hypothesis 2 on moderating effect of conflict was not supported. Conflict did correlate to results and perceived results, but conflict does not have a moderating effect. The validity risks and bias for the Strand-4 results are discussed, with the chapter providing a qualitative analysis of the quantitative data results.

The second surprise in this research was that the alternate methodology resulted in a generally consistent qualitative and quantitative analysis view. The Strand-1 qualitative themes detected the same information as queried by survey questions. The researcher discovered common themes in the 10% sample (1991 to

2016) and the 100% sample in 2015 to 2016. The analysis of the open-ended survey questions showed conflict themes based on Jehn's conflict instrument questions (Strand-3). Two top-level themes were unique to the open-ended portion of the survey. The first unique top-level theme was conflict moved beyond the conflict between individuals on the team to the conflict of factions that tried to replace or take the place of the IETF chair as leader of the IESG as a TMT. Although attempts to remove a leader can contain task and emotional conflict (tension, friction, and anger), this conflict goes beyond the conflict considered by Jehn's (1999) ICS instrument. The second top-level theme involved resistance to change within the group or change within the marketplace influenced by IETF as a SDO among other SDOs. The discussion on Strand-5 in the qualitative analysis of the qualitative data from Strands-1–4 includes a discussion on internal and external resistance to change.

The Strand-5 analysis examines the IESG as making consensus decisions as a TMT of an SDO dedicated to technology change in the IT industry. The IETF goals are to move the IT industry steadily toward Licklider's vision of a "Galactic Network" (Leiner et al., 1999/2003, p. 23). The behaviors of solidarity, conflict, and TI in the IESG TMT during the consensus decision process are qualitatively effective if they help create standards that migrate the IT industry via their decisions on management items (documents published, work selected, and organizational management) and organizational changes. The IETF nominating committee had selected IETF chairs to accomplish particular focuses, so this qualitative analysis examined the IESG decisions per IETF chair. The analysis of the open-ended questions in the survey were used at the same time. The qualitative analysis supported conclusions from the quantitative analysis that Hypothesis 1 of the reduced theoretical was supported, and Hypothesis 2 was not supported. Conflict as a combined construct did not have a moderating effect in IESG interactions in the complex environment of an SDO. The IETF chair's leadership was a critical part of increasing solidarity, decreasing RC, and increasing the helpful type of TC to increase the effectiveness of consensus decisions.

Methodological Changes

The researcher based the original methodology of Strands-1–3 on two assumptions about the IETF data that were incorrect. The first assumption was that the two types of minutes for the IESG contained enough individual behaviors to accurately determine group behaviors on conflict and solidarity. The second assumption was that IESG decisions and yearly statistics on documents published, WG actions (creation, management, or closure), and IETF actions would align within some error. These two assumptions were incorrect due to the impact of the virtual environment that IESG members existed in as part of a virtual team from 1991 to 2016. Since the inception of the IESG within the IETF in 1989, the IESG has operated in a virtual environment with computer-assisted communication across multiple channels (secure email, public email, video conference, and messaging streams) that provide computer assistance for their decision making. The individuals who have served on the IESG are technology-savvy individuals who utilize the tools for virtual communication and decision-making to their fullest extent. The sophistication of this virtual environment has increased steadily since 1989 as the computer network has increased its capabilities.

The compelling pull for refining the research methodology rather than abandoning this research was two-fold. First, geographically distributed TMTs were common in global corporations and international volunteer organizations in 2020. The examination of continuous virtual consensus decision-making would add to the knowledge of consensus decision-making. Determining if a theoretical model correctly predicted how solidarity and conflict interacted within a virtual TMT with continuous actions could help TMTs in global organizations. The second compelling reason was that the IETF had continually improved online data records as an open-standards forum. This evolving and improving source of historical data provided over 26 years of historical data.

This section describes how the researcher discovered problems with the original methodology during the codebook development for qualitative analysis. Table 29 summarizes these investigations into the methodology, the result of those investigations, and the changes to the methodology. Table 29 also reviews the potential risk of errors in the investigation. The researcher resolved the methodological problems by increasing the details and amount of online data. Due to this increased amount of data collection for Strands-1 and 2, the researcher reduced the scope of this research. The researcher reduced the scope of the IESG minutes examined to 10% of IESG minutes from 1991 to 2014 and 100% from 2015 to 2016. The researcher adjusted the methodology for Strand-4 and Strand-5 to fit this reduced scope. Due to this reduced scope, the researcher left the detailed analysis of the IESG behaviors during consensus decision-making that utilizes all IESG minutes from 1991 to 2014 to future research.

Table 29: Research Into Methodological Issues (Strands-1–5)

Strand	Research to refine the problem	Result changes	Test of changes	Risks of error in results
Strand-1	Initial research: IPA of 5% of Minutest to set IPA codebook Step 1: IPA on the format of minutes Step 2-3: IPA analysis of online data for data on decisions, group behaviors, and individual behaviors Step 4: 10% of minutes Analyzed Step 5: 4 months in July	Initial suspicion on theoretical assumption 1 1) Hand alignment of validated formal and narrative minutes. 2) Inclusion of online data on decisions, group behaviors, and individual behaviors in code notes and encodings 3)Better methods to record validated actions and behaviors 4) Improved transfer mechanisms for mixed-mode themes and established discovered themes.	Five steps to refine research methodology Test with 10% of minutes and four months of minutes in 2016	5% of data chosen is an outlier 10% of minutes and four months might have a significant amount of outliers in data
Strand-2	Initial Research: Gather statistics on documents and Working Step 1: Correlate documents approved with RFC Editor's (2016a, 2016b) publications for 10% of data Step 2: Correlate WG approved in 10% of minutes with 10% of existing WGs Step 3: Correlate WG approved in 10% of IESG minutes + 1 BOF call with 10% of existing WGs	Initial data caused suspicion on theoretical assumption 2. Added tests to quantify errors. 1) Include BOF Call minutes (pseudo or actual) in IESG minutes 2) Collect validated actions on an individual action basis, store data in code notes, and encode the validated actions in action themes 3) Create results file from code notes and themes, 4) Use Theme counts for IPA mixed-mode and Strand-2 5) Refine the methodology for Strand-2 after each test	Run added tests on 5% and 10% of data Test with 5%, 10%, 4 months of 2016, All of 2016 and all of 2015	Online data errors Test methodology errors
Strand-3 IPA	A small number of open-ended responses	Alter Mixed mode comparison to link to periods of IETF chairs	Tests None	Small sample size

Strand	Research to refine the problem	Result changes	Test of changes	Risks of error in results
Strand-4	1) Reduction of scope requires split analysis of a) 1991 to 2016, and b) 2015 to 2016. 2) 2016 Survey had few repeat respondents from the 2013 survey	Alter methods of Strand-4 to handle split analysis, and survey	The researcher does an initial run on all data prior to the final run	Split analysis may miss critical data
Strand-5	Reduced scope focuses on 3 summaries: a) 1991 to 2016 b) 2015 to 2016	The single encoding scheme for all data and summaries	Test analysis of portions of 3 summary documents.	Differences in the level of detail on data

Discovery of Internet Engineering Steering Group Processes Differences Than Assumed Process

The discovery of the impact of the virtual environment came after the initial phase of the qualitative coding. The initial phase of IPA analysis for Strand-1 called for codebook creation for the qualitative source by examining 5% of the narrative and formal minutes in combinations with two other reviewers. A panel of three judges formed by the primary researcher and two experts rated the coding within a section.

The initial code analysis found the following issues: (a) high levels of technical jargon, (b) variance in the data kept by the formal minutes during 1991 to 2003, (c) variance in data kept by the narrative minutes in 2005 to 2009, (d) variance in data kept between the formal and narrative minutes, and (e) variance in data in the IESG minutes hindered the discovery of new variables. The high level of technical jargon required additional discussion during the initial reviews of judges of the minutes from cohort years 2011 to 2016. The experience of the primary researcher was able to translate the technical jargon reliably so that the raters could judge the initial encoding of themes relating to group behaviors and group actions in Phase 1 IPA analysis. The group behaviors included solidarity, conflict, OCB, and TI based on the survey questions. The group actions related to effective consensus decision-making included document review and publication, WG creation, management, closure, and IETF management. The group behavior examined linked IETF chair effectiveness to the level of collaboration on decisions during a leader’s tenure. Appendix N contains the IPA encoding codebook.

After the researcher set the initial codebook by considering the narrative and formal minutes from 2011 to 2016, the three raters evaluated the codebook by analyzing 5% of narrative and formal minutes going backward from 2016 to 1991. The researcher selected 5% of the narrative and formal minutes by selecting one of the IESG biweekly minutes per year plus some additional minutes. The researcher selected the minutes each year for the IESG meeting closest to the 31st of May (in late May or early June). During this process of checking the codebook, most interrater differences and discussions centered on identifying different types of conflict that extended beyond TC with differences of opinion. Group discussions among the inter-raters focused on determining if individual actions in the group consensus decision-making process either resolved conflicts via solidarity efforts or accelerated the TC into RCs. The analysis by the interrater group found that the data contained in the narrative minutes 2009 to 2016 provided enough information to determine individual behaviors within the group behaviors that resulted in a decision within a meeting or multiple meetings. Some IESG group decisions did not reach a consensus during a single meeting but reached a consensus between meetings. The narrative IESG minutes did not record between meeting decisions.

The researcher began to suspect a virtual environment with continuous decision-making based on the characteristics of the narrative and formal minutes from 2009 to 2016. The narrative minutes in this period took a snapshot of the comments on documents in the IETF Datatracker, and from reading these minutes, it appeared that the document review and comments were continuous. The meetings seemed to serve as communal decision points but not the only decision points. The formal minutes examined from 2009 to 2016 recorded the result of decisions between meetings, which indicated that consensus decisions were final whenever the IESG reaches a consensus. The formal and narrative minutes also referred to the online IETF Datatracker records with IESG reviews of proposed WG charter changes for new or existing WGs. Unfortunately, neither the formal minutes nor the narrative minutes preserved a copy of this information in an appendix.

The analysis of 5% of the minutes from 2009 to 2016 caused the researcher to consider two questions about the reality of the IESG decision-making process. The questions included the following:

1. Is the virtual environment an artifact of the 2009 to 2016 minutes, or did the IESG's consensus decision-making operate in a continuous virtual decision-process from 1991 to 2016?
2. Are the formal and narrative minutes simply two views on the same online data?

These questions directly impacted the data collection and analysis in Strand-1 and Strand-2. In addition, the analysis in Strands-3-5 methods, which rely on the data collected in Strand-1 and Strand-2, are also impacted. The researcher began investigating these questions in Strand-1's data, leading to an alternate methodology for Strand-1 and investigations into Strand-2. Strand-2 investigations led to an alternate methodology for Strand-2. The alternate methodologies for Strand-1 and Strand-2 led to modifications in the methodology for Strands-3-5.

Strand-1 Investigations

This section describes how the researcher took four steps to investigate whether the formal and narrative IESG minutes contained all the behaviors involved in the IESG consensus decision-making or the minutes only pointed to online records that had the behaviors. The first step in this investigation was to determine the differences in the data contained in the formal and narrative minutes. This step used IPA to look at the file formats and the content. The second step in the investigation was to determine if the minutes referred to a more detailed online dataset from 1991 to 2016. The third step was to determine whether the total data contained in both minutes and online presented enough data on individual and group behaviors to continue the original research. Finally, the fourth step used this altered methodology to analyze the 10% of the minutes from 1991 to 2016, and then 4 months of 2016. After adjustments, the researcher used this altered methodology for all the minutes from 2015 to 2016.

The altered methodology collected additional online data into code memos linked to the IESG minutes. After encoding the appropriate themes for individual behaviors and validated actions into this combined data, this altered method generated group theme counts and transfers the theme counts to SPSS for quantitative evaluations. In addition, the methodology forwarded a list of results, a list of management items, and code notes with per meeting summaries to the qualitative evaluation of Strand-1.

After the entire year of data was analyzed, the process was analyzed to determine if any refinement was necessary to reduce errors. If any refinement was necessary, the corrections were made to the IESG year under consideration before analyzing the following year. This iterative cycle of refinement and test had caused a substantial extension of the time required to triangulate the online data with the survey. Table 30 summarizes the timeline and results of the investigations to refine Strand-1. The only issues related to stand-1 IPA analysis found in the yearly reviews for 2015 and 2016 were IPA codebook additions for behaviors during actions not found in 10% analysis or eight meeting analysis. An example of such an action was the reassignment of ISE documents to IESG members by the IAB chair. The section on the investigation to create a new methodology for Strand-1 and Strand-2 describes how this methodology was confirmed.

Table 30: Summary of Timeline and Results

Strand	Research to refine the problem	Result changes	Test of changes	Risks of error in results
Strand-1	Step 1: IPA on the format of minutes Result: Formal and narrative minutes refer to online data with overlap data. Therefore, uniqueness may be a result of the function of the minutes.	1) Use code notes file to align formal and narrative minutes 2) Validate behaviors in formal and narrative minutes.	How to test: Analyze 10% of the data using the new methodology	Step 1 Risks 10% of data selected may not represent the other 90% file format
	Step 2: Analysis on IESG decisions available in the online sources	3) Validate actions against online data regarding IESG decisions and store data in code notes.		
	Step 3: Determine if the online data regarding each decision has behavioral information	4) Copy online data on behaviors to code notes and encode themes in minutes.		
	Step 4: 10% of minutes analyzed with notes file and validation of behaviors Result: Alternate Strand-1 works with 10%. Virtual continuous decision-making assumption is more likely than the previous assumption.	5) Create Strand-1 Alternate methodology with code notes, theme counts, and interim data files in order to transfer data to SPSS	1) Test with four months of IESG minutes (formal and narrative) from 4/21/2016 to 7/7/2016 2) Use Alternate Strand-1 methodology to discover themes	Step 2 risks 10% of the data selected may not represent the other 90% of the data.

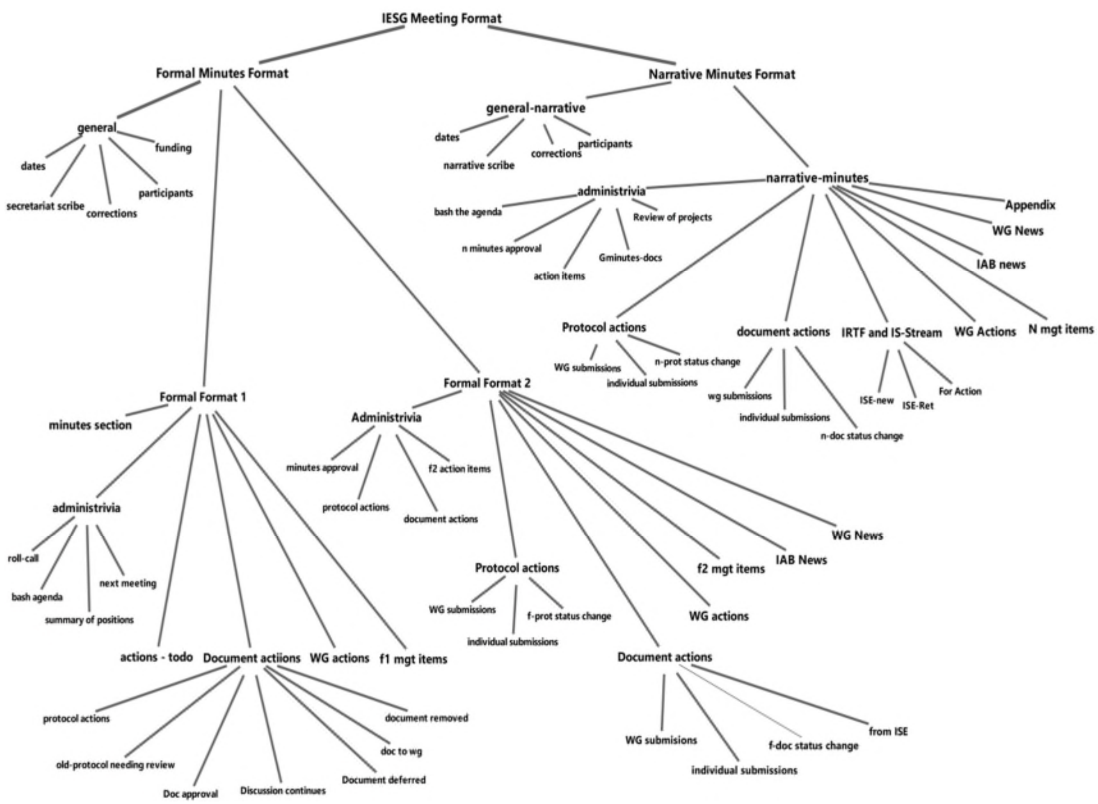
The continuous virtual nature of the IESG decision-making also caused a revision of Strand-2's methods. RFC Editor (2016a) statistics on document publication and IETF data on WGs is accurate for publication statistics. However, the accuracy did not link to the IESG's decision to approve the publication of documents. The discussion on Strand-2 methodology changes section below describes why there are variable delays between some IESG decisions to approve documents publication and publication of document; and between the some IESG action for WG establishment and the reporting of the WG meeting. For example, the approval of a document may have post-approval editing delays before being sent to the RFC Editor and variable delays at the RFC Editor before publication.

Similarly, WGs might have variable delays caused by finding chairs and final edits on charter for new work. The researcher who sought to obtain accurate results must validate each IESG action and result. The original mixed-mode analysis of the Strand-1 data planned to minimize error to the bias of minute takers by comparing it against general statistics. Because there was no online list of validated actions for each IESG, Strand-2 data collection needed to occur regarding each event. The alternate research methodology tied the validated actions (result causing or non-result causing) to the IPA theme encoding.

Interpretive phenomenological analysis on the formats in minutes. The researcher used the IPA process to discover the formats of the formal and narrative minutes. Figure 9 shows the first five levels of the themes under the IESG meeting format theme found in the IPA qualitative analysis of the formal minutes for IESG cohorts 1991 to 2016 (July 1991 to March 2017) and the narrative minutes for cohorts 2005 to 2016 (September 2005 to March 2017). An examination of the levels of themes points to a virtual environment occurring in the IESG cohorts from 1991 to 2016. This section describes how these five levels of themes illustrate the virtual environment.

The topmost level of these summary themes indicated two formats in the formal minutes and one format in the narrative file. The formal minutes used one format (Format-1) from July 1991 to May 2003 and a second (Format-2) after May 2003. This change in format matched the changes between groups of recording scribes for the formal minutes during 2003. After 2003, a small set of scribes from the IETF secretariat had recorded the formal minutes with guidance by the IETF chair. During some IESG meetings from 2015 to 2016, the IETF secretariat asked the IETF chair what should go in the minutes, indicating the IETF chair was guiding what went in the minutes.

Figure 9: Theme node diagrams for Internet Engineering Steering Group meeting format.



Levels 3 to 5 of the weighted node diagrams illustrate the commonalities and differences in the minutes. The third level of the format for formal and narrative minutes contains a general description of the minutes and report on the meeting ("general" and "general-narrative"). The data gathered in the general section included the date of minutes, scribe authoring the minutes, participants, and corrections to the minutes. Format-1 of the formal minutes also contained reported funding sources for generating the IESG minutes in the general section. This recording of funding sources was appropriate for the period during which the National Science Foundation was funding IETF administrative costs.

The fourth level of each of these formats contained themes to identify administrivia, documents actions, WG actions, and IETF management ("F1 mgt item," "f2 mgt item," and "*N* mgt items"). Format-2 of the formal minutes and the narrative minutes' format split the document actions into protocol actions and document actions to differentiate between IETF protocol standards and IETF informational documents. The "action-todo" list was at Level 4 in Format-1 of the formal minutes and often occurred at the end of the minutes. The other two formats (formal Format-2 and narrative format) placed this to-do list at the beginning of the minutes. Placing the to-do list at the beginning of the minutes was convenient in minutes that run 10 to 300 pages. Formal minutes could run for 10 pages, and narrative minutes, with a lengthy appendix, could reach 300 pages.

In addition, these two formats (formal minutes' Format-2 and narrative minutes' format) included information on IAB news and WG status sections. This change in the IESG minutes also represented a change in the IESG meeting format. The format of the narrative minutes differed from Format-2 of the formal minutes by including an appendix and excluding a portion of the administrivia section dealing with decisions reached between meetings ("protocol actions" and "document actions"). Levels 4 to 5 summarized exterior data from the IETF website and IESG decision repositories, such as IETF (2016b) Datatracker or earlier data files.

These levels of themes in Figure 9 indicate a common document approval process recorded by all three formats of the IESG minutes. The differences between Format-1 and Format-2 of the formal minutes were due to a growing formalization of the document approval process. Format-2 of the formal and the narrative minutes shared similar processes but titled the section headings differently.

Table 31: Comparison of Themes Found in IESG Minutes

Level of comparison – types of items	Themes common to all three formats	Themes only common in formal Format-1 and Format-2	Themes only common to Format-1 and narrative minutes	Themes only common to Format-2 in formal and narrative minutes
Level 2 – general & general narrative	date of minutes, scribe, participants, corrections	Funding	(none)	(none)
Level 2: formal Format-1, formal Format-2, narrative	administrivia, document actions, WG actions, mgt items	(none)	(none)	Protocol actions IAB news WG news
Level 4 administrivia			Bash agenda	Minutes approval, action items
Level 4 or 5 protocol actions	*1	*1	*1	WG submission, individual submission, status change
Level 4 or 5 document actions	(none)	None)		WG submission, Individual submission, status change

*1 – protocol actions operate as a theme under document actions in Format-1 of the formal minutes

After examining the format of the minutes found in the 10% of the formal and narrative minutes, the researcher concluded that these minutes documented decisions in a virtual environment where group interactions and decisions occurred at meetings and between meetings utilizing multiple streams of communication. The consistent differences between the two formats of the formal minutes (formal Format-1 [1991 to 2003], formal Format-2 [2003 to 2016]) represented a change in the IESG meeting format and a growing formalization of the document approval process rather than a change to virtual environment decision-making environment. For example, when the formal minutes expanded to add two new sections (IAB news and the WG news) in Format-2 of formal minutes, the narrative minutes added these two sections. The IESG appeared to be adding more time for communicating issues that may impact decisions as the IETF grows. Therefore, the researcher concluded that the virtual environment of the IESG consensus decision-making process had existed since 1991.

The two significant differences between the Format-2 of the formal minutes and the format of the narrative minutes pointed to two different functions in the minutes. First, the formal minutes recorded the IESG actions and the results of the IESG actions. The IESG members reviewed and agreed to this record. The formal minutes, including the documents published (protocol and document) between meetings as formal records data found in the IETF data repositories (IETF Datatracker, IETF website), fit this function. The function of the narrative minutes was to record the details of the conversations that occur within the virtual environment, which includes the meeting and the IETF data repositories, such as IETF (2016b) Datatracker, IETF website, IESG wiki, IANA website, IANA trouble ticket, and IETF secretariat trouble ticket. The recording of the discussions within minutes and the snapshot of the IETF Datatracker discussion captured in the appendix fit this function. These differences also pointed to a virtual environment of continuous decision-making.

Strand-1 Investigation Step 2: Internet Engineering Steering Group's decisions. The second step during the investigation of the impact of the virtual nature of the topology was to examine whether data regarding consensus decision making regarding documents, WGs, and IETF management actions could be found online from 1991 to 2016. This section reviews the qualitative findings of the investigation into finding in the IETF's online records details on individual behaviors and group behaviors. The first type of group behavior entailed deciding to continue discussing a consensus decision or finalize the decision. The second type of data is how individuals interact in making those decisions. The investigation examined online historical records on documents reviewed for publication, WGs chartered to develop new documents, and IETF management actions.

The research found that the historical data contained dates for the approval of documents, WGs, and some IETF (RFC Editor, 2020) actions. Documents could be approved for initial publication or changed from proposed to mature standards (draft and Internet standards). The online data confirmed the approval of documents and changes in the status of documents from 1991 to 2016. For example, the researcher found that the formal minutes on July 30, 1991, record approval of documents, including protocol documents and a document listing the standards approved by the IAB (RFC 1250). The IETF (2016b) Datatracker and the website interfaces (https://datatracker.ietf.org/doc/rfc1250/) record the publication of RFC1250 in August 1991. One of the protocol actions the IESG approved on July 30, 1991, was the movement of the Concise MIB (RFC1212) to draft standards (see Lines 159 to 161 of the minutes; RFC Editor, 2020). The text of RFC1250, the IAB standards list, also confirms this change in Section 6.1.2 of RFC1250 (Postel, 1991, p. 16). During 1991 to 2008, a series of RFC Editor (2020) documents published by the IAB provided the official status of IETF documents (RFC1250, RFC1280, RFC1360, RFC1410, RFC1500, RFC1540, RFC1600, RFC1610, RFC1720, RFC1780, RFC1800, RFC 880, RFC 1920, RFC 2000, RFC 2200, RFC2300, RFC2400, RFC2500, RFC 2600, RFC2700, RFC2800, RFC2900, RFC 3000, RFC3300, RFC3600, RFC3700, and RFC5000). After 2008, the history information in each document was considered the official information (see RFC 7100; RFC Editor, 2020). Since 2008, the official status of IETF documents and publication dates of RFCs has been listed in two online documents: the IETF (2020m) RFC index and the RFC Editor's (2020) *Index*.

The data on WG approvals on the IETF site were under two sources: web pages on each WG and the IETF conference proceedings. The IETF (2016b) Datatracker on the IETF website has a unique page for each WG and lists WGs by area. The web pages listing WGs by area list the concluded WGs (IETF, 2020c) separately from the active WGs (IETF, 2020a). For example, the IP over Asynchronous Transfer Mode (IPatm) WG was discussed in the October 10, 1991 minutes of the IESG (1992) but not approved. However, the online history of the IPatm WG (IETF, 2020g) showed the group as approved by 1/2/1992 and closed on 5/21/1996. The January 2, 1992 minutes of the IESG did not indicate the approval of this WG; thus, the researcher suspected that the IPatm WG was approved in an IESG meeting before January 2nd. Unfortunately, in 1991, several IESG (1992) minutes were not stored as historical records. In such a case, it was helpful to verify during which period the work transitions from a BOF to an official WG using the

IETF conference proceedings. The proceedings of IETF 22 held in November of 1991 (IETF, 1991, pp. 108-110) indicate an IPatm was a BOF, and the proceedings of the IETF 23 held in March 1992 conference (IETF, 1992, p. 101) indicate IPatm was an approved working group with the initial action items occurring in January of 1992. These two pieces of online data validated the IESG actions in the minutes. The researcher found that with careful examination of IETF online data, she could validate the approval of IETF (1991, 1992) WGs from 1991 to 2016.

The research found validation for some of the IESG's (2000) decisions on IETF management actions recorded in the minutes in data on the IETF website or websites associated with the IETF. The IESG web pages on the IETF website contained current and historical information about the IESG (members, areas, and terms), IESG process experiments, IESG statements, appeals on IESG decisions, and IESG sponsored groups (directorates and review teams). The IESG Web site linked to an IESG Wiki with additional information. For example, the IESG (2000) statement on "IESG Guidance on Moderation of IETF Working Group Mailing list" was approved by the IESG during the 7/13/2000 meeting of the IESG and recorded in the IESG formal minutes. The date on the web page confirmed that action within a month of the IESG meeting. The IETF (2020b) BOF Wiki contained IESG decisions on new WGs from 2006 to 2020. Associated websites might confirm the IETF management decisions to allow allocation of protocol numbers (see the IANA website's page on protocol numbers [IANA Protocol Registries]), or reports from the IAB (www.iab.org), or the IESG's interactions with other groups, such as the Internet Society (2020), IEEE, or ITU.

Strand-1 Investigation Step 3: Online data on Internet Engineering Steering Group behaviors. The third step of the investigation into Strand-1 methodology sought to determine if the online data could provide additional data on individual and group behaviors regarding documents, WGs, and IETF management items during the consensus decision-making process. The formal minutes 1991 to 2016 contains data on the individual behaviors during decisions encoded in specific terms. For example, the term "discuss" in the IESG minutes indicates a TC between two IESG members during a consensus decision regarding documents, a WG action, or an IETF management action. Support for an IESG decision might also include suggestions for improvements or comments on problems. For example, on December 12, 1992, the formal IESG minutes record (IESG, 1992) that the approval of a document specifying protocol required the majority of IESG members (nine members) agreeing in a “ballot” via email or the meeting that this protocol standard was appropriate. If an IESG member disagreed, the IESG would continue the discussion to resolve the concerns. This historical record indicates that since December 7, 1992, these ballots have occurred in a written and verbal form on documents, WGs, and IETF management.

The IESG formal minutes from 1991 to 2016 record the occurrence of “discussions” or a “discuss” on IESG decisions on document publication, WGs establishment, or IETF management. The IESG minutes from 1991 to 2003 record any historical data regarding individual discussions in the formal minutes. Starting with IETF data recorded in 1997, the online records differentiate between the date of the IESG’s approval of the document’s publication and the date the document is published. For example, online IETF Datatracker data on RFC2201 (see Ballardie, 1997) indicated this document was approved by the IESG on 6/5/1997 and published as an RFC on 9/1/1997 (IETF, 2020o). The format of the IESG formal minutes transitioned from formal Format-1 to formal Format-2 on 5/29/2003. The formal Format-1 for the IESG minutes shows that the IESG intermixed document, WG, and IETF management actions during a meeting. Formal Format-2 shows the IESG switched to a fixed agenda that progresses from documents actions (protocol, non-protocol documents, and ISE review) to WG actions (new and rechartered WG) to IETF management actions. The IETF added enhancements in 2003 to record the IESG comments on documents. For example, the IESG minutes record ballot comments on RFC3620 for IESG meeting on 9/18/2003 (IETF, 2020n). Between this time in the 2003 cohort (9/2003) and the last meeting of the 2016 cohort (3/16/2016), the formal minutes only provide references to the written ballot comments on documents contained in the IETF (2016b) Datatracker.

This change to augment the IETF (2016b) Datatracker database aligns with a change in the format of the IESG (2020) formal minutes in 2003 to 2005. The

IESG formal minutes changed to report the results of an IESG decision as blocked ("Discusses") or approved rather than including discussions and results. The IESG appeared to have missed the comments regarding discussions because in September 2005, the IESG sponsored creating the IESG narrative minutes. The narrative minutes focused on capturing the comments made during the meeting on documents, WGs, and IETF management. Starting on October 8, 2009, the narrative minutes also include a snapshot of the IESG written comments on a document (denoted as ballot comments). If the IESG reached a consensus decision on a document between meetings, the narrative minutes did not include the individual comments from IESG members. For the 2005 to 2016 cohort, the IETF narrative minutes were the only source for the complete in-meeting discussion regarding documents, WGs, and IETF management items.

The IESG (2020) appeared to have desired an online system to record comments regarding WGs. Therefore, in December 2012, the IETF (2016a) enhanced the IETF (2016b) Datatracker to include IESG ballot comments on WGs. For example, the IETF's (2021b) online web page on sunset4 WG had a historical record that includes IESG (2020) comments (denoted as ballot comments) during the consensus decision to approve the sunset4 WG. However, the IESG formal and narrative minutes from December 2012 to March 2017 did not include these ballot comments. These ballot comments provided information on individual and group behaviors regarding these decisions. During this period, the IESG formal minutes only gave the results (approved or conflicts), and the IESG narrative minutes only provided in-meeting comments. Therefore, the only way to capture the behaviors of the IESG regarding new work during 2012 to 2016 was to include data from the IETF (2016a) Datatracker.

This qualitative investigation into online data demonstrated to the researcher that the IESG used a continuous consensus decision-making process in a virtual environment that allowed both online and in-meeting comments. In some cases, the online comments set the tone for the group discussions. Without including the available online data, the historical analysis would have been flawed. The IETF continually upgraded their IETF (2016a) Datatracker database to provide online data available from 2013 to 2016. Therefore, the IPA analysis utilized these data to set discovered variables.

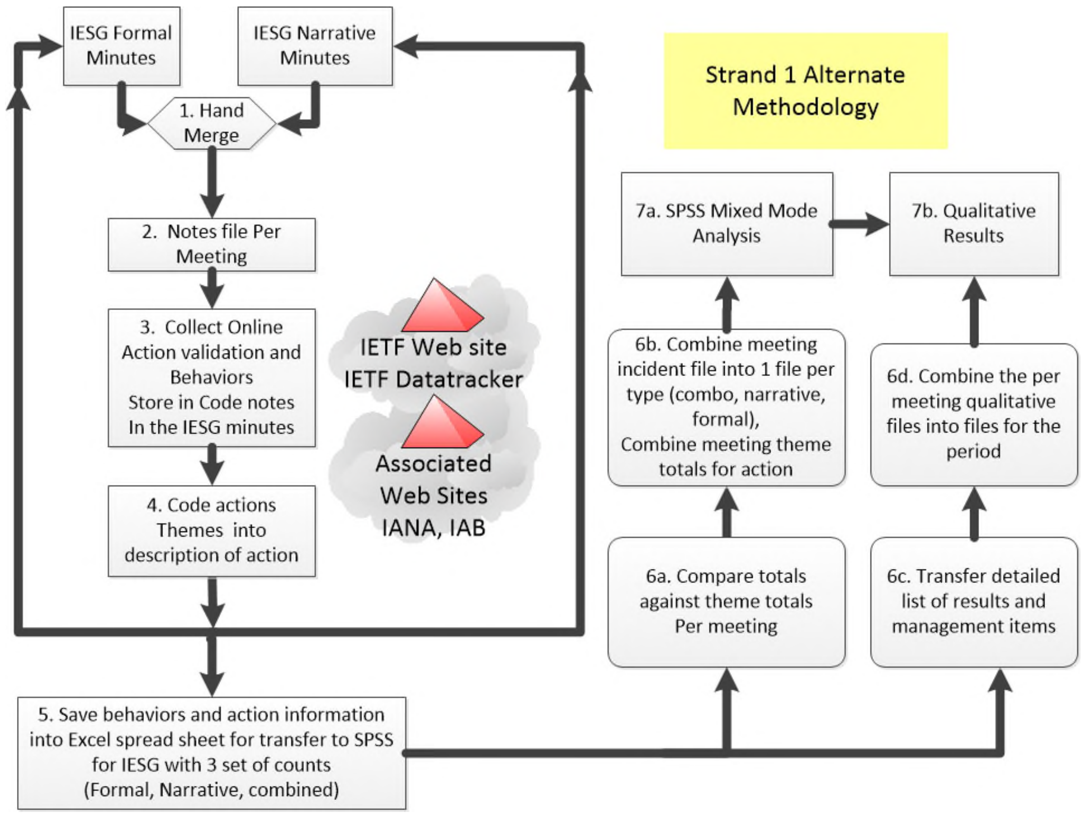
Strand-1 – Alternate methodology. The researcher revised the original methodology for collecting group behaviors. The alternate method copies appropriate information from sources of IETF (2016a) data (e.g., Datatracker, Website, and Wiki on IETF website) and IETF-related data (IAB website, IANA website, IETF secretariat web, and other websites) into code notes associated group decision incidents described in the formal or narrative minutes. Figure 11shows the alternate Strand-1 data collection and analysis flow in seven steps.

Steps 1–3 operated to collect consistent data to augment the minutes. First, in step 1, the researcher prevented inconsistent data in the two types of minutes (formal and narrative) by hand-merging the IESG decision incidents from both IESG minutes files into a sequential list of decisions per meeting. Next, the researcher read through the formal and narrative minutes seeking IESG incidents of decision-making to create a single sequential list of actions. This sequential list of IESG decisions in the notes file allowed the researcher to align the code notes in the formal and narrative minutes.

Step 2 of the Strand-1 data collection process converted this list of sequential actions into a notes file per meeting. Next, the researcher added the line numbers where the IESG formal and narrative minutes recorded each decision to the notes file. The researcher also augmented each decision's entry with the type of decision (document, WG, or IETF management), subtype of decision (e.g., new WG or rechartered WG), source of decision (Area, IESG member sponsoring), behavioral summaries from minutes (e.g., number of conflicts on a document from which IESG member), and the status of the decision (results or consensus discussion continuing).

Based on this information, in Step 3, the researcher collected data from online sources to validate or augment the data and copies the data into the code notes. After copying the data into the code notes, the researcher wrote a summary to direct Step 4's encoding of themes. Step 3 required the researcher make an entire pass through the minutes to validate, augment, and summarize collected data.

Figure 11: Strand-1 alternate methodology.



Step 4 through Step 6 encoded themes, summarized theme counts, and passed these theme counts to the quantitative and qualitative analysis in Step 7. Step 4 encoded the themes for individual behaviors and group actions in each IESG minute file (formal and narrative) per meeting at the appropriate point. The appropriate point depended on the type of decision-making (document, WG, or IETF management) and the data already contained in the minutes for the meeting.

A following subsection describes the methodology for themes encoding for each type of decision-making action. Step 5 stored the individual themes and actions into an Excel spreadsheet for transfer to SPSS. The researcher stored the theme counts in three different Excel spreadsheets for transfer to the SPSS based on three views of the data: formal minutes view, narrative minutes view, and combined minutes view. The formal and narrative views of the data were part of the original Strand-1 methodology. The combination view was the researcher’s merged view of the data. Although MAXQDA (2016) software could transfer theme counts per file, it could not automatically perform the necessary actions to create group behavioral values as summaries of individual behavior actions.

Step 5 contained a great deal of hand processing of theme counts, so the researcher added self-checking features to the Excel spreadsheets for theme counts for group behaviors, group actions, and group decisions. In Step 6, the researcher validated theme count data before sending the data to processing in Step 7. In Step 6a, the researcher used the spreadsheet validation for totals theme counts per meeting to validate incident files before transferring the incidents to SPSS for processing. Step 6b combined all the validated incident files into one file per type (formal, narrative, or combinational). For example, the researcher combined all the incident files for the combinational minutes' view for all IESG meetings for the period. In Step 6c and Step 6d, the researcher combined each meeting's qualitative results and processed these data using Strand-1's qualitative tools (theme grids and weighted node diagrams). Each meeting generated the notes file giving the combined viewpoint on the meeting, a list of results, a summary document on the documents reviewed (approved and not approved), summary documents for each WG handled, and a list of IETF management items. The researcher, in Step 6c, gathered this information and wrote observations in a code note. Step 6d contained gathered qualitative documents for an entire period (e.g., 2015) for qualitative analysis. The researcher used the qualitative tools in MAXQDA (2016) to help create theme counts, theme grids, and weighted node diagrams for steps 6c and 6d.

The researcher, in Step 7, did quantitative and qualitative processing of the data referred to by IESG minutes for a single IESG cohort. First, Step 7a did the mixed-mode quantitative analysis according to the revised codebook found in Appendix H (see alternate methodology). Part of that quantitative analysis confirmed that the theme count totals uploaded match the generated action totals for the data from Steps 6a to 6b. After the researcher completed the quantitative analysis, the researcher wrote a qualitative summary analysis considering if the quantitative and qualitative data supported or failed to support the theoretical hypothesis. The qualitative summary also contained any observations on the data.

Strand-1: Alternate methodology – Saving data in code notes. The online data stored in code notes varied by type. Initial experiments with storing online data showed that duplicating data already in both the formal and narrative IESG minutes did not hold substantial value. Instead, the researcher stored the online data and summaries in code notes attached to the point in the minutes mentioning the IESG act of decision-making. The researcher encoded themes in the text of IESG minutes (formal and narrative) based on the text and the code note summaries. For example, the code notes could have contained most of the data regarding group behaviors for an act of decision-making in the formal minutes, but the researcher encoded the individual and group behaviors in the IESG formal minutes. For the narrative minutes from 2009 to 2016 that contained an appendix, the data regarding the group behaviors would be in the main body of the minutes, the code notes, and the appendix. First, the researcher encoded themes found in the text (main or appendix) and then added the encodings for the additional online data stored in code notes. During 2005 to 2016, some biweekly meetings of the IESG only had the meeting recorded in the formal minutes. In these cases, the researcher created a "pseudo-minutes" file for the narrative minutes from the formal minutes and the online data typically placed in the narrative minutes.

Storing online data regarding WG decisions required a different approach than the documents. For example, starting in 2013, the IETF (2016b) Datatracker contained lengthy "ballot" comments on charters for new WGs or existing WGs rechartering, but the formal and narrative minutes did not provide a copy of this information. First, the researcher stored these lengthy comments on a WG chartering or rechartering decision (denoted as WG ballot) per decision-making action found in the IETF (2016b) Datatracker in a file. The single copy of the online data also lessened the risk of errors due to inconsistent copies.

Second, the researcher encoded the individual behaviors within the group behavior for the chartering of new groups or the rechartering of existing groups for the formal and narrative minutes. The encoded individual themes were associated with either narrative or formal minutes but not both. Thus, the group theme counts per IESG WG decisions were appropriately detected per type of IESG minutes.

Finally, the researcher created the qualitative data file per WG by combining all the WG charter/recharter decision files into one file and adding information regarding the standards documents considered or approved by the WG. WG qualitative analysis in an area benefits from these histories of individual WGs. Some key WGs in each area existed for over 20 years. For example, the mmusic WG in the ART area (IETF, 2020k) had existed since 1993, and the idr WG in the routing area (IETF, 2020e) had existed since 1994. The attitudes of the IESG (2020) members regarding these long-lived WGs varied from strong approval to active disapproval. Chairs often led these two long-lived WGs for decades.

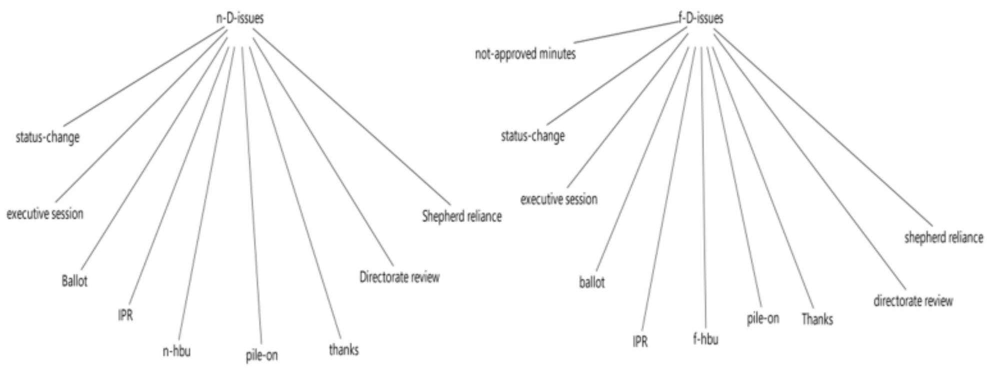
The formal and narrative IESG (2016c) minutes contained information on IETF management decisions, but the information recorded per type of IESG minutes was not always the same. Therefore, the researcher, during her hand-merge process, determined if the information in one type of minutes (e.g., formal) is referenced (however obliquely) in the other type of minutes (e.g., narrative). If the researcher found data referenced by one type of IESG minutes with full details, the researcher added a code note with the extra data. The code note included the other type of minutes as online data. For example, the researcher found that some IETF management decisions had group behaviors recorded in only one type of IESG minutes. The IETF management items had one other problem, the IETF management actions since the text did not always denote repeating.

After reviewing the entire data for 2016, the researcher was suspicious that the IETF management decision actions were repeating. After encoding 2015, it was clear that several IETF management events that were repeating in 2015 and 2016 were simply long-lived consensus decision-making actions. Therefore, the researcher collected a list of IETF management items per year to track these long-lived IETF management consensus decisions. The IPA results for 2015 to 2016 had the yearly list of management items (see Appendix O and P).

Strand-1 – Alternate methodology – Discovery of new themes. The researcher tried to discover new themes in the IPA analysis of 5% and 10% of the minutes from 1991 to 2016. However, the varying formats of the formal and narrative minutes stymied efforts to find consistent data on new themes. The inconsistent data formats made it difficult to determine whether a particular piece of information was a new theme or new data about an old theme. The alternate methodology made detecting variables easier, but the content of the 10% data (2 meetings per year) made themes difficult due to the gaps in the discussion. During the interrater review of the 10% strand data using the alternate methodology, the researcher and the two inter-raters reviewed the new process and the issues around discovering themes. The inter-raters suggested that looking at 4 months of 2016 for some of the disjointed variables, such as “thankfulness for aid” or problems mentioned, might be a good place to solidify discovered themes. This last step in discovering new themes for the IPA encoding codebook occurred during the IPA analysis of four months in 2016.

The IPA analysis of sequential IESG minutes from 4/21/2016 to 7/7/2016 enabled the researcher to read the IESG minutes as conversations that lasted across several meetings of the IESG. This context allowed the researcher to disambiguate new data from old themes and discover themes relating to thankfulness and an IESG member flagging issues. These themes discovered relating to thankfulness include: expressions of thanks, thanks for aid from directorate reviewers, and thanks for shepherds of documents (denoted as shepherd reliance). The discovered themes for IESG members flagging issues included document review themes and organizational issues. The researcher encoded flag-issue themes on documents when multiple IESG members had similar comments (denoted as “pile-on” comments) or when multiple members considered a document a harmless but useless standard. The IESG flagged IPR issues in all types of decisions (documents, WG, and IETF management). Other organizational issues included IESG decisions (denoted as ballots) that indicated strong disagreements, executive sessions for confidential discussions of “sensitive issues,” status changes for documents (e.g., from proposed to full standard), and minutes that were not approved. Figure 10 shows the weighted node diagram for the initial examination of discovered themes in each type of IESG Minutes. The final grouping reorganized these discovered themes to the following master themes: ThankAid, FlagIssue, and status-change.

Figure 10: Theme node diagrams for discovered issues (narrative and formal minutes).



<u>*Strand-2 Methodological Changes*</u>

The researcher also became concerned that the IETF minutes and the resulting online statistics at www.ietf.org did not align with the researcher’s assumptions for Strand-2 data because the IESG had a continuous virtual consensus decision-making environment. The research methodology assumed that the IESG decisions would statistically align with some constant variance with the IETF statistics collected in Strand-2 on documents published, WG actions, and management actions per year. The researcher realized this assumption could be wrong if the IESG operated in a virtual environment with continuous consensus decision-making. The investigation into whether Strand-2’s assumption was wrong looks at the statistics for documents and WGs and compared these statistics against the theme counts for decisions per year with results found in the 5% and 10% of the IESG formal minutes.

The investigation had three steps: (1) compare document theme counts with online document statistics, (2) compare theme counts for working actions against online records for BOFs and WGs, and (3) determine if adding “pseudo” BOF call minutes created valid Strand-2 statistics. The following subsections discuss these three steps in the investigation. Table 32 summarizes the research in these three steps in terms of comparisons made, Strand-2 methods changes to address problems, and how the researcher tested the changes to the methods.

The researcher concluded after this investigation that no online source was valid enough for this research; instead, the total of IESG validated decision actions with results in the IESG minutes is the best basis for Strand-2 statistics per year. The researcher would need to validate each action as unique and corroborated by online data. The researcher also concluded that IESG minutes did not include enough BOF WG actions because formal IESG minutes did not include the three BOF approval meetings per year until the 2016 IESG cohort. Therefore, the researcher created the three “pseudo” BOF call minutes per year and an IPA of the summer BOF call per year to create a complete 10% sample of the IESG actions per year. The researcher analyzed this 10% sample to determine if it could provide accurate Strand-2 statistics for the survey data and concluded it could provide a valid estimate. However, 100% of the IESG minutes were reviewed for 2015 and 2016 to confirm the IPA analysis. This overview section summarizes three steps of investigating the methodology before going through the detailed statistics and the changes to the methodology for Strand-2.

Step 1 of the Strand-2 investigations into methodology compared the IESG approved documents in 5% and 10% of the IESG minutes analyzed with the documents published by the RFC Editor (2016a, 2016b) per year. The investigation used validated counts for approved documents in the IESG minutes. For 5% of the IESG minutes from 1991 to 2016, there was no correlation between documents approved per year and the documents published by the RFC Editor. However, with 10% of the IESG minutes analyzed from 1991 to 2016, there is a positive correlation (r = 0.436, n = 26, and $\rho < 0.05$) between the IESG approving a document and the publication. One might wonder if this correlation represented a simple delay. The qualitative analysis of the graph of the approved IESG minutes in 10% of the data versus the RFC publication rate (shown in Figure 11) did demonstrate a simple relationship of a constant delay between IESG approval and document publication. A subsection below provides the details of this examination into the correlation between IESG approval and RFC Editor publication statistics.

Steps 2 and 3 of this Strand-2 investigation compare the validated WG actions found in the IESG minutes versus the statistics on WGs (concluded and active) for 1991 to 2016. Step 2 compares the 10% of the total working approved versus the 131 validated IESG WG actions in 10% of the IESG (1992, 2020) minutes from 1991 to 2016. WG actions can be related to new work creation, rechartering WG, and managing WG. The online IETF (2016b) statistics at datatracker.ietf.org recorded 768 WGs and 281 BOFs that did not create workings since the IETF began in 1986. These data implied that the IETF held ~1,049 new work BOFs from 1986 to 2020. Approximately 80% of these WG actions occurred from 1991 to 2016 (614 WGs, 839 BOFs). The results of Step 2 indicated that the 131 validated data actions from the 10% of the IESG minutes aligned with the online totals for WGs formed, managed, rechartered, and closed but did detect the correct number of WG actions for BOFs. Step 3 of this investigation determined that the BOF call per IETF contained the extra IESG actions in one meeting. Since these BOF call meetings were closed to the public from 1991 to 2/15/2017, the BOF calls minutes were not made public. However, the researcher determined that the actions in the BOF call meetings were available from other online sources. For example, the IETF plenary proceedings contain a list of BOF scheduled for each meeting, and the IETF BOF wiki (IETF, 2020b) contains data on proposed and accepted new work since IETF 65 (see Arkko, 2016b). By analyzing the content of one BOF call per year from 1991 to 2016, the researcher estimated 10% of the BOF actions per IESG cohort year. Adding this count to the working actions aligned the 10% IESG WGs and BOFs to the Strand-2 counts for WG actions in 10% IESG minutes and BOF calls.

Based on these three steps of additional investigation, the researcher decided to revise the Strand-2 methods to use the validated count of IESG actions in IESG minutes and to include the IESG BOF call pseudo minutes in the yearly IESG minutes. The Strand-2 methods will store the data supporting the validation of each action in code notes, use the action themes to encode the status of actions (result or no-result) within the notes, and generate a list of results. The themes in the IESG theme action group table were split by (a) type (document, WG, or IETF management), (b) subtype of action, and (c) status (result or non-resulting discussion). Appendix O provides the complete list of action themes and theme diagrams. The researcher used the revised Strand-2 methodology with 10% of the IESG minutes before using it on a 100% sample for four months in 2016 (4/21 to 7/7/2016).

The IPA analysis of the 4 months in 2016 showed a continuous process of IESG approvals which were not limited to IESG meetings, plus two additional types of variable delays before publication of documents. The first type of variable delay occurred when an IESG member or authors did post-approval editing or checks before sending the document to the RFC Editor's (2016a) queue. The second type of variable delay occurred when the RFC Editors (2016b) performed editorial changes and reference checks for documents referred to by an RFC. The time for these changes varied per document or document group. Given these results from the 4-month sample of 2016, the researcher concluded that the altered methodology of Strand-1 and Strand-2 was critical to valid results for Strands-1-5. The researcher's changes to Strand-2 methods required adjustments to the methodology in Strands-3-5.

Table 32: Strand-2 Investigations

Research to refine the problem	Result changes	Test of changes	Risks of error in results
Initial Research: Gather statistics on documents, WGs, and IETF management action s	Initial data caused suspicion on theoretical assumption two, so the researcher added the investigation in steps 1 and 2		Online data errors
Step 1: Correlate the number of documents the IESG approved, and documents published by RFC Editor (2016a, 2016b) per year with (a) 5% IESG minutes and (b) 10% of IESG minutes with validated IESG actions.	1) Created revised methodology for Strand-2, which stores the validation for IESG decisions in code notes and encodes each action theme tag with either a result or non-result flag. 3) Use action theme counts with result flag as statistics per yearly cohort.	Tests with data from 5% and 10% of IESG minutes	Online data errors
Step 2: Compare the total number of approved WG actions in 10% of minutes with online records of WGs and BOFs	1) Add creation of pseudo-BOF call minutes for IESG cohorts from 1991 to 2016 for BOF calls without IESG minutes, and include as an IESG minute	Run revised methodology for 10% of data	Online data errors
Step 3: Compare the total number of WG actions in 10% of minutes + 10% of BOF call actions with online records of WG and BOFs	1) Generate a list of results to compare qualitatively against workings meeting	Run revised methodology with four months in 2016	

Strand-2 Investigation Step 1 - Documents. The first step in the Strand-2 investigation into using online statistics as IESG results involved testing that assumption with 5% of the IESG actions found in the minutes for 1991 to 2016. Table 33 shows the statistics generated from 5% IPA (May minutes) and 10% IPA (May and December minutes) versus online statistics. Column A shows the number of documents (RFCs) approved in a May meeting, and column D shows the number of RFCs published per year. The researcher tested the relationship between the documents approved in 1 meeting per year (maydocs) and the RFC Editor (2016a) record of documents published per year (rfcyr) using the Person product-moment correlation coefficient. The researcher performed preliminary analyses to ensure no violation of normality, linearity, and homoscedasticity assumptions. There was no correlation between these two variables.

Table 33: Documents Approved by Internet Engineering Steering Group Versus Request for Comments Published

Year	A. May Minutes Approved (maydocs)	B. December Meetings approved (decdocs)	C. Meetings per Year (mtgs)	D. RFC Editor RFCs (rfcyr)	E. Documents in two meetings
1991	4	6	14	95	10
1992	7	3	38	95	10
1993	11	1	23	175	12
1994	3	5	24	185	8
1995	2	18	16	13	20
1996	29	6	23	170	35
1997	6	7	19	192	13
1998	3	6	21	234	9
1999	18	8	22	259	26
2000	4	5	21	279	9
2001	2	13	23	194	15
2002	4	13	21	220	17
2003	10	8	20	235	18
2004	18	10	23	281	28
2005	14	16	24	327	30
2006	8	8	24	459	16
2007	5	14	22	320	19
2008	14	8	25	290	22
2009	16	21	23	286	37
2010	10	16	24	363	26
2011	17	13	23	390	30
2012	14	13	23	338	27
2013	12	23	22	275	35
2014	11	9	24	327	20
2015	13	16	25	300	29
2016	12	13	24	310	25

The following three influences could have caused the lack of correlation between IESG document approval in 5% of the data: (a) errors based on the 5% sampling size, (b) delays after IESG approval of documents prior to entering the RFC Editor’s (2016b) queue due to virtual nature of IESG consensus decision-making process, and (c) variable times for the RFC Editor’s (2016b) publication process. If errors were due to the 5% sampling size, increasing the sample size to 10% would provide a better basis for a correlation between the approved IESG (2020) documents and the documents published by the RFC Editor (2016b). If the IESG’s (2020) virtual nature caused post-IESG approval delays, only a detailed analysis of the IESG minutes plus all IETF website available data (IETF, 2016b, 2020d, 2021a) and related online data (RFC Editor website, IANA website, IAB website, and IETF secretariat). If the RFC Editor (2016b) processing caused a problem, then delays would occur during processing. This logic pointed the researcher toward two different directions: (a) a 10% IPA analysis and (b) a detailed examination of a short period. The researcher discussed the problem with other experts during the 5% codebook creation process. Without a clear winning strategy, the researcher took both paths to establish an alternate methodology.

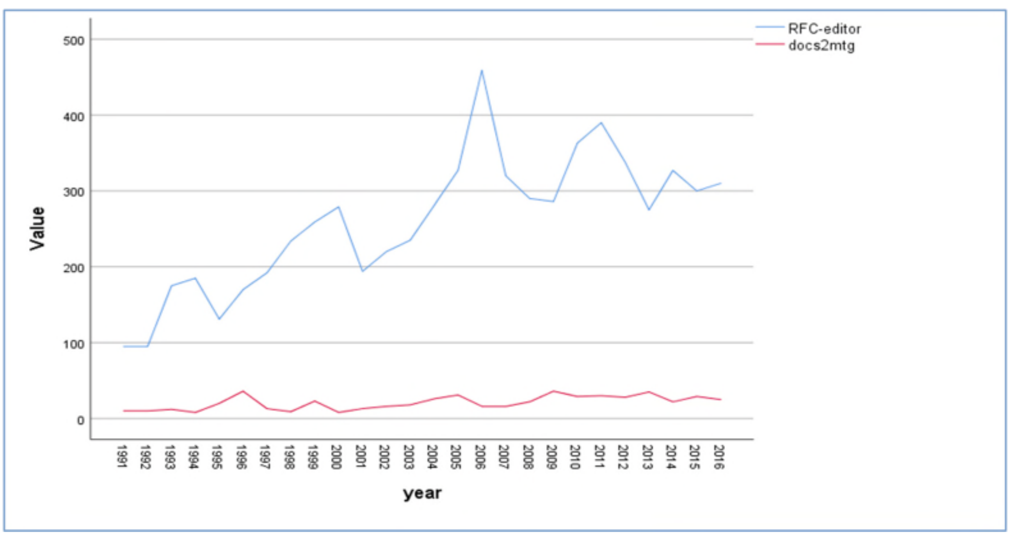
The researcher did an IPA of 10% sample for 1991 to 2016 cohorts and a 100% detailed sample of seven IESG meetings from 4/21/2016 to 7/7/2016. The 10% sample includes 52 formal minutes, and the 100% sample includes seven formal minutes. The 10% sample was created by including the December meeting per IESG cohort year nearest to December 15. The IESG formal minutes have one meeting in 2016 (6/2/2016) included in the 10% and 100% analysis. The total minutes examined in the Strand-2 method investigation was 58 minutes out of 598 minutes. The 10% sample of narrative minutes included 23 minutes from 2005 to 2016, and the 100% sample included 7 minutes from 2016. Like the IESG formal minutes, the narrative minutes had one meeting in 2016 (6/2/2016) included in the 10% and 100% analysis. The total narrative minutes investigated in this Strand-2 investigation was 29 minutes out of 246 narrative minutes.

The theme counts for IESG documented decisions with results from hand-merge of the IESG formal minutes and narrative minutes for the 5% and 10% sample have mixed results. The correlation test indicated no correlation between the yearly theme count for document approvals from a 5% sample using May meetings and the number of documents that RFC Editor (2016a) published per year. The researcher created the 5% December sample by selecting the December meeting closest to December 15 per year and the 10% sample by combining the two 5% samples (May and December meetings). Table 33 lists the number of documents approved in the December meeting per year in Column B and the number approved in two meetings (May and December) in Column E.

After conducting preliminary analyses to ensure no violation of normality, linearity, and homoscedasticity assumptions, the researcher ran correlation tests between (a) theme counts from the 5% sample using December meeting data and the RFC Editor (2016a) statistics on RFC published per year, and (b) theme counts the 10% sample and the RFC Editor's statistics on RFC published per year. The correlation tests indicated a positive correlation between the documents approved in a single December meeting and the RFC Editors' statistics with $r = .391$, $n = 26$, and $\rho < 0.05$. The results also showed a positive correlation between the sum of the documents approved in the two meetings (one in May and one in December), with $r = 0.436$, $n = 26$, and $\rho < 0.05$.

Figure 11 shows a graph of the documents approved in the two meetings (docs2mtg) versus the RFC Editor's statistics. The graphs did not indicate a correlation directly related to a simple time shift between approval and publication. The differences between the IESG rate of RFC approval and the RFC publication rate could be caused by the limited amount of data (10%), or the continuous nature of the virtual environment of the IESG consensus decision-making could also cause these types of relationships.

Figure 11: Internet Engineering Steering Group Request for Comment approvals versus Request for Comment Editor publication statistics.



Strand-2 Investigations Step 2 – Working group actions. The second investigative step compared the WG actions found in the IPA of 10% of the IESG minutes with the online statistics compared for total WGs. The theme counts for IESG WG actions used in this step were encoded using the refinements to the

Strand-2 collection process defined by Step 1 of the investigation. The researcher encoded the theme counts in the IPA of two meetings per year (May and December) and generated theme count totals for WG actions using a mixed-mode analysis tool in MAXQDA. Table 34 contains the number of WG actions in each of the two meetings per year (May and December) and the total for the two meetings. In addition, the researcher gathered online statistics on WG and BOFs from the IETF web pages at datatracker.ietf.org/group. Finally, the number of IESG approved WGs and proposed work (BOF meetings) did not result in a WG.

Table 34: Approved Internet Engineering Steering Group Actions

	A. May minutes WG actions (maywg)	B. December minutes WG actions (decwg)	C. Total Actions in 10% of IESG meetings
1991	2	0	2
1992	0	0	0
1993	3	0	3
1994	1	2	3
1995	1	2	3
1996	1	2	3
1997	2	0	2
1998	1	0	1
1999	4	2	6
2000	0	2	2
2001	1	4	5
2002	1	1	2
2003	5	2	7
2004	2	1	3
2005	2	8	10
2006	4	1	5
2007	7	1	8
2008	5	1	6
2009	10	5	15
2010	5	5	10
2011	5	1	6
2012	4	4	8
2013	4	5	9
2014	4	0	4
2015	1	2	3
2016	2	2	4
Total	77	53	130

Table 35: Working Groups and Birds of Feather Meetings in the Internet Engineering Task Force (1986 to 2020)

Area	ART	GEN	INT	OPS/NM	RTG	SEC	TSV	Total
WG	73	14	81	81	60	72	65	453
BOFs	5	24	32	25	18	29	42	176
Older Areas	APP + SAP	IPng	OPS	NM	RAI	Sub-IP	USR	total
WG	208	12	13	24	25	7	14	315
BOF	75	3	6	2	7	1	11	105
Total WGs								768
Total number of BOFs that did not become a WG								281
Total WG + BOFs								1049
Estimate for 1991 to 2016 (80%) – WG								614
Estimate for 1991 to 2016 (80%) – BOF + WG								839

The IETF (2020d) process creates new WGs by holding BOFs meetings before the IESG approved a charter for the technology standard work. After completing some tasks, a WG would recharter to undertake new work or close. The IESG members managed WGs through this process. Therefore, IESG WG actions could involve approving new work (BOFs), new charters, revised charters proposed for existing WGs, closing existing WGs, or management of WGs. The researcher gathered data from the IETF online statistics and IETF web pages to determine the number of WGs per area from 1989 to 2020 and the number of BOFs per area from 1989 to 2020. Table 35 shows the totals the researcher gathered per area. The IESG created 768 WGs across all IETF areas and sponsored 281 new work proposals (BOFs) that did not result in WGs. Therefore, a conservative estimate for the WG actions is the sum of the WGs plus the BOFs that did not result in a WG. For this investigation, the researcher estimated the total WG Actions (WG + BOF) to be 1,049 actions (1981 to 2020) and estimated that 80% of the total WG actions occurred during 1991 to 2016. This estimate is a conservative estimate for BOFs because some WGs held two 2 BOFS before finalizing a charter for new work.

The researcher compared these estimates of WG actions from 1991 to 2016 with the theme count totals from the 10% sample and detected a lower than expected number of theme counts for BOFs. Table 36 shows that 10% of the total WG actions would be 84 BOFs WG actions and 61 WG established actions. The theme counts detected in the 10% IPA were 7 BOF counts and 52 new WG charter actions. The estimate of 61 WG established and the 52 new WG were close. The processing of writing standards went through cycles of defining new work and finishing the work.

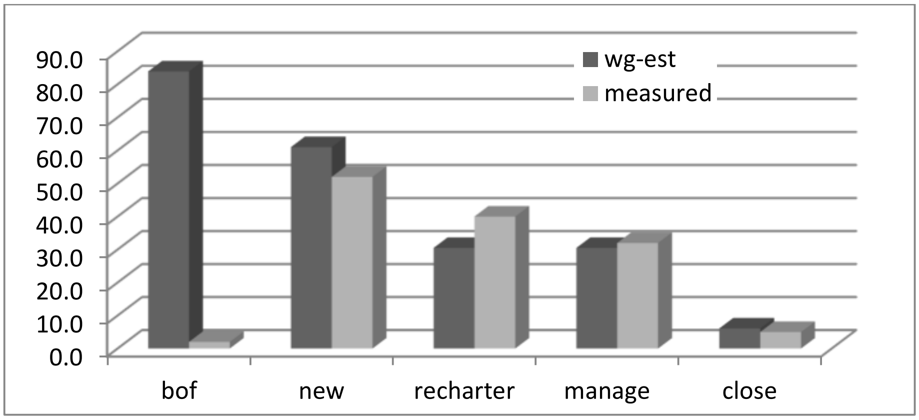
At the end of a cycle, a WG could either close or recharter. The researcher estimated the recharter, management, and close WG statistics based on the online statistics for chartering new WGs. The researcher estimated that 50% WGs recharter per year and the other 50% need management actions. About 10% of the IETF's WGs close per year, and these WG were part of the 50% of WG, which required management actions before they close.

Table 36 shows the estimates for rechartering, management, and closure of WGs from the online statistics. The estimates for WG's recharter, management, and closure are close to the measured data. The estimated actions for a 10% sample are 212, but the theme count total for WG actins from the 10% IPA is 131 WG actions. These 131 validated WG actions include 2 BOF actions (2%), 52 (40%) on new WGs actions, 40 actions (31%) on rechartering actions, 32 (24%) actions for WG management, and five (4%) actions to close WGs. The BOF proposals were out of line with the total BOF measures, as the bar graph in Figure 12 shows.

Table 36: Estimated Versus Actual Counts

Category	Percentage of total WG actions (10%)	Working group estimate for 10% of data	IESG approves measured in 10% of data	Percentage of total WG actions
BOF	10%	84	7	5%
New	10%	61	52	40%
Recharter (5%)	5% of WG	31	40	31%
Manage (5%)	5% of WG	30	26	24%
Close (1%)	1% of WG	6	5	4%
Total number		212	131	

Figure 12: Work group estimated versus work groups measured.



Strand-2 Investigations Step 3 – Adding BOF call minutes. The BOF approval actions were underreported in the IESG (1992, 2020) meeting minutes because the BOF call minutes were missing. The IESG (1991 to 2016) held a "BOF coordination call" (denoted as BOF call) before each IETF meeting, where BOF proposals were approved. These BOF meetings were considered private meetings for the IESG until IETF-98. Since February 2016, the BOF coordination called for IETF-98 to IETF-112 had been open meetings with IESG minutes. Since IETF-65, the BOF wiki had provided the work proposals for the next IETF (2020b). The past BOFs held at each meeting could be found in the IETF Proceedings for IETF-1 to IETF-107. Because the IETF meetings since 1991 were held three times per year, each BOF call represented one-third of the BOF actions for a cohort.

Data was collected on the number of BOFs held at the summer meeting of IETF (2020b) from 1991 to 2016 from the IETF (2020d) proceedings and the proposed BOFs from 2007 to 2016 from the BOF wiki. Table 37 lists the BOF approved for the summer meeting of the IETF from 1991 to 2016 and the proposed BOFs for the July meeting from 2006 to 2016. The total number of BOFs held in the summer IETF (1991 to 2016) was 284. If one estimates the total number of BOFS as three times the single BOFs, the total number of BOFs would be 852 BOFs. The estimate for BOF WG actions for IPA 10% of the IESG minutes with BOF calls is 85 BOF WG actions, and the estimate for the BOF WG actions based on the online statistics is 84. If the 85 BOF theme counts from pseudo-BOF call minutes were added to the existing counts in the IESG minutes, the estimated total number of IESG actions is 215. Therefore, if the 10% IPA analysis included 2 minutes and the pseudo-BOF call minutes, the WG actions were close to the 10% estimate from the online statistics.

Figure 13 illustrates how including the BOF call data aligns the 10% IPA analysis with the estimate of WG Actions 10% based on the total number of WGs and BOFs. The researcher concluded that including one pseudo-BOF call minutes in the 10% IPA analysis is critical to creating accurate statistics for the 10% estimate of 1991 to 2016. The researcher also concluded that the 100% IPA needed to include three pseudo-BOF or real BOF call minutes was critical to the 100% IPA analysis creating accurate WG actions.

Figure 13: Estimated work group actions versus measured actions.

Table 37: Birds of Feather Meeting Counts

Estimated BOF Call Date for the summer meeting	Number of BOFs approved	Estimated BOF Call Date	Number of BOFs approved	Number of BOF proposals
6/15/1991	5	6/01/2006	7	7
6/15/1992	21	6/15/2007	5	15
5/20/1993	12	6/15/2008	6	7
5/20/1994	16	6/2/2009	6	7
5/20/1995	10	6/18/2010	5	7
5/01/1996	15	6/23/2011	6	7
6/15/1997	13	6/15/2012	4	4
7/5/1998	23	6/21/2013	14	15
6/1/1999	13	6/1/2014	8	12
6/15/2000	14	6/1/2015	9	10
6/15/2001	11	6/1/2016	12	14
6/01/2002	10			
6/01/2003	15			
6/01/2004	11			
6/01/2005	13			
1991 to 2005 total	202	2006 to 2016	82	105
Total BOFs at Summer IETF meeting (1991 to 2016)		284	Estimate for all BOFs (1991 to 2016)	840

Table 38: Working Group Actions

Category	Percentage of total WG actions (10%)	Working Group Estimate For 10% of data	IESG approves measured in 10% of data	Percentage of total WG actions
BOF	10%	84	92*1	41%
New	10%	61	52	25%
Recharter (5%)	5% of WG	31	40	19%
Manage (5%)	5% of WG	30	32	12%
Close (1%)	1% of WG	6	5	2%
Total number		212	215	

*1 – 85 BOF call minutes WG actions plus 7 BOF actions already in IESG minutes.

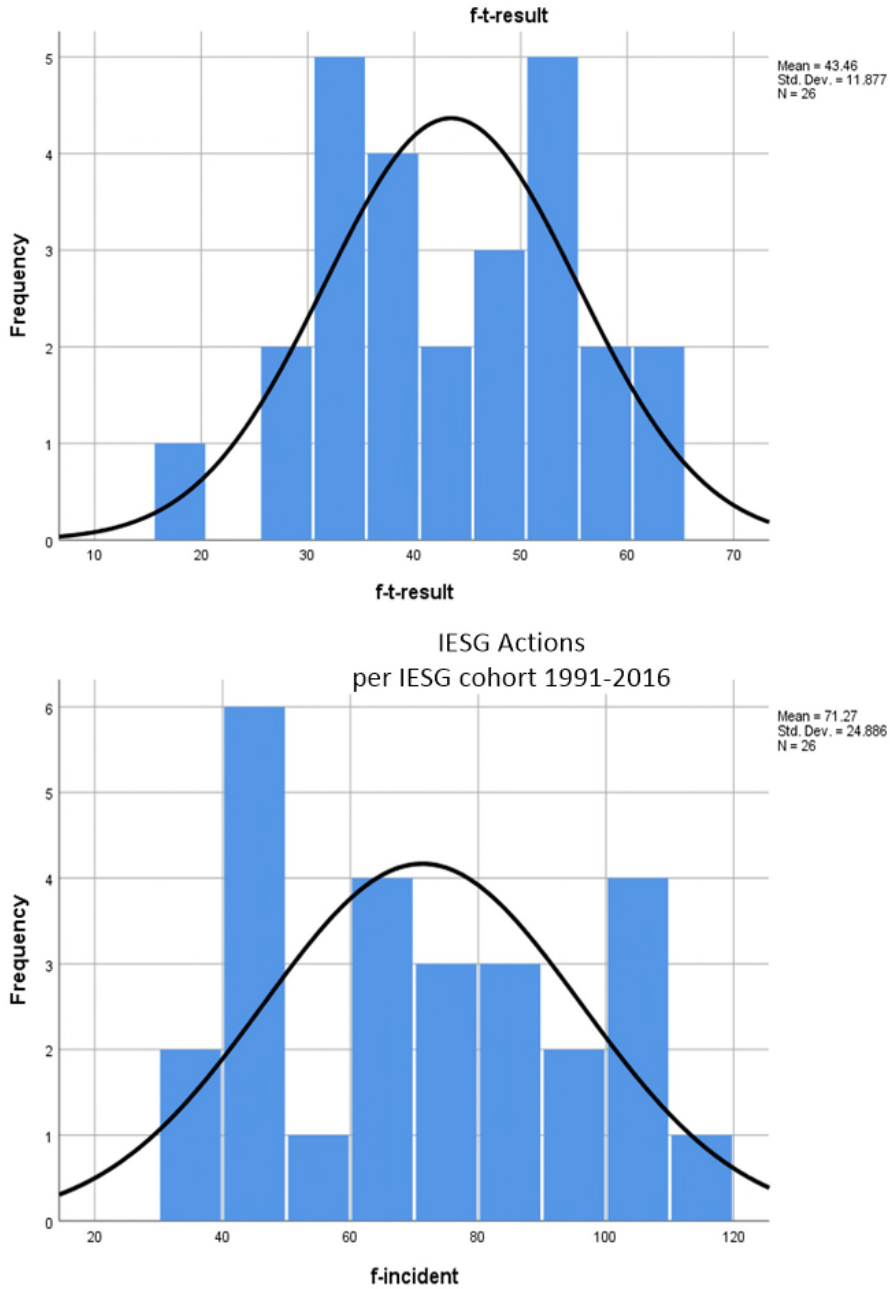
Strand-2 – Final checks of statistics. The altered methodology for Strand-2 provides validated statistics on documents and WGs that align with online statistics. Does this alternate methodology provide results per IESG cohort from the 10% sample that the Strand-3 data analysis can use to test the theoretical model? If the data sampled by the 10% sample has normality for the IESG actions with results when summarized per IESG cohort, then the tests for suitability for multivariate analysis have a criterion variable with statistical normality. The researcher ran a statistical analysis of the IESG actions with results per cohort from the 10% analysis test normality using tests for skewness, kurtosis, $z_{skewness}$, and $z_{kurtosis}$, and a histogram. Table 39 shows the results of these tests, and the values for $z_{skewness}$ and $z_{kurtosis}$ are within the 0.01 significance level (±1.96) suggested by Hair et al. (2010) for the 0.05 error level (p. 72). Figure 14 shows a histogram of the IESG actions with results per IESG cohort with a curve with normality overlaid on the histogram. The curve indicates the IESG actions with results per IESG cohort generally follows the normality curve.

The researcher did two additional checks on the validated IESG actions data before accepting the IESG results from the 10% sample. The first test was to test the IESG actions to see if the data for all IESG action per IESG cohort also had normality. The researcher ran the same normality tests (skewness, kurtosis, $z_{skewness}$, zkurtosis, and a histogram) on the 10% sample of IESG actions (with and without results)s she ran the IESG actions, with results. Table 39 shows the results of these tests. Because both $z_{skewness}$ and $z_{kurtosis}$ were within normality bounds (±1.96) for the IESG actions, the IESG action constructs within the 10% sample had normality. The second check was on the normality of the components of the IESG decision-making actions per IESG cohort (document publication, WG actions, and IETF management) for theme count data totals from a combination of the hand-merged IESG minutes and pseudo-BOF call. Each of the IESG decision-making actions components per IESG cohort had normality at the .05 error level ($\pm$1.96), as shown in Table 39. These initial statistics on the 10% sample of the IESG solidified the alternate methodology, but the researcher still tested this alternate methodology on four sequential months (seven IESG meetings) in 2016.

Table 39: Skewness and Kurtosis for Results and Actions per IESG Group Cohort

Variable from actions 10% of IESG minutes per IESG Cohort (*N = 26*)	skewness	$z_{skewness}$	kurtosis	$z_{kurtosis}$
IESG Results	-0.010	-0.020	-0.546	-0.568
IESG Actions (all)	0.086	0.179	-1.269	-1.321
Document Actions	-0.005	-0.011	-1.324	-1.378
WG actions	0.211	0.439	-0.464	-0.482
IETF Management Actions	0.146	0.304	-0.579	-0.603

Figure 14: Histogram of Internet Engineering Steering Group actions and results.



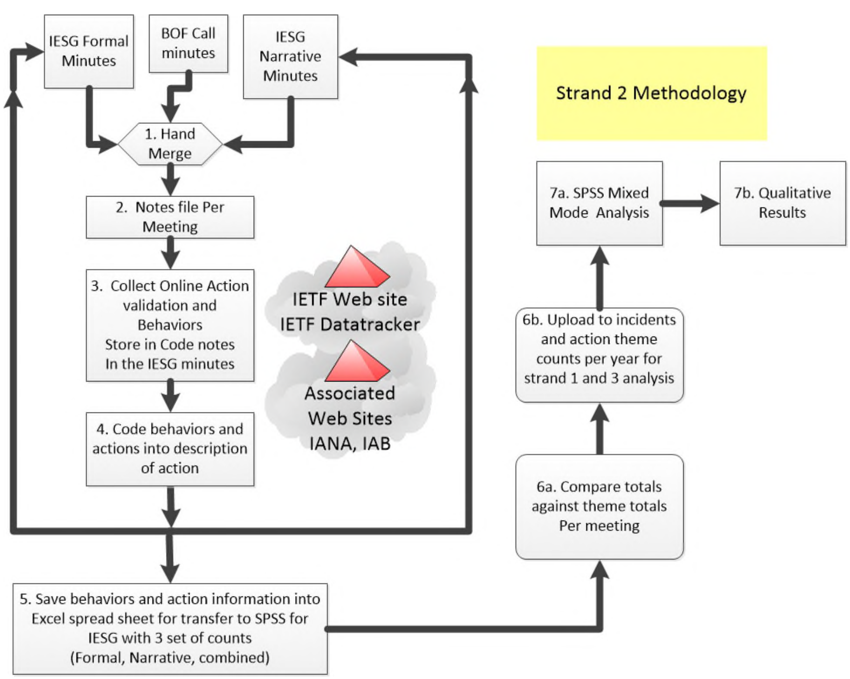
Alternate methodology for Strand-2 data collection. The alternate method for Strand-2 data collection gathers statistical information per each IESG action. The researcher collects Strand-2 statistical data for each IESG from the IESG minutes (formal and narrative) and the BOF call minutes (pseudo and real). The researcher creates the pseudo-BOF call minutes IESG cohorts from 1991 to 2016 from the online data from the IESG wiki or the IETF proceedings. However, the official BOF call minutes only existed from the 2016 IESG cohort. The researcher treated the pseudo-BOF call minutes as formal minutes for the 1991 to 2016 IESG cohorts and narrative minutes from 2005 to 2015 IESG Cohorts. The researcher treated the BOF minutes from the 2016 IESG cohort as narrative minutes. Figure 15 shows a diagram of the alternate method for Strand-2. After collecting formal, narrative, and BOF-call minutes, the researcher went through Steps 1 to 6a of the alternate method per IESG meeting until all IESG meetings, then generating the statistics for all years using Steps 6b and 7a.

In Step 1 of this new method, the researcher read through the formal and narrative minutes for an IESG meeting to create a hand-merged list of IESG actions for a meeting. In Step 2, the researcher used a single note file to align the IESG incidents of decision-making found in the formal and the narrative minutes. Many incidents of IESG decision-making actions within a meeting were the same because the IESG formal and narrative minutes described a single meeting with two exceptions. First, the formal minutes report actions that occurred between meetings. Second, the narrative minutes contained detailed descriptions of the conversations during the meeting and provide copies of online data regarding decisions on documents. Any group decision incident mentioned in formal and narrative minutes received one entry in the notes file per meeting.

In Step 3 of the processes, the researcher verified the IESG consensus decision-making actions that caused a measurable result against data on the appropriate source in the IETF website, the IETF (2016b) Datatracker, the RFC Editor (2020) website, or other public sources related to the IETF (IAB website, IANA website, and IETF secretariat). The exact method of validation varied per type of results. The researcher verified IESG (2000) actions on documents by examining the history information on each document on the IETF (2016b) Datatracker, the IETF (2020m) RFC publication list, and the RFC Editor’s (2016a) list of published RFCs. The IETF (2016b) Datatracker contained the time each document was approved by the IESG (2000) and published by the RFC Editor

(2016a). The RFC Editor (2016a) also indicated the date an RFC was published. The researcher in Strand-2 saved the time of document approval and the publication date. The combined Strand-1 and Strand-2 collection process per document in the alternate methodology used the following steps: (a) collected data on the behaviors (group and individual) from the IESG (2019) comments from the IETF website and IETF (2016b) Datatracker and hand-merge, (b) created a "notes" file per meeting, (c) collected validation references for document approval by the IESG and publication by RFC Editor (2016a) and included both pieces of information in the code notes associated with each document, (d) encoded the behaviors (Strand-1 encoding) and the document action as resulting in approval or not (Strand-2 encoding), and (e) saved the theme counts in an Excel file for transfer to SPSS and the results per meeting in a file for qualitative processing. Appendix N provides the codebook for the methodology and samples for the meeting notes file and the document code notes. The theme count totals for actions per meeting replaced the Strand-2 statistics for Strand-1 and Strand-3 quantitative analysis.

Figure 15: Strand-2 alternate methodology.



The researcher detected acts of IESG decision-making on WGs in the IESG minutes in the WG charter section, the WG recharter section, the WG news section, and the IETF management sections during Steps 1 to 2 of the Strand-2 process. Then, in Step 3 of the Strand-2 process, the researcher sought validation for these WG-related actions on the IETF web under the pages that described a WG or descriptions in the IETF proceedings that describe a WG. The IETF (2016b) Datatracker had evolved since the 1990s to contain a historical record of the IESG discussion on WG charters and the dates of the approval of charters. Before 2013, the historical data on WGs was only contained on the IETF website on pages per WG or in the IETF proceedings for IETF meetings (3 times per year). The depth of this historical data varied from indicating when the WG met in an IETF meeting to details on chartering. After 2013, the WG webpage on the IETF website linked to the online IESG consensus discussion on the charter approval process (proposed and approved). The researcher gathered this information and included it as code notes or a separate document. Appendix N contains an example of a code note regarding a WG.

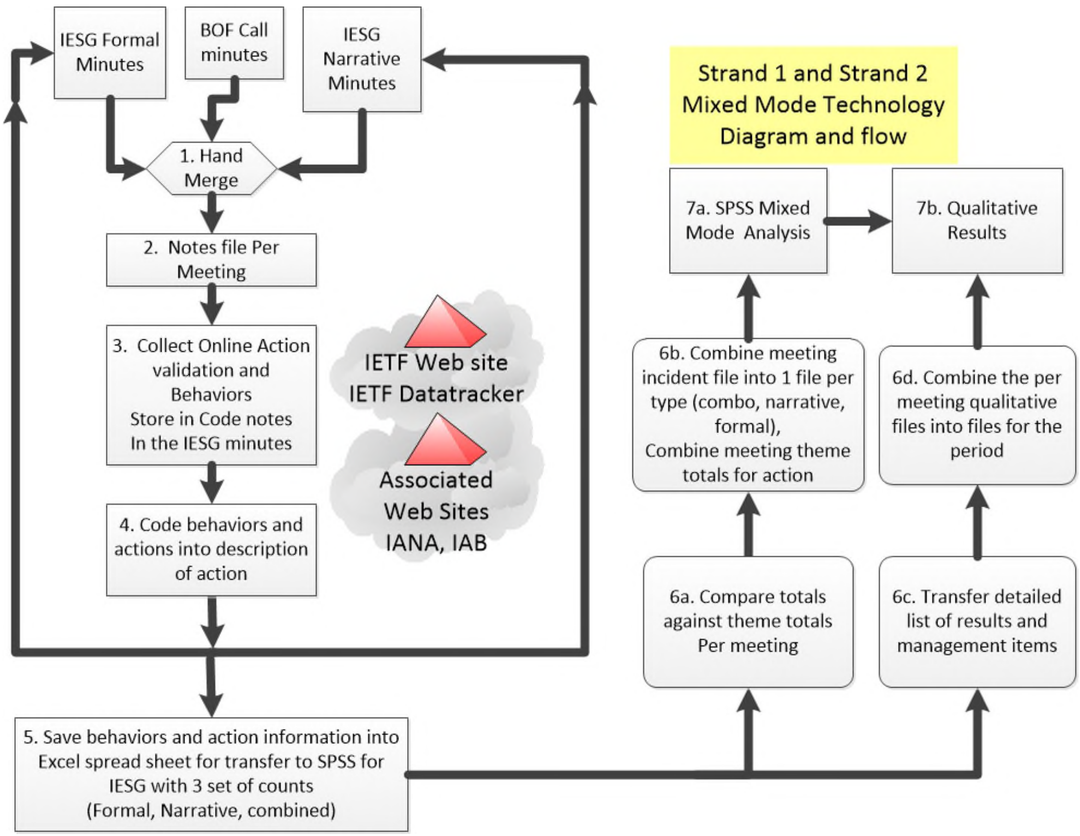
The Strand-2 data on the IETF management decisions of an IESG cohort may only exist in the IESG minutes, or additional historical records may to the decision. The Strand-2 data collection sought validation for each IETF Management consensus decision-making action that caused a measurable result. The researcher started the search for validation by examining where the original IESG minutes recorded the individual IETF management consensus decision in the minutes. For example, the researcher found IETF management decisions in IESG minutes in the sections on actions between meetings, IAB news, IETF management, and the WG news. These locations may provide a hint where the researcher could find validation on a portion of the IETF Web site (data or email lists or wikis), IETF-related websites (IAB web page, RFC Editor website, IANA website, and IETF secretariat), or website of other organizations (ISOC, IEEE, or ISO study groups). The researcher's ability to find data to validate IETF management items as unique and not a repeat was a concern, so the researcher monitored how this validation in three waves of detailed analysis on the 100% data for 2015 to 2016 (4 months of 2016, 12 months of 2016, and 12 months of 2015).

Strand-1 and Strand-2 data collection and interpretive phenomenological analysis coding method. The combined alternate methodology for Strand-1 and Strand-2 started with the formal and narrative minutes. The combined process had seven steps, as shown in Figure 16. During Step 1, the researcher read the formal and narrative minutes for one IESG meeting and sought to determine a list of IESG decisions. During Step 2, the researcher then recorded a list of the incidents of consensus decisions in the “note” file for the IESG meeting that integrated the incidents from the formal and narrative minutes. During Step 3, the researcher went through each incident of group consensus decision-making listed in the code-notes file collecting data about the individual behaviors, group behaviors, and validating group actions. Next, the researcher stored the data collected in Step 3 per incident a code note attached to the location in the IESG minute text describing this event. Next, in Step 4, the researcher evaluated the information collected and encoded the themes for individual behaviors and actions for each incident within an IESG minute (formal or narrative). Finally, for Step 5, the theme counts for individual behaviors, group behaviors, actions, and results per incident IESG actions were stored in an Excel file. This Excel file had the following worksheets: (a) per incident behaviors (individual and group behaviors) and Strand-2 data (decision-making actions and results), (b) qualitative results, (c) collaboration statistics, (d) theme count totals per meeting, and (e) validation checks. The validation checks aided the researcher's efforts to check the manual transfer of data.

After the researcher processed all the group decision-incidents in the IESG minutes for a single meeting, Step 6 prepared the data for mixed-mode analysis of the IPA theme counts (Step 7a) and the qualitative analysis (Step 7b). Step 6a involved the researcher validating that the theme count totals were the same on the spreadsheet as in the reports from MAXQDA on individual actions, group actions, and meeting totals. The researcher also validated the qualitative list of results. After the validation step, the researcher numbered the group incidents sequentially within a meeting and a year. Finally, an incident data file was prepared to pass to SPSS for three types of data: incidents in the combinational of IESG minutes data, formal minutes, and narrative minutes. The alternate methodology retained the three forms (formal, narrative, and combination) to determine how a single source (formal or narrative) might bias the outcomes from the “hand-merged” combination results. Step 6a transfers the data to SPSS. Step 7a processed the data based on the IPA codebook (Appendix B, revised codebook).

The researcher analyzed the qualitative data from the IPA analysis in Strand-1 and Strand-2 in Steps 6b and 7b. Step 6b transferred for each meeting the qualitative results. These results included a detailed list of results on each group decision results (documents, WG, and IETF management) actions, the note file for the meeting, a code note about what was significant in the documents, detailed information on WGs, a list of management items, and a summary note for the meeting. The researcher combined the IESG meetings files for processing in Step 7b, during which the researcher examined the actions of each IESG cohort in the period examined. In Step 7a, the researcher ran mixed-mode quantitative multivariate analysis tests (descriptive and multivariate statistics) on the theme counts (Strand-1 and Strand-2) to determine if the theoretical model hypotheses were supported. The results of the quantitative analysis were also qualitative Step 7b. Step 7b provided a qualitative summary of the data per period (10% 1991 to 2016, 100% 2015, 100% 2016).

Figure 16: Strand-1 and Strand-2 alternate methodology.



<u>Four Month Validation of Alternate Methods for Strand-1 and Strand-2</u>

After completing the analysis with 10% of the IESG minutes and the BOF calls, the next step in validating the alternate methodology for Strands 1 and 2 entailed a detailed examination of a group of 8 IESG meetings in 2016 for 4 months in 2016. The IESG meetings on the following dates in 2016 were analyzed using the alternate methodology for Strand-1 and Strand-2: 4/21 5/5, 5/10, 5/19, 6/2, 6/16, 6/30, and 7/7. The validation sought to (a) compare the validated statistics on documents against the online public summaries of statistics per IESG cohort, (b) compare validated actions on WGs against actual WG actions, and (c) determine if the researcher could discover using the altered methodology. The researcher compared the validated IESG document decisions against non-validated statistics at the following locations: the RFC Editor's website and the IETF website. The term non-validated did not question the accuracy of these public statistics about recording publication dates. Instead, the term non-validated indicated that these statistics did not link to approval by that year's IESG cohort.

These investigations indicated that the IESG approval did not match the non-validated summary statistics for documents. The IESG document approval indicates that an IESG's continuous virtual consensus decision-making process may reach consensus. Although document decisions clustered around IESG meetings, the IESG made document decisions between meetings using virtual consensus decision-making. After approval, the time between the IESG approval and publication varied due to the following reasons: (a) post-approval editing before entering the RFC Editor (2016a, 2016b), (b) poor quality of the text causing a lengthy editing process, or (c) dependencies on other documents (described as "MISREF" missing references).

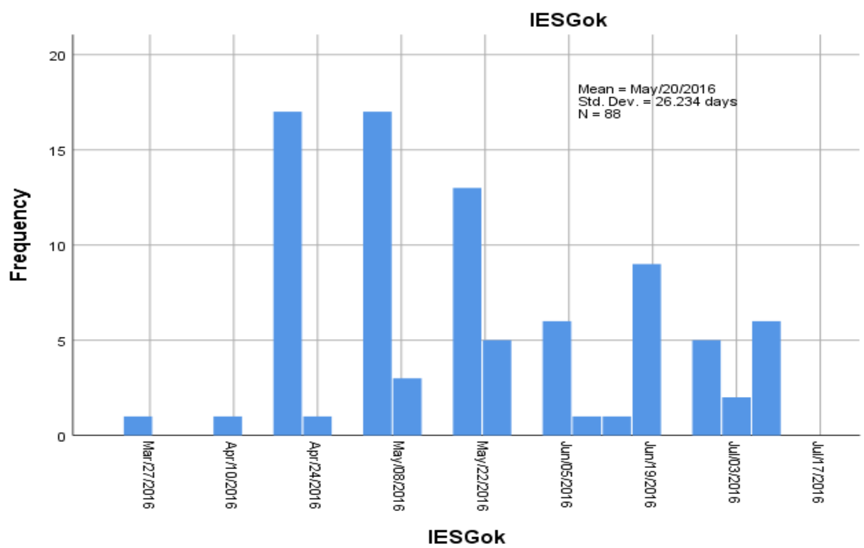
Table 40: IESG Approval

Strand	Research to refine the problem	Results and changes	Changes and Next Steps	Error risks
Strands 1 and 2	Research: Combined Strand-1 and Strand-2 data collection for eight meetings from 4/21/2016 to 7/7/2016 Investigations: 1) validated document actions compared with public online statistics 2) validated WG actions compared with public online statistics except for BOF calls 3) variables could be discovered	The analysis showed that the IESG operated in a continuous virtual consensus decision-making process, the online document and working statistics did not represent cohort actions, and the altered method provided enough stability to discover variables.	The researcher tracks the IETF management items to find any repeated items.	8 IESG meetings in 2016 meetings may be an outlier

Strand-2 document statistics of 4 months in 2016. During the eight IESG meetings from 4/21 to 7/7 in 2016, the IESG came to a consensus decision on 91 documents (90 publish the documents and one do not publish the document). By January 1, 2020, the RFC Editor (2020) had published only 88 of the 90 documents the IESG had approved. Because these two documents may be outliers, the researcher only examined the 88 published RFCs. The IESG (2020) made consensus decisions to publish these 90 documents in seven out of the eight meetings. The eighth meeting was a special meeting on May 10, 2016, where the IESG took a retreat to discuss strategic management discussions.

During this retreat, the IESG placed the normal processes within a meeting to review documents, WGs, and IETF management items on hold. Figure 17 shows a histogram of the dates for IESG approval of these 88 documents. The IESG document approval pattern shown in Figure 17 had maximums during the IESG meetings, indicating the IESG approves documents were in a continuous decision pattern because the IESG approved four to six documents between meetings. The pattern confirmed the researcher's suspicion that the IESG operated in a virtual consensus decision-making pattern.

Figure 17: IESG approves document for publication.



The IESG document publication process had post-document approval delays due to post-approval editing. For example, the IESG added 86 out of the 88 documents approved between 4/21/2016 to 7/7/20216 to the RFC Editor's (2016a) queue by August 2016. However, the IESG post-approval follow-up on two documents occurred after August 2016 (10/25/2016 and 2/8/2017). The IESG group's trust in the post-approval editing was an effective way to process documents, but it was another reason the researcher's original assumption on Strand-2 statistics was incorrect.

The editors in the RFC Editor (2016b) did not have a constant rate of publication or a constant time for processing a document. Figure 18 shows the publication date for the 88 documents referred to in the meetings from 4/21/2016 to 7/7/2016. The RFC Editor (2016a) group published most of the 88 documents between May and October 2016, but the delays in the RFC Editor (2016b) processing caused other documents to be published16 months later. A document may encounter delays in the RFC Editor's (2016b) processing due to dependencies on other documents or editorial issues. For example, the RFC Editors (2016b) held an RFC published as a standard (proposed, draft, or full standard) until all documents with which an RFC publication depended on for normative requirements were published.

Due to this mandate, RFC Editor (2016b) did not publish an IETF standard document until the IESG approved the standards with which it depended (normative standards). The RFC Editor's (2016b) delayed publication of two documents out of the 88 documents after August 2016 was due to a lengthy wait for normative requirement documents. Similarly, when an RFC document was declared "historic," the RFC Editor (2016b) assured that all documents referring to the now historic document indicated the historic status of that document. The five documents published in December 2017 were declared historic by the IESG on May 19, 2016 (see RFC Editor, 2020).

Figure 18: IESG sends documents to Request for Comments Editor.

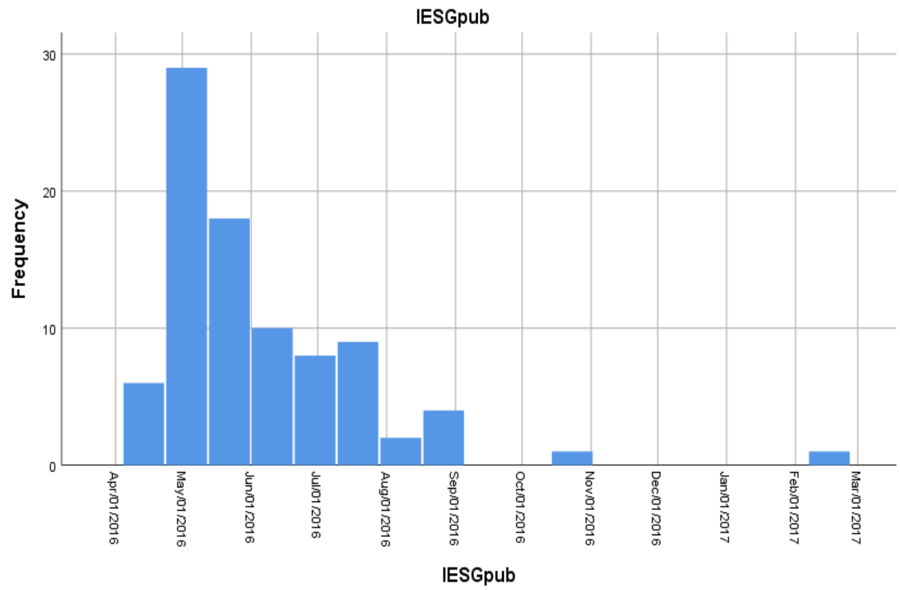
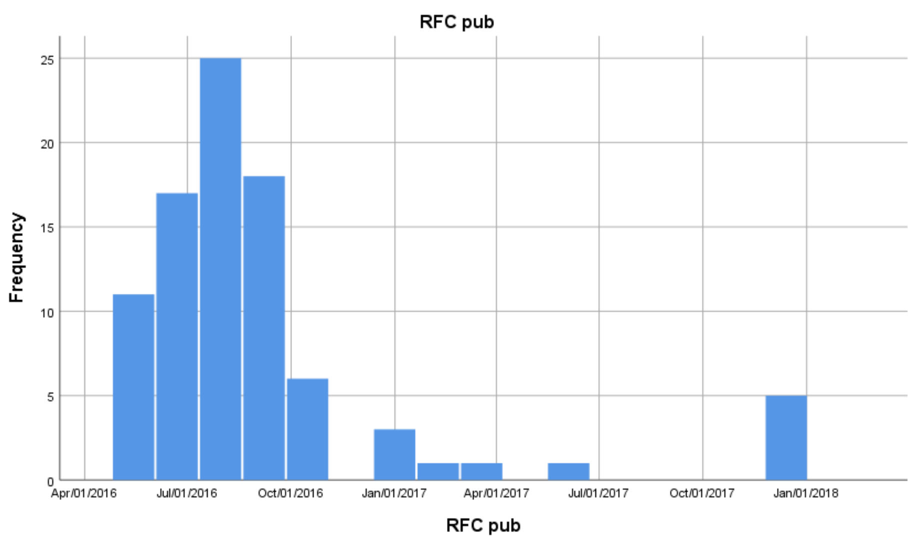


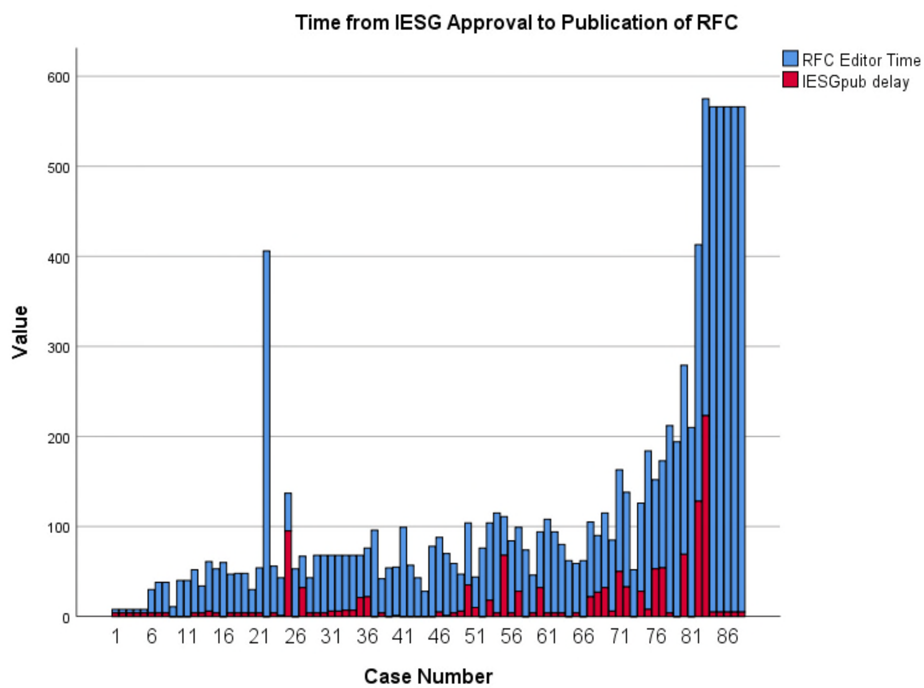
Figure 19: RFC editor's publication of documents.



One artifact of the IESG's continuous decision-making processes was that the time for the IESG to review a single document varies based on the content and the IESG discussions. In addition, the RFC Editors' (2020) processing time varied depending on whether a document was published as an individual standard or as a group of documents. Figure 20 shows how the IESG publication delays and the

RFC delays may combine to create publication delays of 200-580 days. These lengthy delays could cause errors in the yearly statistics on documents publication. For example, a document was approved in October 2016 with 200 days of delay appears in the 2017 statistics, but the 2016 IESG cohort reviewed and approved the document. As Figure 20 illustrates, 16 of the 88 documents (18%) approved from April 21 to July 7, 2016, had delays of over 150 days after approval before the document was published. This variance in the publication delays for documents could explain why the assumptions about online statistics were flawed in the original Strand-2 methods.

Figure 20: Time (days) from IESG approval to RFC Editor's publication for 88 documents.



Strand-2 working group statistics of 4 months in 2016. Two public documents reported the WG and BOF statistics: IETF Chair 7's report at IETF 96 (Arkko, 2016b) and the IETF BOF wiki (IETF, 2020b). This investigation compared the WG actions reported by the IESG in the 4 months of 2016 and the IETF (2020b, 2020d) BOF Wiki against the report given by the IETF (2016a, 2016c) chair. The IETF chair's report included BOFs for the IESG sponsored groups and new WGs approved by the IESG. Table 41 presents the data compared WG actions presented by the validated minutes and the IETF chair.

The IESG minutes reported 14 IESG actions comprised of creating four new WGs (eight IESG decisions), rechartering two workings (via three IESG actions), closing one working (via one IESG action), and replacing two chairs (two IESG actions). The IESG minutes did not mention IESG decisions to hold three BOFs. The public IETF reported eight IESG actions comprised of three BOFs and five new WGs.

Table 41: Working Group Actions

WG category	WG actions in IESG minutes	IETF chair's report
BOF call pseudo minutes	BOF Slots for pending WGs: babel, lamps (spasm), mtgvenue, and "sipbrandy." [IETF management action] BOFs: ITS, IMTG, L4S, LEDGER, LPWAN, LURK, PLUS, QUIC [8]	BOFs: ITS, IMTG, L4S, LEDGER, LPWAN, LURK, PLUS, QUIC (8)
New WG	babel (2), lamps (2) mtgvenue (2), and "sipbrandy" (2) [8 actions]	babel, dispatch, lamps (spasm), mtgvenue, "sipbrandy" [5]
Rechartered WG	lisp (1), manet (2)	None mentioned
Closed WG	mif (1)	None mentioned
Management of WG	I2rs new chair (1) Lamps new chair (1)	None mentioned
Total WG decisions	22	13

These two reports provided substantially different views on the same IESG (2020) actions, as the online data showed. Validating each action using data at the IETF website helped determine why the reports differed. The new WGs approved were the same except for the dispatch WG. The 2015 cohort approved the dispatch WG during the 3/3/2016 IESG meeting, but it did not meet until IETF (2016c). The IESG minutes recorded two decisions per WG: (a) approving the charter for public review in the IETF (denoted as IETF review) and (b) approving the WG to begin operation with a charter. The minutes indicated that the four WGs mentioned by the IETF chair and the IESG minutes (babel, lamps, mtgvenue, and "sipbrandy") had two IESG decisions per WG during the period 4/6/2016 to 7/7/2016.

The IETF chair's meeting report did not mention WGs that rechartered changes in chairs for WGs or closed because his report focused on involving participants in new work. However, the IETF (2020i) website did confirm the lisp WG rechartering on 5/6/2016, and the manet WG rechartering on 7/1/2016 (IETF, 2020j). The mif WG pages recorded the closure of the mif WG on 5/10/2016 (IETF, 2020l). The management items to add chairs could find validation in the WG pages for IETF (2020h) Chair 6 operating as LAMPS WG chair on 7/22/2016 and I2RS WG chair change on 6/2/2016 (IETF, 2020k). The IETF Chair's report or the IESG BOF wiki did not mention these WG rechartering and closure actions, but these were valid actions for IESG members. Based on this analysis, the researcher determined that the alternate Strand-1 and Strand-2 methods were valid for use with the 100% sample (2015 and 2016).

Strand-1 and Strand-2 Methodology Reviews After 2015 to 2016

The researcher found alternate Strand-1 and Strand-2 methods produced consistent encodings for the 100% sample for 2015 and 2016. However, after reviewing the encodings, the researcher noted three areas of concern: (a) inconsistency of encoding for rare events, (b) lack of unique identifiers for IETF management, and (c) data from theme counts for behaviors lacking normality. The researcher addressed the concern for encoding inconsistency for rare events by doing an extra review of the encodings for all rare events. Secondly, the lack of clear identification of unique IETF management items confirmed that some IETF management acts of decision-making (with and without results) were duplicates. Because the researcher used the IESG acts of decision-making with results in the 100% sample in 2015 and 2016 to validate the 10% sample of results for the same period, the results must be accurate. Therefore, the researcher created a list of IETF management items and included a cross-reference to each mention. Using this list, the researcher reviewed and adjusted the encoding for IETF management decisions (actions and results) for 2015 and 2016.

The researcher discussed the non-normality of the theme counts for 2015 and 2016 with research advisors and investigated best practices for historiometric mixed-mode methods. From one viewpoint, predictor themes (solidarity) and moderator (conflict) were unlike survey methods, so the lack of normality did not impact the results as long as Strand-2 statistics had a normal distribution over time (as shown with 10% analysis). A second viewpoint was that the alternate methodology for Strand-1 and Strand-2 matched the leadership research that utilized historiometric mixed-mode methods. Ligon et al. (2012) pointed out “at least 4 types of leadership data available through historiometry: 1) situational influences/conditions, 2) behaviors, 3) developmental events, and 4) sustainable performance” (p. 1106).

This research sought to map leadership consensus decision-making behaviors to the yearly sustainable performance per IESG cohort, which fit this model. Parry, Mumford, Bower, and Watts (2014) pointed out that historiometric research requires six methodological concerns: “1) theory, 2) sources, 3) controls, 4) samples, and 5) predictors and 6) criteria” (p. 141). This researcher tested a theory that increases in solidarity behaviors (with conflict moderators) caused a volunteer organization’s leadership to increase sustainable performance. The alternate methodology followed the best practices suggested by Parry et al. (2014) to select sources, set controls, and pick samples with sufficient possibility for predictors and criterion variables. The researcher concluded that the non-normality of the theme count data was less important than the combination of (a) qualitative validation of behaviors and results for the 10% sample and 100% sample, and (b) use of best historiometric practices.

Discovered variables remained ancillary qualitative and quantitative constructs helping to determine the construct validity of solidarity. The criterion variable, IESG actions with results, was based on individually validated actions, and this variable had a normal distribution for the 10% sample from 1991 to 2016. This research appeared to match the six methodological controls of the best historiometric practices that allowed quantitative analysis of the mixed-mode results of qualitative analysis (theme counts or Likert scales) to use descriptive and multivariate (correlation and regression) statistical methods.

Historiometric research had used regression on ratings (judgments) and objective indicators. For example, Ligon et al. (2012) reviewed 61 papers, and over

25 of these papers used some form of regression on a combination of ratings combined with objective indicators. Based on these practices in published research, the researcher's alternate methodology used the behavioral theme counts totals for predictor variables in multivariate statistical analysis for Strand-1 and Strand-4. In addition, the Strand-2 counts of the validated results as criterion varied in Strand-1 and Strand-3.

Potential Errors in Alternate Strands 1 and 2 Methodology (MAXQDA Limits)

At the start of the project, the MAXQDA (2016) program limits in 2016 required the researcher to transfer the theme counts data to interim Excel spreadsheets rather than a direct transfer. Although the researcher created the Excel spreadsheet with checking functions, any manual transfer could have resulted in errors. The spreadsheet uploaded to SPSS contains group behaviors per incident of IESG consensus and Strand-2 results per incident. The Excel spreadsheets from 2015 to 2016 also contained individual behaviors behind the group behaviors. The researcher compared the total theme counts for meetings and per year in the spreadsheet with the theme count totals from MAXQDA's (2016) mixed-mode analysis.

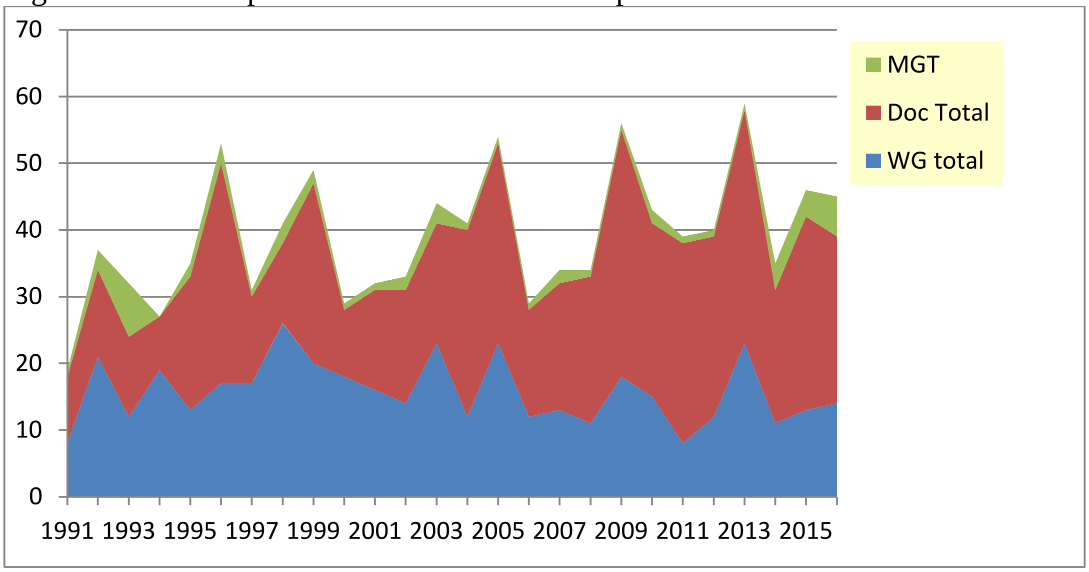
The manual transfer of Strand-1 data was needed to track the behaviors per incident of group decision-making rather than per IESG meeting. The researcher believed that consistent results with the 100% analysis for 2015 and 2016 would confirm the 10% Strand-1 and Strand-2 analytical results using the alternate methodology. Due to this belief, the researcher recorded the individual and group behaviors per incident of IESG decision-making for 2015 and 2016 for solidarity, conflict, OCB-GC, OCB-Altruism, TI, and the discovered variables. Because the group behaviors were the sum of individual behaviors, the Excel spreadsheets generated these totals per incident, meeting, and year for 2015 and 2016. The researcher compared the meeting and yearly theme count totals against the MAXQDA (2016) totals when validating the theme totals. Although the multiple methods of hand-checking on the generation and transfer of theme counts required a significant amount of time, the effort was necessary. The researcher detected errors detected in the hand transfer of data. Future researchers will need to improve this methodology by either using the MAXQDA (2016) programmatic interface to generate totals in a section of the source material or by creating a file per incident.

Strand-3 Methodology Changes

This research's original methods for Strand-3 planned to use statistics from the RFC Editor and IETF website. Strand-2 data collection investigations demonstrated that the online statistics did not provide accurate statistics for an IESG cohort. However, the complete analysis of the years 1991 to 2016 was beyond the timeframe of this research. Therefore, the alternate methodology for Strand-3 used the validated action theme counts found in 10% of the IESG minutes (biweekly plus BOF calls). The discussion on the Strand-2 revised methodology indicated the validated actions (with and without results) had a normal distribution. Figure 21 illustrates how the number of decisions in each category (WG, documents, and management items) varies within the 10% sample from 1991 to 2016. This variance shows differences between the IESG cohorts.

The small response for the 2017 survey required using the small model (solidarity, conflict, and IESG consensus decision-making results). In addition, several IESG members indicated they did not take the survey because they had responded to the survey on the same topic in 2013. Due to these comments, the researcher included data on the analysis of the 2017 survey and the 2013 survey. The researcher also examined if she could generate an estimate of a combination of the 2013 and 2017 survey responses to give additional qualitative information on the surveys. The qualitative analysis of the open-ended questions also had a smaller response, with several responses repeating over a few years. The smaller response tended to group responses closer to the IETF chair's term. Therefore, the researcher summarized the theme counts for the IPA analysis of the survey's open-ended questions on conflict for the period associated with an IETF chair's term.

Figure 21: Results per IESG cohort in 10% sample.



Strand-4 and Strand-5 Methodological Changes

The researcher reduced the Strand-4 triangulation due to the reduced research scope and the small responses from the 2017 IESG survey. The scope had five detailed datasets: (a) Strand-1 mixed-mode qualitative analysis using 10% of minutes from 1991 to 2016, (b) Strand-1 mixed-mode analysis for the 100% sample from 2015 and 2016, (c) Strand-3 survey from 2017, (d) Strand-3 survey from 2013, and (e) theme counts from the IPA analysis of the open-ended questions from the 2017 survey. Strand-4 methodology pulled the data from these five sources to triangulate to determine if the research hypotheses were supported. All triangulation of the models used the reduced model (solidarity, conflict, TI, and results of effective consensus decision-making).

The alternate method for Strand-5 had a reduced scope to match the reduced scope of the analysis. The data collection for Strand-5 collected a final summary from six sources: (a) Strand-1's mixed-model analysis of the IESG minutes, (b) Strand-2 summarization of the historical data on WG and IETF chairs, (c) Strand-3's qualitative of the open-ended questions on the survey, (d) Strand-3's qualitative analysis of the survey data, and (e) Strand-4's analysis of the qualitative data.

Strand-5 qualitative analysis used the triangulation of two theories to whether the qualitative data supported (a) the reduced theoretical model on solidarity and (b) theories regarding change leadership. The triangulation of two analysis theories helped determine if this research's solidarity model explained the group behaviors that enable the IESG as leaders of the IETF and SDO that helped improve Internet technology. The qualitative analysis used IPA analysis to detect master themes related to this research's reduced theoretical model.

For the second theoretical basis, the researcher used content analysis to determine if the IESG's leadership helped the IETF be an effective change organization in the global IT industry to accomplish the IETF's vision of creating a better Internet. For the content analysis, the researcher first examined the historical data on WGs and IETF chairs to determine if there was a progression of standards that led to improvements in Internet technology. If that progression existed, the researcher examined how the IETF chair and IESG influenced change in the global IT industry. For the change analysis, the researcher used SWOT (strength, weakness, opportunity, and threat) to determine the environment of the IESG cohorts under each IETF chair's leadership and a comparison of the vision of an IETF chair and the IETF's accomplishments under the IESG cohorts under that IETF chair. In the change analysis, the researcher considered how solidarity and conflict impacted the IESG leadership per IETF chair's tenure.

Qualitative Analysis of Need for Methodological Changes

The qualitative test for a methodology was to answer the following two questions: (a) Does the revised theory resemble reality, and (b) does the revised list of data analyzed represent reality? The theory-reality question could be restated for this research as the following: Does a continuous virtual consensus decision-making theory match the historical realities? The data analyzed versus reality entailed the following: Does the data in the IESG minutes and BOF call augmented by online data at the IETF match reality? One must turn to the history of the beginning IETF and the changes in online tools since 1987 to answer the theory-reality question.

The National Science Foundation (NSF) selected the IETF in 1987 as the organization to standardize technology and operations for regional networks to connect the NSF-funded research backbone. As a part of the NSF's cost-reduction plan for high-energy physics, this backbone connected high-energy physics sites in the United States and Europe, replacing costly direct links. Researchers and universities who participated in the NSF-funded network were in the United States and Europe, so the IETF and IESG were spread across multiple time zones. Group decisions could not wait for costly in-person meetings but were in teleconferences. The global dispersion of IETF members and IESG members existed from 1991 to 2017.

Based on this historical start, the evolution to a continuous consensus decision-making process using online forums to discuss documents and WGs decisions was an evolution toward more effective decision-making. A continuous process would be necessary to effectively manage a standards organization whose business is quickly creating and fielding standards. In such an environment, the biweekly teleconferences were times for summarization or specific debates rather than a detailed decision review. The online discussion forum for the IESG was called a ballot system. This online system was recognized as early as December 7, 1992, in the IESG minutes, as making the IESG more effective in making decisions on protocol actions. The IETF ballot concept fit a virtual consensus decision-making group.

The ballot concept allowed the comments on documents, WG charters, and management items to be shared online prior to a meeting. This online ballot forum allowed IESG members to provide comments and state their disposition on an IESG decision. This disposition toward approval could be "yes," "no," or "discuss," where "discuss" meant the individual had a concern about an item that needed discussion before approval. A "discuss" indicated at least a TC due to a difference of opinion, but other types of conflict levels might also be involved. The ballot system allows IESG members to start the consensus process in the virtual environment. This virtual consensus decision-making might include individual or group behaviors of solidarity or conflict. If IESG reached a consensus on a decision in the premeeting ballot process, the IESG noted the approval in their meeting. The consensus to approve could be full accord or tentative based on some condition. For example, the tentative conditions in documents included having specific edits made to a document or adding a note with text attached to the document or instructions for the RFC Editor (2016b) during the editing process. Tentative approval for a WG involved approving a WG based on changes to the charter for the WG. Tentative approval for a management item included checking, validation, or research for the management action. If the IESG failed to reach a consensus on a decision during the meeting, a follow-up discussion could occur in the virtual environment or a subsequent meeting. Individual and group behaviors of solidarity and conflict occurred in this virtual environment. The IESG continued to discuss a document or WG charter in the virtual ballot environment to resolve a “discuss” ballot until the IESG reached a group consensus (approve or reject).

The original tools of the early IESG (2000) started with the tools that many networks used to share documents and information on decisions. The early IESG minutes (1991 to 1993) made oblique references to the shared computer files, email, or sources. These oblique references were understandable as the goal of the minutes was to describe the IESG actions and decisions. Early IESG minutes assumed the IETF (2016b) secretariat and the RFC Editor (2016a, 2016b, 2020) kept some email exchanges, public files with versions of the documents (Internet-Drafts or RFCs), and private files. Other early mail documenting IESG activities could be obtained from universities (if one has the time) or from the NSF archive. Unfortunately, other historical records and email exchanges were on servers at defunct companies, so those historical records were no longer available. The growth of the IETF caused the IETF to develop the IETF (2016b) Datatracker aid communication within the IESG communication and provide transparency on IESG debates and decisions to the IETF community.

Initially, the IETF (2016a) systems kept track of versions of the files for documents and the WG charters. The IETF (2016b) Datatracker was enhanced in 2001 to start recording events in the life cycle of a document and versions of the documents. Continuous upgrades in the IETF (2016b) Datatracker resulted in IESG comments on the document being kept in the data tracker starting in 2002. A summary of these comments began to appear in documents approved in July 2003 (see RFC3550 as the first document; RFC Editor, 2020). The IETF (2016b) Datatracker began keeping information on comments from individual IESG members on WGs in working ballots in August of 2012 (see IPSECME WG ballot; IESG, 2019). Since 2002, the IETF (2016b) Datatracker information had been an online forum for communication. The data recorded in the written IESG (2019) minutes contained pointers to the IETF's (2016b) online forums. Conclusions made about the IESG minutes gave inconsistent results until the researcher examined the other data available in the IETF (2016b) Datatracker.

Results from Strands 1-4

This research's alternate mixed-mode methodology used historiometric best practices to prepare the five datasets, which consisted of two IPA analyses of the IESG minutes (10% sample and 100% sample), the IPA analysis of the 2017 survey's open-ended questions (OEQ) on conflict, and the two surveys (2017 survey and 2013 survey) for quantitative analysis. Ligon et al. (2012) described six interrelated steps in which five of these steps ("model specification," "sample plan formation," "content-coding scheme development," "material preparation," and "coding logics;" p. 1118) should occur before quantitative analysis using descriptive statistics, and multivariate analysis begins. This section summarizes the five datasets' descriptive statistics and statistical analysis according to research completed in Strands-1–4.

Strand-1 resulted in the report on the descriptive statistics, reliability of scales, suitability for statistical analysis (correlation and HRM), correlations, and HRM modeling for the research on historical records in the IESG minutes. The descriptive statistics provided data on actual sample sizes, the individual behaviors found in IPA analyses, the discovered themes, and theme counts for behaviors under study by theme count totals and weighted diagrams. The researcher created group behaviors as the sum of individual behaviors under study (solidarity, conflict, TI, and OCB) exhibited in an individual's action(s) on a group consensus decision.

The 10% sample (1991 to 2016) had a mean of 11.7 individual behavioral actions per IESG decision, and the 100% sample had a mean of 11.0 individual behavioral actions per decision. Strand-1 discovered themes specific to the IESG, which thanked others who aided the IESG member(s) or Flag issues for other IESG members. The researcher combines the two discovered themes, “ThankAid” and “FlagIssue,” into “Discovered IESG” (D-IESG) in the reduced model.

Strand-2 descriptive statistics provided the per IESG cohort actions and results estimated by 10% sample for 1991 to 2016 per cohort and discovered in the 100% sample from 2015 to 2016 (see Table 53). The researcher used Strand-2 theme counts as counts in the statistical analysis Strands-1–4 because the published historiometric research used theme counts similarly The descriptive statistics in Table 53 also indicate the results as a percentage of the theme counts for decisions per IESG cohort. The IPA analysis of the 10% sample of IESG minutes found that 61% of the decisions had a measurable result rather than continued discussion. The measurable results for the 100% sample for the 2016 IESG cohort were also 61% of the decisions. The 100% sample of the 2015 cohort indicated the decisions with measurable results were 64% of all decisions. Strand-2 results included the demographics of who made decisions, the type of decisions made, and the results. The demographics of the 10% sample of formal minutes indicated that the IESG delegated 13% of the decisions to a single person but that two or more people made 87% of the decisions as consensus decisions. Similarly, in the 100% sample of IESG minutes of the 2015 IESG Cohort, the IESG delegated 17.4% of the decisions to one person and 82.6% to two or more people. The IESG cohort in 2016 (based on the IESG minutes) delegated 13.6% of the decisions to one person and 86.4% of the decisions to two or more people.

The statistical analysis of Strand-3’s 2017 and 2013 surveys were presented as part of the Strand-4 comparison of the statistical analysis for Strands-1–4. Appendix Q provides an in-depth analysis of the 2017 survey (Section Q.3) and the 2013 survey (Q.4). The Strand-3 analysis of the 2017 survey found that in some facets, this survey acted as a retest of the 2013 survey, but it performed as a unique survey in other ways. Appendix Q also has an in-depth analysis of the risks due to the retesting factor, leading to fewer respondents. The Strand-4 comparison examined the descriptive statistics for historical data in the 10% sample of the IESG (1992, 2020) minutes from 1991 to 2016 (Dataset 1), the 100% sample from 2015 to 2016 (Dataset 2), the 2013 survey (Dataset 3), the 2017 survey (Dataset 4), and the quantitative data from the open-ended conflict questions in the 2017 survey. The Strand-4 triangulation results included a risk analysis on data validity risks and bias, as discussed later in this chapter, and qualitative analysis of Strand-4's quantitative results.

Strand-1 Results

The historical records for the IESG minutes analyzed for the 10% sample from 1991 to 2016, and 100% sample from 2015 to 2016 had two types: formal and narrative. These minutes were hand-merged into a single logical stream to produce a unique set of events to be analyzed. IESG (2020) minutes were references to the portions of the IETF's (2019) online databases, so the researcher's data from the online databases were compiled into the IESG minutes as code notes. The formal IESG minutes were available from July 1991 to March 2017 (end of 2016 Cohort year). Some narrative IESG minutes were available in September 2005, but full coverage for narrative minutes only began with the 2007 IESG cohort. The 10% sample analyzed 78 minutes of the IESG formal minutes comprised of 52 minutes (2 per year) and 26 BOF minutes (1 year), and 35 of the narrative minutes (23 IESG minutes and 12 BOFs). The 100% sample for 2015 contained the historical records pointed to by the minutes for 23 IESG biweekly meetings, one IESG retreat, one IESG meeting at IETF (2016c) 95, and three BOF minutes. The 100% sample 2016 included IESG for 23 IESG biweekly meetings, one IESG retreat meeting, and three BOF minutes.

Table 42: Samples – Interpretive Phenomenological Analysis and Surveys

IPA study	Cohort years	Total minutes formal	Total minutes narrative	Total minutes BOF	Minutes analyzed formal	Minutes analyzed narrative	Minutes analyzed BOF	Minutes analyzed total
10%	1991 to 2016	599	246	78	52	23	26 (12)	78 (35)
100%	2015	25	25	3	25	25	3	28
100%	2016	24	24	3	24	24	3	27
OEQ	1989 to 2016	16 responses		32 comments		19 Task	32 Relationship	

Survey Date	Cohort years	Potential IESG Responses: IESG Members All	IESG Members active	Cohort slots All	Cohort slots active	IESG Responses: Valid People	Valid Years	Cohorts Covered
2017	1989 to 2016	97	81	356	317	25	28	94
	1991 to 2016	96	80	337	299	25	26	88
2013	1989 to 2013	89	76	315	289	41	25	129
	1991 to 2013	88	75	295	269	41	23	125

Survey Date	cohorts	IETF chairs totals: All	Active	Cohort slots All	Cohort slots active	IETF chair responses: Valid People	Valid Years	Cohort slots
2017	1989 to 2016	7	5	28	23	4	15	15
	1991 to 2016	7	5	26	23	4	15	15
2013	1989 to 2013	7	6	28	20	5	16	16
	1991 to 2013	7	6	26	20	5	16	16

Note. There were seven IETF chairs (1991 to 2016), but three of the chairs served as members.

One question in the methodology was which type of minutes provided the best references to the online data to study the behaviors of solidarity, conflict, TI, and OCB in consensus decision-making. The IPA analysis determined the number of people who could act in a group behaviors per IESG decision by encoding the individual actions of IESG with an individual behavior action (IBA) theme, and the sum of IBA themes is the number of people who participated in a decision. Each IBA in the historical records either denoted OCB behaviors or a lack of OCB behavior (OCBnb). The sum of the individual theme counts per behavior studied is the group theme. The IPA of the 10% formal minutes from 1991 to 2016 found 21,643 IBAs in 1853 decisions in 26 years, with 11.7 IBAs per decision and 71.3

decisions per year. The IPA analysis of the narrative minutes in the 10% sample 2005 to 2016 found 11,687 actions in 12 years with a mean of 13.1 individual actions per decision and a yearly mean of 89.3 actions. The narrative minutes had more individual actions (IBAs) per decision since the narrative minutes do not record as many decisions since they failed to record between meeting decisions. The IESG cohorts often reached consensus decisions between meetings through a portion of the IESG (usually two to five people) to resolve open issues or TCs. From 2005 to 2016, the formal minutes pointed to online data that replicated most of the text from the narrative minutes. However, the narrative minutes contained in-meeting discussions, unlike the formal minutes.

After determining the link to the online data in the 10% analysis, the researcher carefully tracked differences between the data referred to by the formal minutes, the narrative minutes, and the researcher's hand-merge of the two minutes (denoted as the combinational). For the IESG minutes from the 2015 to 2016 cohorts, the richness of the online data made the hand-merge or combinational minutes the preferred data source. The combinational minutes of the 2015 IESG cohort had 8816 IBAs in 820 decisions and a mean of 10.8 IBAs per decision. The combination minutes of the 2016 IESG cohort had 8721 IBAs in 785 decisions and a mean of 11.1 IBAs per decision. The average number of individuals exhibiting one of the studied behaviors (solidarity, conflict, TI, and OCB) is between 10.8 and 11.7 individuals in both IPA studies. This consistency indicates data collection process for Strands-1 and 2 selected a consistent set of source material to be analyzed by IPA encoding. This statistic indicated a successful design and implementation of the alternate methodology to select and prepare consistent historical material to be analyzed for these behaviors.

Table 43: Strand-1 – Theme Analysis

Dataset	Study	IBA	Decision	Cohort Years	IESGMtg	Decisions by type		
						Doc	WG	IETF-mgt
DS1	10%	21643	1853	26	78	945	505	403
DS2	100%	17543	1605	2	55	923	195	487
	2015	8816	820	1	28	474	111	235
	2016	8721	785	1	27	449	84	252

Note. There were seven IETF chairs in the period from 1991 to 2016, but three of the chairs served as members.

This section lists the theme counts for the reduced model behaviors found in the 10% sample and the 100% sample in Table 53. The solidarity behavior in the 10% sample had a yearly mean theme count of 257.9 themes, which was 31% of the yearly mean for IBAs, and the solidarity theme counts per year range between 23% to 36% of total IBAs per year. The solidarity behavior in the 100% sample for 2015 had 3,244 themes detected in 8,816 IBAs (37% of the total IBAs). Similarly, the solidarity behavior in the 100% sample for 2016 had 3,249 themes detected in 8721 IBAs (37%). Solidarity, as a percentage of individual actions within group decisions, was similar in the 10% and 100%, but higher. The conflict behavior in the 10% sample had a yearly mean theme count of 72.7 themes, which was 9% of the IBAs, and theme counts ranged between 4% to 15% of the total IBA theme count per year. The conflict behavior found in the 100% samples was 433 themes in 2015 (5% of all IBAs) and 564 themes in 2016 (6% of all IBAs). The theme counts for conflict discovered in the 100% analysis of the IESG minutes from 2015 and 2016 fit within the lower end of the range of yearly conflict themes estimated by the 10% sample. The open-ended questions on conflict indicated damping of recording of conflict in the IESG minutes. The TI behavior in the 10% theme count had a mean yearly theme count of 295.8 (36% of the total IBAs) and yearly theme counts, which range between 29% to 44% of the total IBA. The TI behavior theme count from the 100% sample from 2015 was 2,926 themes (33% of the IBAs), and from 2016, it was 3,075 (35% of the IBAs). The TI theme count totals for the IESG cohorts in 2015 and 2016 fit within the middle of the range of TI theme count totals for IESG cohorts from the 10% sample. The yearly mean for OCB theme counts in the 10% was 773.8 or 93% of the total IBAs, and the OCB theme counts per year ranged between 77% to 95% of the total IBAs. The theme counts for the OCB behavior detected in the 100% sample of 2015 IESG minutes was 8,096 (92% of the total IBAs for 2015), and the theme counts for the 100% of the 2016 IESG minutes was 8,195 (94% of the total IBAs). The theme counts for OCB for 2015 to 2016 fit at the top of the estimated range for OCB theme counts from the 10% sample. The descriptive statistics on the theme counts showed that the IPA encodings for the reduced model and the full model fell within the same ranges. The descriptive statistics were essential to the Strand-1 analysis because the scale reliability tests on theme counts did not have the same results as Likert scales.

Qualitative validity of theme counts. This qualitative analysis used theme grids and theme counts to validate the encoding themes. The researcher generated theme grid and weighted diagrams for the theme counts for behaviors in the full model and the reduced model for the behavioral counts in the 10% sample from 1991 to 2016 (Dataset 1) to determine the typical pattern of the theme counts. A full description of these theme counts and weighted diagrams can be found in Appendix O Section O.2. The weighted theme counts drawings for solidarity, conflict, TI, and OCB from the 10% sample are included in Figures 24 to 27 for the formal and narrative minutes. These theme counts show a similar pattern for formal and narrative minutes per behavior.

Across all the behaviors, the solidarity, conflict, TI, and OCB patterns, the weighted links in the theme counts were thicker for behaviors exhibited during the discussion of document consensus decisions. For example, solidarity themes for HS and VS had thicker links for solidarity questions that indicated solidarity to aid others to finish tasks. The process worked best during the consensus process if there was mutual aid to reach a consensus decision. Conflict themes had the thicker lines for TC due to differing opinions and RC, which caused tension. TI theme node diagrams indicated a heavy dependence on co-workers, and OCB indicated a heavy concentration of doing actions required by formal requirements. All of these qualitative patterns pointed toward the descriptive analysis being true. The researcher generated theme grids and weighted node diagrams from the theme counts of the behaviors in the full model and the reduced model for the 100%

sample from 2015 to 2016 (Dataset 2). The weighted diagrams for solidarity, conflict, TI, and OCB from the 100% sample were visually identical to the ones between the narrative and formal minutes.

The researcher checked the themes discovered in the 10% sample of the IESG minutes to determine if these themes were in the 100% sample (2015 to 2016). These discovered themes might be specific to the IESG review process. The researcher grouped the discovered themes under the following top-level themes in the full model: ThankAid, FlagIssues, status-change, and OCBnb. The researcher found the ThankAid themes were detected in circumstances when the IESG members publicly thanked other people for helping the group come to a consensus. The FlagIssues themes occurred when the IESG members flagged issues for other IESG members to consider before coming to a consensus decision. The status-change theme occurred when the IESG members changed the status of a document, such as raising a standard from a proposed standard to a full standard. The “full” standard status indicated an IETF standard whose technology had been implemented and deployed in many networks. The OCBnb theme indicated when IESG members did not perform organizational duties expected of them during the consensus decision process.

The researcher combined the ThankAid, FlagIssues, and change-status into the reduced model's Discovered-IESG theme (D-IESG). The total number of D-IESG themes found in the 10% sample was 1035 (5% of the total IBA), with an average of 39.8 themes per year. The total number of D-IESG themes found in 2015 was 1379 themes (14% of the IBAs), and the total D-IESG themes found in 2016 was 1234 (14% of the IBAs). The D-IESG themes in the 10% sample had a range of 0 to 138 themes per year (0%-14% of the IBAs). However, the minutes for the 1991 to 2004 IESG cohorts and associated online databases lacked the detail to detect these themes. Figure 26 contains the weighted node diagram for the D-IESG theme, which shows that most of these themes came when the IESG members thanked authors and directorate reviewers and flagged issues in documents with multiple comments (“pile-on” comments) and changed status from proposed standard to historic status. The IESG may decommission no longer used standards by declaring the standards as “historic.” The D-IESG theme group provides concurrent and discriminant theme patterns from solidarity in the historical record tied to the IESG functions from the descriptive statistics. The correlation results for the 10% analysis and the 100% confirm these descriptive statistics.

The OCBnb theme had a mean of 60.9 themes per year in the 10% sample (7% of the IBAs), but the OCBnb themes occurred in 826 of decisions (45%). This statistic means one or more of the IESG members who should have participated in the decision-making process did not participate in 45% of the decisions. The decisions delegated to a small group only expected members to participate in the decision-making process. The theme count for the OCBnb in the 2015 IESG minutes data was 1,032 (12% of IBA) in 449 decisions (55% of the 820 decisions), and the theme count for OCBnb in the 2016 IESG minutes was 787 (9% of the IBA) in 414 decisions (53% of the 785 decisions). The OCBnb theme was grouped with the OCB theme in the weighted node diagrams below to contrast the two themes.

Does the OCBnb theme indicate the “free-rider” condition exists in all the IESG cohorts where some IESG members ride along in the IESG consensus decisions without providing expected input? This interesting question was set aside because the focus of the OCBnb theme was to confirm that the OCB construct was detected correctly in the IPA encodings. The weighted node diagrams for OCB and OCBnb in Figure 25 show that the OCBnb theme occurs more frequently during document review cycles. Because document review and evaluation takes time, an overloaded IESG might skip one or two documents per meeting. The appendices provide the full details on the theme count totals, the weighted node diagrams, and the descriptive statistics for discovered themes for the full theoretical model (ThankAid, FlagIssue, and OCBnb) and the reduced model themes (Appendix O for the 10% sample and Appendix P for the 100% sample of 2015 to 2016).

Figure 22: Solidarity behaviors (from Dataset 1: 10% analysis).

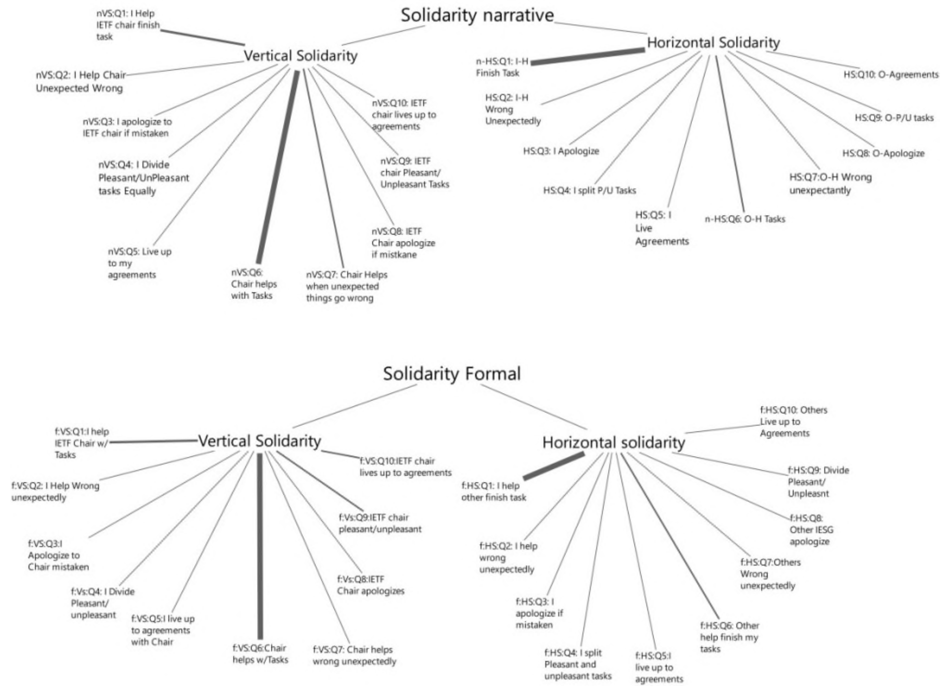


Figure 23: Conflict: Weighted node diagram (10% analysis).

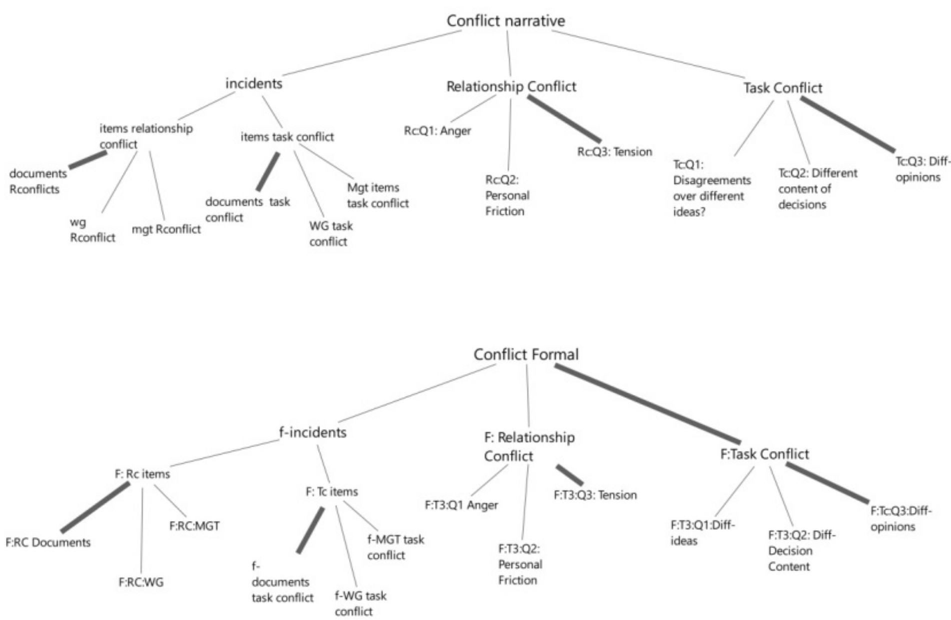


Figure 24: Task interdependence weighted node diagram.

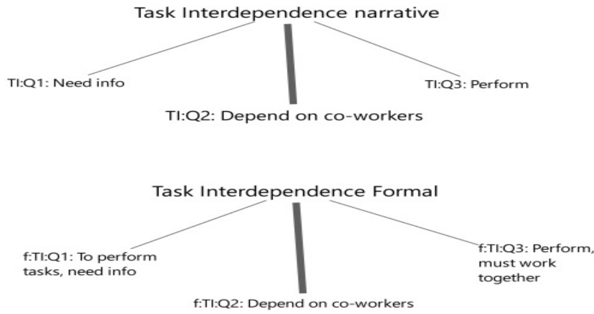


Figure 25: Organizational citizenship behaviors weighted node diagram.

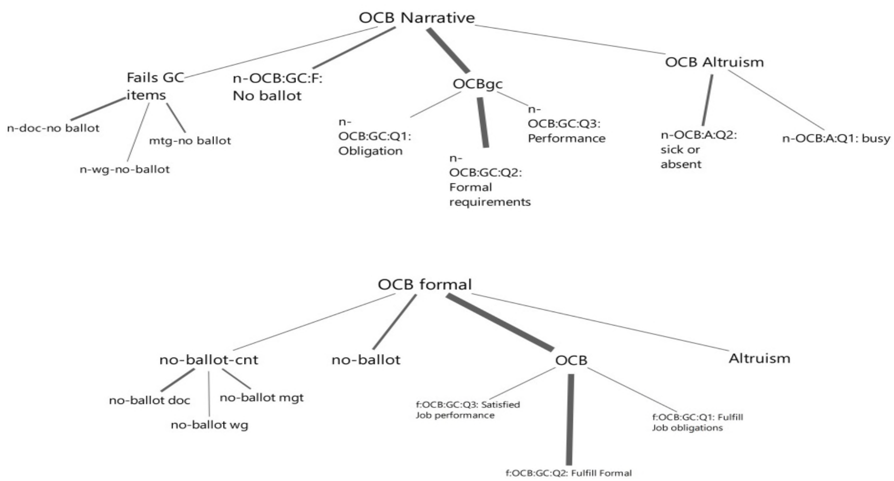
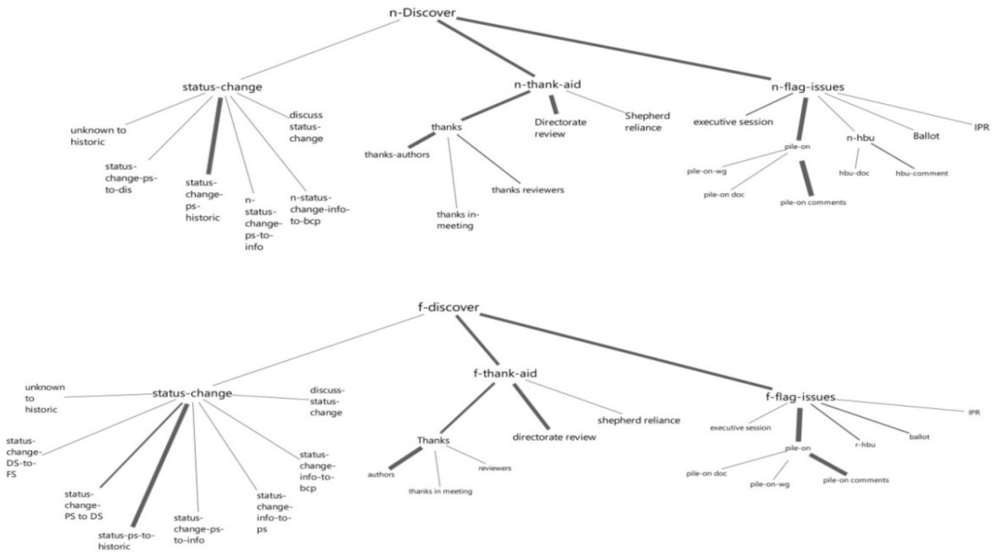


Figure 26: Discovered Internet Engineering Steering Group weighted node diagram.



Strand-1 – Suitability for multivariate statistics. Suitability for multivariate statistical analysis assumed a valid scale for the behavior plus behavioral data with normality, homoscedasticity, linearity, and errors results with a normal distribution. The theme counts as a quantitative measurement of behaviors in a historical text had different characteristics than Likert scales used in historical text reviews or survey responses. Theme counts depended on the existence of the behaviors within a textual document.

As described, the descriptive statistics on the behavioral theme count and the weighted diagrams suggested the behavioral scales for the full and reduced model are valid. The researcher conducted scale reliability tests on the theme counts for the reduced model behaviors in the 10% sample per meeting and per year (see Appendix O Section O.6.1.2) and 100% sample per meeting for each of the 2 years (2015 to 2016; see Appendix P Section P.6.2.2). Table 54 contains the Cronbach alpha on standardized items values per scale (S, C, TI, OCB, and DI) for Datasets 1 and 2 (10% IPA analysis 1991 to 2016 and 100% analysis 2015 to 2016) and Datasets 3 and 4. The 10% sample per meeting and the 100% samples per meeting have reliability scale numbers in the range of 0.606 to 0.916 except for conflict, which ranges (0.463 to 0.884). Due to the differences, the theme count reliability scales contained information rather than disqualifying factors for the multivariate statistics.

The researcher conducted normality, homoscedasticity, and linearity tests on theme count data in Dataset 1 (10% sample) and Dataset 2 (100% sample) for behavioral variables and behavioral variables against the results found in Strand-2. The Strand-2 data were gathered as theme counts in the revised process. Section O.6.4 in Appendix O contains the details of the normality, homoscedasticity, and linearity tests for the full and reduced model for Dataset 1 (10% sample), and section P.6.4 in Appendix P contains these results for Dataset 2 (100% sample). These sections also contain a quantitative and qualitative analysis of the error terms to determine if the errors have constant variance, independence, and a normal distribution. The analysis of Dataset 1 found that the reduced model behaviors (S, C, TI, OCB, and DI) were suitable for multivariate analysis per year (1991 to 2016, $N = 26$), but only (S, TI) were suitable per meeting (1991 to 2016, $N = 78$). The analysis of Dataset 2 found the reduced model behaviors of (S, C, TI, OCB, and DI) were suitable per meeting for 2015 ($N = 28$) and 2016 ($N = 27$).

The results from Strand-1's correlation and hierarchical regression modeling. The researcher conducted correlation tests and HRM modeling tests on behavioral data from theme counts found suitable for multivariate analysis for the reduced and full model in Dataset 1 (10% sample) and Dataset 2 (100% sample). However, this section only discusses the reduced model correlations and HRM results for Strand-1's data analysis of Dataset 1 (10% sample 1991 to 2016) and Dataset 2 (100% sample 2015 to 2016) used in the Strand-4 comparison of the Datasets 1 to 5 (two IPA analysis, two surveys, and one IPA of open-ended conflict questions). The reduced model behavioral variables examined in correlation and HRM are (S, C, TI) and two alternate models based on OCB (OCB, C, TI) and Discovered-IESG (DI, C, TI). The behavioral data examined in correlations and HRM modeling are the theme counts from hand-merged formal minutes (1991 to 2016) for Dataset 1 and the hand-merged combinational minutes for Dataset 2 (2015 to 2016). Appendix O contains the results of the correlation and HRM modeling tests for Dataset 1 (10% sample 1991 to 2016) in Sections O.6.3 and

O.6.5, respectively. Appendix P contains the complete details of the correlation results for Dataset 2 (100% for 2015 and 2016) in Sections P.6.3 and P.6.5. Table 44 summarizes the following correlation tests for the reduced model and the alternative reduced models: (a) between behavioral variables and (b) between behavioral variables and results. Table 45 contains key results of HRM modeling for Strand-1.

Table 44: Strand-1 – Correlation Results

Model	1991 to 2016 (*N = 26*)	2015 per meeting (*N = 28*)	2016 per meeting (*N* = 27)
The reduced model between behaviors	S-C** (0.716) S-TI** (0.938) C-TI** (0.727)	S-C** (0.616) S-TI** (0.991) C-TI** (0.679)	S-C** (0.588) S-TI** (0.986) C-TI** (0.591)
Reduced model behaviors-results	S-R** (0.845) C-R* (0.409) TI-R** (0.738)	S-R** (0.804) C-R* (0.545) TI-R** (0.798)	S-R** (0.661) C-R* (0.544) TI-R** (0.632)
OCB model between behaviors	S-OCB** (0.902) C-OCB** (0.728) TI-OCB** (0.938)	S-OCB** (0.919) C-OCB** (0.692) TI-OCB** (0.918)	S-OCB** (0.895) C-OCB** (0.684) TI-OCB** (0.877)
OCB model behaviors-results	OCB-R** (0.784)	OCB-R** (0.855)	OCB-R** (0.844)
DI model between behaviors	S-DI** (0.683) C-DI* (0.448) TI-DI** (0.568) DI-OCB** (0.577)	S-DI** (0.892) C-DI** (0.721) TI-DI** (0.568) DI-OCB** (0.842)	S-DI** (0.690) C-DI** (0.494) TI-DI** (0.687) DI-OCB** (0.719)
DI model Behavior-results	DI-R* (0.433)	DI-R* (0.743)	DI-R* (0.649)

** - correlation significant at $\rho < 0.01$, * - correlation significant at $\rho < 0.05$
R = Results

The behavioral data for the reduced model from the theme counts from Dataset 1 (1991 to 2016), and Dataset 2 (2015 to 2016) had a strong positive correlation between variables (S, C, TI) at a significance of $\rho < 0.01$. These reduced model behaviors also had a medium to strong positive correlation with the results at a significance of $\rho < 0.01$, with one exception. Conflict (C) had a medium correlation with results (0.409 for 1991 to 2016, 0.545 for 2015, and 0.544 for 2016) at a significance of $\rho < 0.05$. The positive correlation of conflict might come from positive theme counts for conflict. The correlation tests on the two alternate models based on OCB and Discovered-IESG (DI) had similar results to the reduced model with one exception. The correlation tests found a medium to strong positive correlation between DI and C with a significance of $\rho < 0.05$ with 2015. The correlation tests found that solidarity had a stronger positive correlation than DI in Dataset 1 (1991 to 2016) and Dataset 2 and a strong positive correlation than OCB in Dataset 1 (1991 to 2016). OCB had a stronger positive correlation with results in Dataset 2 (0.855 in 2015 and 0.844 in 2016) than the solidarity variable had with results (0.804 in 2015 and 0.661 in 2016). The correlations were similar but different, indicating a unique construct for solidarity.

The researcher ran HRM tests using the HRM4 modeling sequence for the reduced model on the theme counts from Dataset 1 (10% sample) and Dataset 2 (100% sample). These HRM tests found solidarity explained 44% to 65% of the variance in the results per IESG cohort at a significance of $\rho < 0.01$. The solidarity explained 62.2% of the variance in results per IESG cohort in Dataset 1 (10% sample for 1991 to 2016). Solidarity explained 64.6 % of the variance in the results per meeting for the 2015 IESG cohort and 43.7% of the variance in results per meeting for the 2016 IESG cohort. The reduced model without the TI control variable (HRM5 modeling sequence) explained 77.8% of the variance in Dataset 1 (1991 to 2016) with solidarity explaining 62.2% (beta = 1.163) and conflict explaining 16.4% (beta = -0.556) at a significance of $\rho < 0.01$, but the results for the 100% sample in Dataset 2 remained the same. These results support Hypothesis 1 for the reduced model but fail to support Hypothesis 2.

The researcher tested two alternate models using hierarchical regression models using the modeling sequences (HRM4 and HRM5). The first alternate model substitutes OCB for solidarity and the second alternate model replaces solidarity with DI (Discovered IESG behaviors). The HRM tests substituting OCB for solidarity in the HRM4 and HRM5 modeling sequence found that OCB explained 61.5% to 73.1% of the variance in the results per IESG cohort for Dataset 1 and Dataset 2 at a significance of $\rho < 0.01$ with one exception. The HRM-4 modeling used on the data from 2016 in Dataset 2 (100% sample) found the OCB alternate model explained 76.3% of the variance in the per meeting results for the IESG 2016 results at a significance of $\rho < 0.05$. This modeling results found that OCB explains 71.2% of the variance in per meeting results (beta = 1.254), and conflict explains 5.6% of the variance (beta = -0.467). These results pointed to a difference between solidarity and OCB to explain the variance in 2016 for the IESG minutes. The HRM tests substituting DI for solidarity in the HRM4 modeling sequence explained 67.3% of the variance (DI: 18.8%, TI: 35.8%, and C: 12.7%) in Dataset 1 (1991 to 2016) at a significance of $\rho < 0.01$, 63.9% of the variance for 2015 at a significance of $\rho < 0.05$ (DI 52.2%, TI 8.7%), and 42.1% of the variance in 2016 at a significance of $\rho < 0.01$. The HRM modeling using the HRM5 modeling sequence without the TI control variable found only DI explained the variance in the results per meeting from the 100% sample (Dataset 2). The results of the alternate models pointed to a significant difference between the solidarity, OCB, and Discover-IESG (DI) behaviors. These results supported the construct validity of solidarity.

Table 45: Strand-1 – Reduced Model Hierarchical Regression Modeling Results

Dataset	Dataset 1 (formal, per year)	Dataset 2 (combinational, per meeting per year)		Hypotheses validated	
N	1991 to 2016 (*N = 26*)	2015 (*N = 28*)	2016 (*N* = 27)	H1	H2
HRM4 (S, TI, C, SxC)	62.2% variance at $\rho < 0.01$ $F(1,24) = 39.539$ S: 62.2%, beta = 0.789	64.6% variance at $\rho < 0.01$ $F(1,26) = 47.478$ S: 64.6%, beta = 0.804	43.7% variance at $\rho < 0.01$ $F(1, 26) = 19.434$ S:43.7%, beta = 0.661	3 Yes	3 No
HRM4-alt1 (OCB, TI, C, OCBxC)	61.5% variance at $\rho < 0.01$ $F(1,24) = 38.325$ OCB: 61.5%, beta = 0.784	73.1% variance at $\rho < 0.01$ $F(1,26) = 70.762$ OCB: 73.1%, beta = 0.855	76.3% variance at $\rho < 0.05$ $F(2,24) = 38.553$ OCB: 71.2%, beta = 1.254 C: 5.6%, beta = -0.467	3 Yes[*1]	3 No[*1]
HRM4-Alt2 (DI, TI, C, DIxC)	67.3% variance at $\rho < 0.01$ $F(3, 22) = 15.064$ DI: 18.8%, beta = 0.047 TI: 35.8%, beta = 1.090 C: 12.7%, beta = -0.520	63.9% variance at $\rho < 0.05$ $F(2,25)= 22.155$ DI: 55.2%, beta = 0.125 TI: 8.7%, beta = 0.685	42.1% variance at $\rho < 0.01$ $F(1,25) = 18.189$ DI: 42.1%, beta = 0.659	3 Yes[*1]	3 No[*1]
HRM5 (S, C, SxC)	77.3% variance at $\rho < 0.01$ $F(2, 23) = 39.150$ S: 62.2%, beta = 0.789 C:15.1% , beta = -0.556	64.6% variance at $\rho < 0.01$ $F(1,26) = 47.478$ S: 64.6%, beta = 0.804	43.7% variance at $\rho < 0.01$ $F(1,26) = 19.434$ S:43.7%, beta = 0.661	3 Yes	3 No
HRM5-Alt1 (OCB, C, OCBxC)	77.8% variance at $\rho < 0.01$ $F(2,23) = 40.401$ OCB: 61.5%, beta = 1.163 C: 16.4%, beta = -0.363	73.1% variance at $\rho < 0.01$ $F(1,26) = 40.401$ OCB: 73.1%, beta = 0.855	71.2% variance at $\rho < 0.01$ $F(1,25) = 40.401$ OCB: 71.2%, beta = 0.844	3 Yes[*1]	3 No[*1]
HRM5-Alt2 (DI, C, DIxC)	18.8% variance at $\rho < 0.05$ $F(1, 24) = 5.543$ DI: 18.8%, beta = -0.433	55.2% variance at $\rho < 0.01$ $F(1,26) = 32.027$ DI: 55.2%, beta = 0.743	42.1% variance at $\rho < 0.01$ $F(1,25) = 18.189$ DI: 42.1%, beta = 0.649	3 Yes[*1]	3 No[*1]

<u>*Strand-2 Results*</u>

The total number of decisions and decisions with results per IESG cohort were detected in IESG minutes and validated against online sources at the IETF website (www.ietf.org) and the IETF online database (datatracker.ietf.org). Table 53 contains the number of decisions and the number of decisions with results per cohort year. Descriptive statistics were run on the decisions to determine who was involved in the decision (passively and active participation) and what types of decisions were made by an IESG cohort. Active participation in consensus decision-making involves an individual exhibiting one or more behaviors during the decision. Passive participation involves the person agreeing to the decision. The researcher tracked active participation using the IBA variable (Individual behavior actions). The group delegated some decisions to one person or a group of people.

Table 46 provides the statistics on who was involved in the decision in Dataset 1 (10% sample) and Dataset 2 (100% sample), and Table 47 provides the statistics on the types of decisions found in Dataset 1 and Dataset 2. This section examines whether the results from these descriptive statistics for Datasets 1 to 2 and weighted theme nodes grids for action themes support the premise that Strand-2 mixed-mode methods created valid and consistent data for actions and results.

Dataset 1 (10% sample), and Dataset 2 (100% sample) have similar ranges for the number of decisions, who made the decisions, types of decisions, types of successful decisions, and ranges for successful decisions. The average number of consensus decisions per IESG cohort is 713 for the 10% sample for the IESG cohorts from 1991 to 2016, and 61% of these decisions had a measurable result. The IPA detected 820 decisions in the 2015 IESG minutes, with 64% of these decisions having a measurable result. The IPA analysis of the 2016 IESG minutes found 785 decisions, and 61% of these decisions had a measurable result. The active participation ranged from 11.6 people per decision in Dataset 1 (1991 to 2016), 10.6 people per decision in the 2015 IESG minutes, and 11.1 people in 2016 IESG minutes. The IESG delegated between 13% to 17% of the decisions to a single individual in the IESG minutes examined. The remainder of the decisions were consensus decisions made by the whole IESG or a subgroup (2 to 14).

Interestingly, the IESG 2015 delegated 17% of the decisions and had a slightly higher success rate (64%). The demographics of the types of the decisions were between 51-58% documents (51% for 1991 to 2016, 58% for 2015, 57% for 2016), 11% to 27% WG actions (27% for 1991 to 2016, 14% for 2015, and 11% for 2016), 22% to 32% IETF management (22% for 1991 to 2016, 29% for 2015, 32% for 2016). These quantitative demographics showed similar results to the weighted theme count diagrams for 1991 to 2016 and 2005 to 2016 shown in figures 27 and 28, where the majority of the IESG decisions focused on IETF standard documents. The weighted node diagrams show the IESG minutes have the ratio of the types of decisions is consistent, but the earlier years in the IETF had more reports of the WG actions due to large BOFs. The demographics of the decisions culminating in measurable results also followed this pattern, with most results occurring in document decisions (29% to 38% of total decisions), followed by WG actions (9-22% of total decisions) and IETF management (9% to 14% of total decisions).

There are two reasons Dataset 1 had a larger percentage of WG actions. The first reason was that the online data on early years (1991 to 2005) only reported successful BOF meetings in the IETF proceedings, so rejected BOF proposals were under-reported. This under-reporting was a shortcoming of the IESG minutes and associated IETF online data, but estimating any additional data added risk to the analysis. The second reason was that the IESG approved more BOFs during the early years (1991 to 2006) than the later years. Based on the quantitative and qualitative similarities in Datasets 1 and 2, the researcher concluded that Strand-2 data on results is reliable.

Table 46: Strand-2 – Decisions Demographics

Decisions	2015			2016			‘91 to ‘16
	Formal	Narrative	Combo *1	Formal	Narrative	Combo	Formal (Est. *4)
Yearly total	820	679	820	785	667	785	713
Who decides (IBA)	8771	8521	8816	8721	8454	8721	8324
Mean IBA per decision	10.70	12.55	10.75	11.11	12.67	11.11	~11.7
Mean IBA per meeting	313.26	303.82	314.36	323.00	313.11	323.00	277.47
Meetings per year	28*2	28*2	28*2	27*3	27*3	27*3	*4

People involved in decision	Level of Involvement in Decisions (by percentage)						
	2015			2016			‘91 to ‘16
	Formal	Narrative	Combo	Formal	Narrative	Combo	Formal (Est *4)
1	17.6%	13.8%	17.4%	13.6%	12.3%	13.6%	13%
2-5	15.4%	5.3%	15.4%	16.8%	6.3%	16.8%	
6-9	0.4%	0.1%	1.2%	0.6%	0.3%	0.6%	
10-14	0.0%	0.0%	0.0%	0.1%	0.1%	0.1%	
15	59.8%	68.5%	56.7%	56.9%	67%	56.9%	87%
16-18	7.0%	12.2%	10.1%	11.3%	13.3%	11.3%	
19-20	None	None	None	0.5%	0.6%	0.5%	

Theme counts for decisions	2015			2016		
	Formal	Narrative	Combo	Formal	Narrative	Combo
1 person	144	94	143	107	82	107
2to 20 people	676	573	677	678	585	678

	Dataset 1: Formal Minutes (1991 to 2016) Decisions			Dataset 1: Narrative Minutes (2005 to 2016) Decisions		
10% sample predictions (IBA)	% of total	Mean per meeting	Mean per year *4	% of total	Mean per meeting	Mean per year (estimate)*4
1 person	13%	3.16	97	16%	4.11	120
2 to 20 people	87%	20.60	616	84%	21.29	621

*1 Combo = Combinational minutes

*2 - 23 meetings, 3 BOF calls, 1 retreat, 1 IETF 95 meeting

*3 - 23 meetings, 3 BOF calls, 1 retreat

*4 – Estimate based on 2 meetings + 1 BOF call per year

Table 47: Strand-2 – Types of Decisions

	2015				2016			
Decisions[*1]	Formal	%	Narrative	%	Formal	%	Narrative	%
All types	820	100%	679	100%	785	100%	667	100%
Documents	474	58%	333	49%	449	57%	331	50%
WG	111	14%	111	16%	84	11%	84	13%
IETF mgt	235	29%	235	35%	252	32%	252	38%
Results	526	64%	426	63%	478	61%	383	57%
Documents	313	38%	213	31%	295	38%	200	30%
WG	102	12%	102	15%	72	9%	72	11%
IETG mgt	111	14%	111	16%	111	14%	111	17%
No Result	294	36%	253	37%	307	39%	284	43%
Documents	161	20%	120	18%	154	20%	131	20%
WG	9	1%	9	1%	12	2%	12	2%
Management	124	15%	124	18%	141	18%	141	21%

Decisions	2015 to 2016				DS1:1991 to 2016		DS1: 2005 to 2016	
	Formal	%	Narrative	%	Formal	%	Narrative	%
All types	1605	100%	1346	100%	1853	100%	889	100%
Documents	923	58%	664	49%	945	51%	416	47%
WG	195	12%	195	15%	505	27%	247	28%
IETF mgt	487	30%	487	36%	403	22%	226	25%
All Results	1004	63%	383	60%	1130	61%	456	51%
Documents	608	38%	20	31%	539	29%	176	20%
WG	174	11%	72	13%	416	22%	182	20%
IETG mgt	222	14%	11	16%	175	9%	87	11%
No Results	601	37%	537	40%	723	39%	433	49%
Documents	315	20%	251	19%	406	22%	240	27%
WG	21	1%	21	1%	89	5%	65	7%
IETF mgt	265	16%	265	20%	228	12%	128	14%

Figure 27: Action theme counts from Dataset 1.

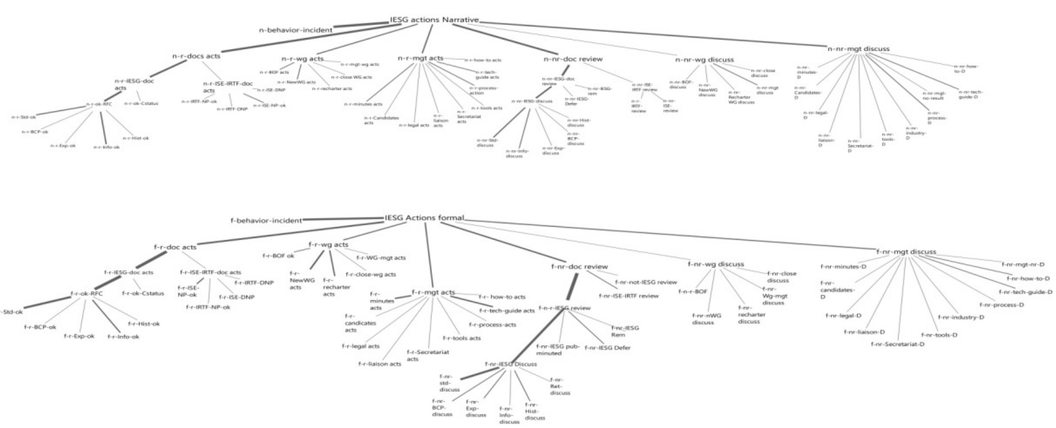
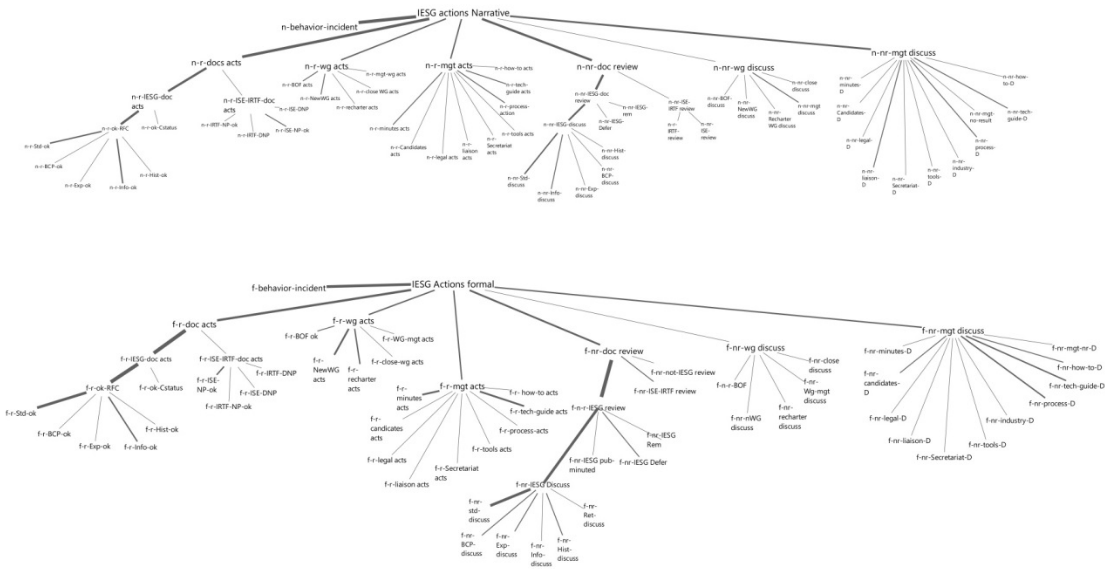


Figure 28: Actions theme counts from Dataset 2.



Strand-3 Results

Strand-3 results encompassed the 2017 survey (Dataset 4), the 2013 survey (Dataset 3), and the mixed-mode analysis of conflict open-questions from the 2017 survey. The 2017 survey received 25 valid responses covering 28 IESG cohorts and 94 cohort slots. The researcher defined one cohort slot as one IESG member position for one IESG cohort year. The 2017 survey received four valid IETF chair responses covering 15 IESG cohorts. The 2013 survey received 41 valid IESG responses covering 25 years (1989 to 2013) and five IETF chair responses covering 16 IESG cohort years. The open-ended questions received 16 responses with 32 comments that created 19 TC themes and 32 RC themes (see Table 42). The researcher conducted analytical tests for scale reliability, descriptive statistics, multivariate analysis suitability, correlation, and HRM modeling on the behavioral data from the 2013 and 2017 surveys for all IESG responses, IESG Cohort means, and the IETF chair responses. Because the IPA analysis only covered 1991 to 2016, t these analytical tests covered two periods: one starting in 1989 (1989 to 2013 and 1989 to 2016) or 1991 (1991 to 2013 and 1991 to 2016).

Table 52 summarizes the Strand-3 analysis test results and a similar test for the IPA analysis (10% and 100%). Table 53 summarizes Strand-3 descriptive statistics per cohort and similar results for Dataset 1. Table 54 the scale reliability test results for Dataset 1-4 plus the reliability scores (Cronbach alpha from published research on solidarity (HS, VS), conflict (TC, RC), and TI. Appendix Q contains the full details of the analysis for the open-ended questions (Section Q.2), the 2017 survey (section Q.3), and the 2013 survey (Section Q.4). Table 48 lists HRM modeling sequences for variants of the reduced model with a dependent variable of results from Strand-2 for the 2013 and 2017 survey and a dependent variable of perceived results (PR). The HRM-4 modeling sequence is the original reduced model with solidarity (St, TI, Ct, and SxC), and the HRM-5 modeling sequence is the HRM-4 without TI as a control variable. Table 49 and Table 50 provide the results of the HRM modeling tests. This section reviews Strand-3 results.

The researcher tested the behavioral data from 2013 and 2017 surveys for scale reliability (S, TI, OCB, PR, and C; 2017 only) for two time periods (1989 to 2016) and (1991 to 2016). The scale reliability tests on the reduced behavioral data (S, C, TI, and OCB) from the 2017 survey for both periods (1989 to 2016 and 1991 to 2016) found Cronbach alpha values for standardized items between 0.800 to 0.916. Table 54 contains the results of the scale reliability tests for the responses from the 2017 survey. These Cronbach alpha values were equal to or higher than the Cronbach alpha values found by previous research studies into solidarity and OCB by Koster and Sanders (2006, p. 258), solidarity and TI for teams by Sanders and Schyns (2006b, pp. 543–544), and conflict using Jehn’s (1995) scale by Pearson et al. (2002, p. 122).

The 2013 survey was used to conduct reliability tests for data from two periods (1989 to 2013 and 1991 to 2013) and found similar results for the reduced behavioral models (S, TI, and OCB) except for OCB (see Table 54). The scale reliability tests on OCB behavioral data from 1989 to 2013 have a Cronbach alpha on standardized items value of 0.685, and the same OCB data from 1991 to 2013 has a Cronbach alpha of 0.669. These values were below the 0.700 value Koster and Sanders (2006, p. 258), so these scales were unreliable. The researcher created the IESG perceived results scale (PR) as an instrument unique to the IESG. The scale reliability tests for PR found that Cronbach alphas for standardized items ranged between 0.765 and 0.869, so these scales can be considered reliable.

The descriptive statistics were run on the reduced model behaviors to determine mean values for all responses from the IESG members, ranges for the IESG cohort means, IETF chair mean values for all responses, and ranges from IETF chair responses. Table 52 list these values for the 2013 and 2017 surveys. All instruments in the 2013 and 2017 surveys used 7-point Likert scales, allowing easy comparison. The mean scores for all IESG responses for the cohort solidarity, TI, and OCB were 0.2 to 0.6 higher in the 2013 survey than in the 2017 survey but centered around *agree* or *somewhat agree* on the Likert 7-point scale. The mean scores per IESG cohort for the solidarity, TI, and OCB behavior had a more extensive range in the 2017 survey (*undecided to agree*) than the 2013 survey (*agree somewhat to agree*). The PR mean score is 0.2 lower in the 2013 survey than the 2017 survey, and the mean scores per IESG cohort for PR have a more extensive range in the 2017 survey (*undecided* to *agree strongly*) than the 2013 survey (undecided to agree). The conflict scale for all IESG responses on the 207 survey scored 2.96 (1991 to 2016) that a respondent *somewhat disagreed* that conflict existed during their IESG cohort. The group means score per IESG cohort for conflict ranged from 2.39 to 3.75 (*disagree* to *undecided*) on the 2017 survey.

The IETF chair responses from the 2013 survey and 2017 survey were similar due to the small number of IETF chair responses. There were four valid IETF chair responses for the 2017 survey and five valid IETF chair responses for the 2013 survey. The range of means of IETF chair responses per IESG cohort for the behaviors (solidarity, TI, and OCB) on the 2013 and 2017 surveys were similar, centering around the *agree* to *strongly agree* values. The mean for conflict in the 2017 IETF chairs survey responses was 2.93 (*somewhat disagree*), with a per IESG cohort range from 1.5 to 3.5 (*strongly disagree* to *some disagree*). The descriptive statistics pointed to responses that aligned with the reduced model (S, C, and TI) and the alternate, reduced model based on OCB (OCB, C, TI). The IETF chair responses had a higher average score (6.24) on the 2017 survey than on the 2013 survey. The descriptive statistics for the 2013 and 2017 surveys supported the reduced theoretical model.

The descriptive statistics for the Perceived Results Scale from the IESG responses and the IETF chair responses aligned Strand-2's results as a percentage of all decisions. The mean value for PR from all IESG responses for the 2013 survey was 5.13 (*somewhat agree*), and for the 2017 survey, the mean was 5.34 (*somewhat agree*). This score aligned with the Strand-2 statistics, which indicated that 61% of the decisions had results. The mean for perceived results from the IETF chairs survey was 5.60 in the 2013 survey and 6.24 in the 2017 survey. The IETF chairs had more positive views of their results than the IESG members. The range for PR was 4.17 to 5.17 (*undecided* to *agree somewhat*) in the IESG responses from the 2013 survey, and the range was 4.17 to 6.33 (*undecided* to *agree strongly*) in the IESG responses from the 2017 survey. The PR scores in the IETF chairs responses were 4.2 to 7.0 in the 2013 survey and 6.0 to 7.0 in the 2017 survey. This range of PR scores showed a substantial variance in the group opinions of IESG cohorts on whether their IESG cohort effectively made consensus decisions. These opinions varied from being undecided if their cohort was effective to strongly agreeing that their IESG cohort was effective in decision-making. This variance of opinions matched the variance in the Strand-2 statistics on the percentage of decisions (40% to 96%) that culminated in results.

The researcher ran suitability tests for the reduced mode variables (S, C, and TI) and the alternate model variable (OCB) using Strand-2 theme counts for results (R) as the dependent variable and PR as an alternate dependent variable. These suitability tests included normality, linearity, and homoscedasticity tests, plus tests for non-normality of error terms. Because suitability tests in the historiometric studies used Likert-7 for behavioral scales and counts for results, this methodology aligns with past research. The data from all IESG responses from 2013 had no behaviors suitable for correlation or HRM testing. On the other hand, solidarity and PR were suitable for multivariate analysis for all IESG (2020) data groupings (e.g., 2017 all IESG responses, 2013 IESG cohort means, 2017 IESG cohort means). The TI variable was only suitable for multivariate analysis when using the data from the 2013 survey summarized as cohort means and the 2017 survey data summarized per cohort means (1991 to 2016). This researcher examined group data from the IESG cohort means (1991 to 2016); thus, the researcher could test the reduced model hypotheses using the 2013 and 2017 survey data. The OCB behavior was only suitable for the 2017 survey for 1991 to 2016. This multivariate suitable test result indicates that solidarity is different from OCB.

The researcher investigated the relationships between the between behaviors in the reduced model (S, C, and TI) and the alternate model (OCB, C, and TI) and between these behaviors (S, C, TI, and OCB) and the results (real and perceived) using the Pearson product-moment correlation coefficient for the survey groups which were suitable for correlation. Before testing for correlation, the researcher determined if the survey data had enough statistical power to make correlation tests meaningful. The rule of thumb for simple regression is 15 participants per predictor in a survey (Pallant, 2010, p. 150). Therefore, a simple correlation between two variables or between a variable and the dependent variable (results) was within the statistical power of the data from the 2013 and 2017 surveys.

The researcher ran correlation tests on data for all responses for IESG cohort slots for the 2013 survey (N = 129) and the 2017 survey (N = 94) and mean values for the IESG cohort responses per IESG cohort for the 2013 survey (1989 to 2013; N = 25) and the 2017 survey (1989 to 2016; N = 28). Based on all IESG cohort response data, no intra-behavior correlation existed for the 2013 survey, but the 2017 survey solidarity negatively correlated with conflict (-0.627 for 1989 to 2016 and -0.558 for 1991 to 2016). However, the group means per IESG cohort from the 2013 survey for solidarity positively correlated with Strand-2 results (R) and Strand-3's PR. This positive correlation between solidarity and results (R) had a value of 0.517 for 1989 to 2013 IESG cohorts at a significance level of $\rho < 0.01$ and 0.470 for 1991 to 2013 IESG cohorts at a significance level of $\rho < 0.05$. The positive correlation between solidarity and perceived results based on the group means from the 2013 survey is 0.531 for 1989 to 2013 and 0.541 for 1991 to 2013. The correlation tests on the group mean per IESG cohort from the 2017 survey responses found intra-behavior correlations between solidarity and conflict (S-C), TI (S-TI), OCB (S-OCB), and PR. These intra-behavior correlations based on cohort group means from the 2017 survey indicated that solidarity negatively correlates with conflict at a value of -0.722 (1991 to 2016) but positively correlates with TI and OCB at a significance level of $\rho < 0.01$. Based on the same group means, the conflict negative correlates to TI (-0.718 for 1991 to 2016 with a significance of $\rho < 0.01$) and TI positively correlates to OCB (0.476 for 1991 to 2016 with a significance of $\rho < 0.05$). Based on the 2017 survey's group means, none of the behaviors correlate with Strand-2 results (R), but all of the behaviors correlate with the PR behavior (S, C, TI, OCB) for 1991 to 2016. Correlation positively correlates with PR (0.764 for 1991 to 2016) and negatively correlates with conflict (-0.396 for 1991 to 2016) at a significance of $\rho < 0.01$. Based on the group mean per IESG cohort, PR positively correlate with the Strand-2 results (R) with a value of 0.538 for the 2013 survey (1991 to 2016) at a significance level of $\rho < 0.01$. These correlation results support the reduced model fit for the Strand-2 results for the 2013 survey and the reduced model for perceived results for the 2013 and 2017 surveys. One of the open questions is why the 2017 survey did not correlate with Strand-2 statistics on results.

The researcher ran HRM model sequences for the reduced model on the 2017 and 2013 survey data. The dependent variable for these HRM modeling tests was the Strand-2 results, and the alternate dependent variable was perceived results from the survey. The researcher ran the modeling sequences in Table 48 on the 2013 and 2017 survey responses summarized per IESG cohort as a group mean and all IESG responses from the 2017 survey. The HRM modeling sequence HRM-4 found behaviors reported on the 2017 survey did not predict the variance in Strand-2 results for 1989 to 2016 or 1991 to 2016. The 2013 survey behaviors summarized per IESG cohort group mean found solidarity predicted 28.1% of the variance in the Strand-2 results at $\rho < 0.01$ for the period 1989 to 2013 and 29.3% of the variance in the Strand-2 results at $\rho < 0.05$ for the period 1991 to 2013. The alternate model HRM-4 modeling sequence using OCB did not predict the Strand-2 results for 2013 or 2017. The HRM-4-PR modeling tests on the 2017 survey behavioral data summarized per IESG cohort as a group mean found solidarity predicted 50.8% (1989 to 2016) to 58.3% (1991 to 2016) of the perceived results at $\rho < 0.01$ and 49.8% (1989 to 2016) to 52.0% (1991 to 2016) of the perceived results when all IESG responses considered. The HRM-4-PR modeling tests on the 2013 survey found the behavioral data summarized per cohort mean predicted 28.1% (1989 to 2013) to 29.3% (1991 to 2013). The researcher tested the operation TI as a control variable by removing the TI from the HRM-4 modeling sequences to create the HRM-5 series of modeling runs (HRM-5-R, HRM-5-PR, HRM-5-R-Alt-OCB, and HRM-5-PR-Alt-OCB). The HRM-5 modeling tests found similar results to HRM-4 runs, so removing the TI variable did not seem to impact the results significantly. The HRM modeling tests in Strand-1 and Strand-3 found support for Hypothesis 1 but did not support Hypothesis 2.

The open-end conflict questions on the 2017 survey revealed the duality between the RC recorded in the IESG minutes and intra-group conflict, which percolates under the surface. The open-ended comments reported that IESG members within the IESG tried to have the IETF chair removed or actively worked against the IETF chair during some periods. During other periods, factions within the IESG leadership actively worked against the majority of the IESG members. During other periods, IESG members expressed conflict passively by working against one or more consensus decisions. The IETF chair's ability to lead the group to address the changing technology and standards environment qualitatively seems to have changed the levels of conflict and stress.

Due to this qualitative observation, the researcher summarized the conflict themes for IESG cohorts under the leadership of an IETF chair. The IPA of the open-ended questions found 19 themes for TC and 32 themes for RC in 16 responses. Even though the IETF chair periods range from 2 to 6 years, the average number of themes per IETF chair is 2.7 TC themes and 4.6 RC themes. The small number of themes restricts the mixed-mode quantitative analysis to descriptive statistics and the plot in Figure 31.

Figure 29: Strand-3 theme counts for relationship conflict and task conflict in open-ended questions.

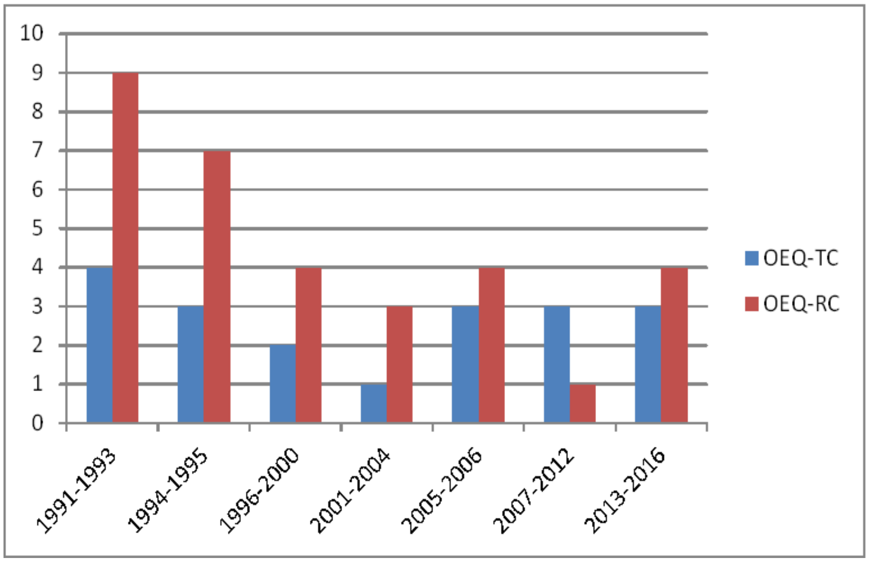


Table 48: Strand-3 Reduced Model HRM-4 and HRM-5 Sequences

Sequence	Predictors with Dependent Variable Result				
HRM-4-R Steps	HRM-4-R 2017 steps	HRM-4-R 2013 steps	HRM-4-R-Alt-OCB 2017 Steps	HRM-4-R-Alt-OCB 2013 Steps	Hypothesis Proved
1 (a)	Predictors: (Constant), St	Predictors: (Constant) St	Predictors: (Constant), OCB	Predictors: (Constant), OCB	H1
2 (b)	Predictors: (Constant), St, TI	Predictors: (Constant), St, TI	Predictors: (Constant), OCB, TI	Predictors: (Constant), OCB, TI	
3 (c)	Predictors: (Constant), St, TI, Ct	-	Predictors: (Constant), OCB, TI, Ct	-	H2
4 (d)	Predictors: (Constant), St, TI, Ct, SxC	-	Predictors: (Constant), OCB, TI, Ct, OCBxC	-	
HRM-4-PR Steps	HRM-4-PR 2017 steps	HRM-4-PR 2013 steps	HRM-4-PR-Alt: OCB 2017 steps	HRM-4-R-Alt-OCB 2013 Steps	Hypothesis Proved
1 (a)	Predictors: (Constant), St	Predictors: (Constant), St	Predictors: (Constant), OCB	Predictors: (Constant), OCB	H1
2 (b)	Predictors: (Constant), St, TI	Predictors: (Constant), St, TI	Predictors: (Constant), OCB, TI	Predictors: (Constant), OCB, TI	
3 (c)	Predictors: (Constant), St, TI, Ct	-	Predictors: (Constant), OCB, TI, Ct	-	H2
4 (d) Not done	Predictors: (Constant), St, TI, Ct, SxC	-	Predictors: (Constant), OCB, TI, Ct, OCBxC	-	
HRM-5-R Steps	HRM-5-R 2017 steps	HRM--R 2013 steps	HRM-5-R-Alt-OCB 2017 Steps	HRM-5-R-Alt-OCB 2013 Steps	Hypothesis Proved
1 (a)	Predictors: (Constant), St	Predictors: (Constant) St	Predictors: (Constant), OCB	Predictors: (Constant), OCB	H1
2 (b)	Predictors: (Constant), St, Ct	-	Predictors: (Constant), OCB, Ct	-	H2
3 (c)	Predictors: (Constant), St, Ct, SxC	-	Predictors: (Constant), OCB, Ct, OCBxC	-	

Sequence	Predictors with Dependent Variable Result				
HRM-5-PR Steps	HRM-5-PR 2017 steps	HRM-5-PR 2013 steps	HRM-5-PR-Alt: OCB 2017 steps	HRM-4-R-Alt-OCB 2013 Steps	Hypothesis Proved
1 (a)	Predictors: (Constant), St	Predictors: (Constant) St	Predictors: (Constant), OCB	Predictors: (Constant), OCB	H1
2 (b)	Predictors: (Constant), St, Ct	-	Predictors: (Constant), OCB, Ct	-	H2
3 (c)	Predictors: (Constant), St, Ct, SxC	-	Predictors: (Constant), OCB, Ct, OCBxC	-	

Table 49: Strand-3 HRM-4 Results (2017 Survey and 2013 Survey)

Real Results	HRM-4-R cohort mean		HRM-4-R all responses		HRM-4-R-Alt-OCB cohort mean	Hypothesis	
2017	1989 to 2016 (*N* = 28)	1989 to 2016 (*N* = 26)	1989 to 2016 (*N* = 94)	1991 to 2016 (*N* = 88)	1991 to 2016 (*N* = 26)		
Total variance [*1]	None	None	None	None	None	HS1	HS2
Significant variables	-	-	-	-	-	N:5	N:5
2013	1989 to 2013 (*N* = 25)	1991 to 2013 (*N* = 23)	1989 to 2016 (*N* = 94)	1991 to 2016 (*N* = 88)	cohort mean (*N* = 26)	Hypothesis	
Total variance	Step 1: 26.7% $\rho < 0.01$ S (26.7%) $F(1,23) = 8.388$	Step 1: 22.1% $\rho < 0.05$ S (22.7%) $F(1, 21) = 5.947$	-	-	-	HS1	HS2
Significant variables	at $\rho < 0.01$ S (beta = 0.517)	at $\rho < 0.01$ S (beta = 0.470)	-	-	-	Y:2	N:2
Perceived Results	HRM4-PR Cohort Mean		HRM4-PR All Responses		HRM-4-PR-Alt-OCB cohort mean	Hypothesis	
2017	1989 to 2016 (*N* = 28)	1991 to 2016 (*N* = 26)	1989 to 2016 (*N* = 94)	1991 to 2016 (*N* = 88)	1991 to 2016 (*N* = 26)		
Total variance [*1]	Step 1: 50.8% $\rho < 0.01$ S (50.8%), $F(1, 26) = 26.872$	Step 1: 58.3% $\rho < 0.01$ S (58.3%), $F(1, 24) = 33.574$	Step 1: 49.8% $\rho < 0.01$ S (49.8%), $F(1, 92) = 91.355$	Step 1: 52.0% $\rho < 0.01$ S (52.0%), $F(1, 86) = 93.257$	Step 1: 27.2% $\rho < 0.01$ OCB (27.2%) $F(1, 24) = 8.966$	HS1	HS2
Significant variables	at $\rho < 0.01$ S (beta = 0.713)	at $\rho < 0.01$ S (beta = 0. 0.764)	at $\rho < 0.01$ S (beta = 0.706)	at $\rho < 0.01$ S (beta = 0.721)	at $\rho < 0.01$ OCB (beta = 0.522)	Y:5	N:5

Perceived Results	HRM4-PR Cohort Mean		HRM4-PR: All Responses		HRM4-PR-Alt-OCBCohort mean	Hypothesis	
2013	1989 to 2016 ($N = 25$)	1991 to 2016 ($N = 23$)	1989 to 2016 ($N = 94$)	1991 to 2016 ($N = 88$)	1991 to 2016 ($N = 26$)		
Total variance[*1]	Step 1: 28.1 $\rho < 0.01$ *S* (28.1%), $F(1, 23) = 9.008$	Step 1: 29.3% $\rho < 0.01$ *S* (29.3%), $F(1, 21) = 8.708$			Step 1: 27.2% $\rho < 0.01$ OCB (27.2%) $F(1, 24) = 8.966$	HS1	HS2
Significant variables	at $\rho < 0.01$ *S* (beta = 0.531)	at $\rho < 0.01$ *S* (beta = 0.541)			at $\rho < 0.01$ OCB (beta = 0.522)	Y:3	N:3

Table 50: Strand-3 HRM-5 Results (Theoretical Reduced Model Without Task Interdependence)

Real Results	HRM-5-R Cohort mean		HRM-5-R All responses		HRM-5-R-Alt-OCB Cohort mean	Hypothesis	
2017 (DS-4)	1989 to 2016 (N = 28)	1989 to 2016 (N = 26)	1991 to 2016 (N = 94)	1989 to 2016 (N = 88)	cohort mean (N = 26)		
Total variance	None	None	None	None	None	HS1	HS2
Significant variables	---	---	---	---	---	N:5	N:5
2013 (DS-3)	1989 to 2016 (N = 25)	1991 to 2016 (N = 23)	1991 to 2016 (N = 94)	1989 to 2016 (N = 88)	cohort mean (N = 26)	Hypothesis	
Total variance	Step 1: 26.7% $\rho < 0.01$ S (26.7%), F (1, 23) = 8.388	Step 1: 22.1% $\rho < 0.01$ S (22.1%), F (1, 21) = 5.947	-	-	-	HS1	HS2
Significant variables	at $\rho < 0.01$ S (beta = 0.517)	at $\rho < 0.01$ S (beta = 0.470)				Y:2	N:2
Perceived Results	HRM5-PR Cohort mean		HRM-5-PR All Responses		HRM5-Alt-PR Cohort mean	Hypothesis	
2017 (DS-4)	1989 to 2016 (N = 28)	1991 to 2016 (N = 26)	1991 to 2016 (N = 94)	1989 to 2016 (N = 88)	1991 to 2016 (N = 26)		
Total variance	Step 1: 50.3% $\rho < 0.01$ S (50.8%) F(1, 26) = 26.872	Step 1: 58.3% $\rho < 0.01$ S (58.3%) F(1, 24) = 33.574	Step 3: 52.2% $\rho < 0.05$ S (49.8%), C (2.4%), F(2, 91) = 49.666	Step 1: 52.0% $\rho < 0.01$ S (52.0%) F(1, 86) = 93.257	Step 1: 27.2% $\rho < 0.01$ OCB (27.2%) F(1,24) = 8.966	HS1	HS2
Significant variables	at $\rho < 0.01$ S (beta = 0.721)	at $\rho < 0.01$ S (beta = - 0.764)	at $\rho < 0.05$ S (beta = 0.603) C (beta = -0.185)	at $\rho < 0.01$ S (beta = 0.713)	at $\rho < 0.01$ OCB (beta = 0.522)	Y:5	N:5

2013 (DS-3)	1989 to 2013 (N = 25)	1991 to 2013 (N = 23)	1991 to 2016 (N = 129)	1989 to 2016 (N = 125)	1991 to 2016 (N = 26)	Hypothesis	
Total Variance	Step 1: 28.1% $\rho < 0.01$ S (28.1%), $F(1, 23) = 9.008$	Step 1: 29.3% $\rho < 0.01$ S (29.3%), $F(1, 21) = 8.708$				HS1	HS2
Significant variables	at $\rho < 0.01$ S (beta = 0.531)	at $\rho < 0.01$ S (beta = 0.541)				Y:2	N:5

Strand-4 Quantitative Results

The Strand-4 baseline data analysis included statistical analysis of the responses to the two surveys (2013 survey and 2017 survey), data from mixed-mode analysis of the theme counts from the IPA analysis of two groups of IESG minutes (10% sample from 1991 to 2016 and 100% from 2015 to 2016), and theme counts from the IPA analysis of the data from the open-questions on conflict from the 2017 survey. The triangulation of the analytical data considered this baseline data as the following five datasets: (a) Dataset 1 containing the theme counts from the IPA analysis 10% of the IESG meetings from 1991 to 2016, (b) Dataset 2 containing the theme counts from the IPA analysis of 100% of the IESG meetings from 2015 and 2016, (c) Dataset 3 containing the data from the 2013 survey responses, (d) Dataset 4 with data from the 2017 survey responses, and (e) Dataset 5 from the IPA analysis of the open-ended questions from 2017 survey. Using the reduced theoretical model, the researcher analyzed these quantitative datasets at the group level (IESG cohorts) per year. Table 52 summarizes the baseline data for the Datasets 1 to 4 (surveys and IPA analysis). An IETF chair leads the IESG as a TMT for a sequence of years, so the descriptive statistics also consider the IESG groups under one IETF chair. Table 55 summarizes the IESG Cohort data per period of leadership of an IETF chair. This section contains the Strand-4 summaries by categories (sampling, descriptive statistics, reliable scales, readiness for HRM and correlation, correlation results, and HRM results).

Strand-4: Sample sizes. The surveys and the IPA analysis examined IESG behavior from the IESG inception in 1989 to 2016. Due to the number of IESG members from the IESG cohorts in 1989 and1990 who were inactive, the researcher was pleased to have responses in 2013 and 2017 surveys from 1989 to 1990. Unfortunately, the historical records for individual IESG meetings were unavailable for 1989 to 1990, but the IETF (1991, 1992) conference proceedings from 1989 to 1990 provided details on the IESG decisions and their results. Using this historical record, the researcher did a Strand-2 IPA analysis to calculate estimates for 1989 and 1990 IESG cohorts for the number of decisions and their results. Because the estimate for 1989 to 1990 for the number of IESG decisions and their outcome was based on the IETF (1991, 1992) proceedings and not officially recorded minutes, the statistical analysis separated the analysis into two groups: data that includes 1989 to 1990 (1989 to 2013 and 1989 to 2016) and data that included only IESG minutes 1991 to 2016 (1991 to 2013 and 1991 to 2016). The 2017 survey had 25 valid IESG member responses (31% of 81 active members), which included data for 94 cohort slots (30% of active member cohort slots) for 1989 to 2016 (28 years) and 88 cohort slots for 1989 to 2016 (26 years) and four valid IETF chair responses for 15 cohort slots for the years 1996-2000 and 2005 to 2016 (N = 15, years = 15). The 2013 survey had 41 valid IESG responses, including data for 129 cohort slots for 1989 to 2013 (25 years) and 125 cohort slots for 1991 to 2013 (23 years), and five IETF chair responses for the two periods (1994 to 2000 and 2005 to 2013).

The IPA analysis found 21,643 individual behaviors in 1853 decisions in Dataset 1 in the 10% analysis of the IESG (1992, 2020) minutes from 1991 to 2016 (see Tables 213 to 214) and 17,537 individual behaviors in 1605 decisions in Dataset 2 in the 100% analysis for 2015 to 2016 (see Tables 294 to 295). The researcher examined each individual's behavior for solidarity (S), conflict (C), task interdependence (TI), and OCB behaviors using the survey questions as encoding rules. In addition, the researcher encoded each decision with the type of decision (document publication, WG action, or IETF management) and the result status. The mean number of decisions per year in the 10% sample in Dataset 1 was 71.3 for the 3 IESG meetings analyzed. Dataset 2's 100% analysis discovered 820 decisions in 28 meetings during the 2015 IESG cohort year and 785 decisions in 27 meetings during the 2016 IESG cohort year. The 1,853 decisions in Dataset 1 were 51% document decisions, 27% WG actions decisions, and 22% IETF management decisions. The 100% analysis in Dataset 2 found the topics of the decisions were 58% on documents, 12% on WG actions, and 30% on IETF management decisions. The 10% found that 61% of the decisions caused a measurable result (29% document, 22% WG actions, and 9% IETF management). The 100% analysis found that 63% of the decision caused a measurable result in 2015 to 2016 (38% documents, 11% WG, 14% IETF management). The lower WG count was understandable due to the reorganization completed in 2015 of the IETF areas. Table 53 lists the Strand-2 data results per IESG cohort used in the IPA 10% and survey analyses. The IPA analysis 10% sample aligned with the 100% sample found in 2015 to 2016, where the 10% analysis estimated 780 for 2015 versus 820 decisions discovered in the 100% analysis and estimated 710 decisions for 2016 versus the 785 decisions discovered in the 100% analysis. Table 53 also provides the IPA theme counts from the 10% analysis and the 100% analysis for the IBA, the reduced theoretical model (solidarity, conflict, and TI), and the alternate model (OCB) per IESG cohort. The purpose of the IESG minutes was to record the organizational actions of the IESG, so most actions demonstrated organizational citizenship behavior. In contrast, the survey responses were the opinions of the IESG members and the IETF chair on the IESG's group behaviors per IESG cohort.

Table 51: Strand-4 – Survey and IPA Sampling

Survey		IESG members totals				IESG valid responses		
Date	Cohorts	All	Active	Cohort slots All	Cohort slots active	All received	Years	Cohort slots
2017	1989 to 2016	98	82	357	319	25	28	94
	1991 to 2016	97	81	337	299	23	26	88
2013	1989 to 2013	89	76	315	289	41	25	129
	1991 to 2013	88	75	295	269	41	23	125
Survey		IETF chairs totals				IETF chair responses		
Date	cohorts	All	Active	Cohort slots All	Cohort slots active	Valid received	Years	Cohort slots
2017	1989 to 2016	7	5	28	23	4	15	15
	1991 to 2016	7	5	26	23	4	15	15
2013	1989 to 2013	7	6	25	20	5	16	16
	1991 to 2013	7	6	23	20	5	16	16
	Study	IBA	decisions	Cohort Years	IESG Mtgs	Decisions by type: Doc	WG	I-mgt
DS1	10%	21643	1853	26	78	945	505	403
	100%	17543	1605	2	55	923	195	487
DS2	2015	8816	820	1	28	474	111	235
	2016	8721	785	1	27	449	84	252

Note. Seven IETF chairs led the IESG during 1991 to 2016, but three of these IETF chairs served as members before serving as IETF chairs.

Strand-4: Descriptive statistics. The descriptive statistics for the behaviors from the survey response and the IPA analysis of the historical records confirmed the existence of solidarity, task interdependence, and OCB found within each IESG cohort. The mean scores for the behaviors for all IESG responses for 1989 to 2016 differ 0.01 to 0.20 from the mean scores from 1991 to 2016 on a Likert 7-point scale (1 = *strongly disagree* to 7 = *strongly agree*), so only the 1989 to 2016 scores are presented in this discussion. Table 52 contains the values for both periods (1989 to 2016 and 1991 to 2016).

The 2013 survey solidarity mean score for all IESG respondents was 5.42 (*agree somewhat* to *agree solidarity exists*) with a cohort mean range of 4.95 to 6.18 (*agree somewhat* to *agree*). The 2013 survey solidarity mean score for all IETF chairs was 6.13 (*agree*), and the range was 5.70 to 6.65 (*agree* to *agree strongly*). The mean score for solidarity for all IESG 2017 survey's responses on solidarity was 5.73 (*agree that solidarity exists*), and the mean score for solidarity per IESG cohort ranged from 4.95 to 6.18 (*somewhat agree* to *agree*). The solidarity behavioral mean score for all the 2017 IETF chair respondents was 6.18 (*agree*), and the range was 5.70 to 6.65 (*agree* to *agree strongly*). The 2017 survey had a broader range than the 2013 survey for cohort means and IETF chair means. The IESG and IETF chair self-reported opinions on solidarity aligned with the historical record found in the IPA analysis, where 31% of the individual behaviors in the 10% analysis from 1991 to 2016 had solidarity, and 37% of the individual behaviors had solidarity in the 100% analysis for 2015 to 2016.

The task interdependence control variable had more variability than solidarity. The mean of all IESG responses for the 2013 and 2017 surveys was from Likert scores indicating *agree* (2013 survey TI score: 6.14 from 1989 to 2016 and 2017 survey TI score: 5.57 from 1989 to 2016), but IESG cohort mean ranges differed. The 2013 survey IESG cohort means for TI range from 5.67 to 6.67 (*agree* to *agree strongly*), and the 2017 survey IESG cohort means for TI range from 3.83 to 6.00 (*undecided* to *agree strongly*). The responses from the IETF chairs in the 2013 and 2017 survey indicated *strongly agree* that TI existed in the IESG cohorts. The IPA analyses found that 20% of the IBA had task interdependence in the10% sample (1991 to 2016), and 33-35% of the IBAs had task interdependence in the 100% sample (2015 to 2016). These results point to variance in the amount of TI per IESG cohort year.

OCB was measured concurrently as an alternate construct to solidarity. The descriptive analysis of all IESG (1992, 2019, 2020) responses for OCB had a mean of 6.27 in the 2013 survey (1989 to 2013) on a 7-point Likert scale, with a range in the IESG cohort mean value per year of 5.25 to 6.40 (*agree somewhat* to *agree*), and the 2017 survey had a mean of 6.09 (1989 to 2016), with a range in the IESG cohort mean values per year of 5.50 to 6.53 (*agree*). The mean score for OCB from all IETF chair responses from the 2013 survey was 6.44 (*agree*), and from all IETF Chair responses from the 2017 survey, the OCB mean score was 6.11. The IPA found OCB in 92% of the IBAs in 10% sample of the IESG minutes (1991 to 2016) and 92% to 94% of IBAs in the 100% sample. The high amount of OCB in the IESG (2020) minutes was due to the IESG minutes being a historical record of the actions of the IESG decision-making at required meetings.

The researcher only examined conflict in the 2017 survey and the IPA analyses (10% and 100%). The IPA analysis found conflict themes in 8% of the IBAs in the 10% sample of the IESG minutes (1991 to 2016) and 5% to 6% of the 100% sample (2015 to 2016). The conflict scale on the 2017 survey had a mean of 2.97 for all IESG responses on a 7-point Likert scale. This score means the common opinion of IESG opinion was to "somewhat disagreed conflict existed" in their IESG cohort. The group means per IESG cohort year had a range of 2.39 to 3.75 (*disagree* to *undecided*). Comparing this to the qualitative reports of conflict, it appears that conflict is socially unacceptable, so reports may be damped from public records while intra-group conflict exists. The IESG cohorts group means scores of 3.75 (*undecided*) as a group opinion probably indicates conflict existed.

Strand-4: Suitability for multivariate analysis. The behavioral data from the surveys formed reliable scales for the behaviors in the full model (HS, VS, RC, TC, and TI) and the reduced model (S, C, and TI) for all IESG responses and the IETF chairs (see Table 54) with one caveat. The IETF chairs so uniformly strongly agreed that task interdependence was present and that chairs' controlling variable was a constant. The scale reliability tests alternate construct OCB for the 2017 and 2013 surveys found the full model constructs of OCB-GC and OCB-altruism (OCB-A) were borderline reliable in the 2017 survey but unreliable for the 2013 survey. The alternate reduced model construct OCB was reliable in 2017 survey data but not in 2013 survey data. The scale reliability tests on the OCB constructs (OCB-GC, OCB-A, OCB) for IETF chairs responses to the 2013 survey and the 2017 survey found OCB and OCB-GC constructs were reliable, but OCB-a was only reliable in the 2017 survey. The PR questions (CDQ1, CDQ2, and CDQ3) were treated as a scale for the reduced model and tested for scale reliability. The PR scale was borderline reliable for the IESG responses from the 2017 survey ($N = 94$ or $N = 88$), reliable for the IESG responses from the 2013 survey ($N = 129$ or $N = 125$), reliable for the IETF chair responses from the 2017 survey ($N = 17$), and unreliable for the 2013 IETF chair responses ($N = 16$).

The Strand-4 baseline found the survey datasets did not have sufficient size to use the full model for behaviors from either all IESG respondents, all IETF chairs, or the per IESG cohort, so the reduced model was used for Strand-4 analysis (see Table 52 for the cohort size summary). The statistical analysis of the reduced theoretical model used the Strand-2 data collected by the IPA 10% analysis as an estimate for results. The Likert scale behavior data from both surveys (2013 and 2017) and the theme count behavioral data from the IPA of the 10% sample (1991 to 2016) used the Strand-2 estimate for results. The behavioral data for the 100% sample of 2015 to 2016 used the Strand-2 data collected on the 100% sample. The surveys (2013 and 2017) also used the PR behavior as an alternate dependent variable. After testing for statistical requirements for normality, homoscedasticity, linearity, and non-correlated errors, only some behavioral data were ready for correlation and HRM testing. The focus of the reduced model is the IESG group data defined by IESG cohort means per year. The IESG group data (cohort mean) could be used in the 2013 and 2017 survey to test solidarity (S), task interdependence (TI), and perceived results (PR) for the period 1991 to 2016 plus conflict (C) and OCB for the 2017 survey for the period 1991 to 2016. The data from the IPA analysis of 10% of the IESG (1992, 2020) minutes from 1991 to 2016 and the 100% of the IESG minutes 2015 to 2016 could also support correlation tests and HRM modeling for S, C, TI, and OCB.

Strand-4: Reduced model correlation and hierarchical regression modeling results. The researcher investigated the correlations between the reduced model behaviors (S, C, and TI) and results (PR and results [R]) using the Pearson product-moment correlation coefficient. Solidarity behaviors have a positive correlation with results ranging from 0.470 (2013 cohort mean at a significance of $\rho < 0.05$) to 0.845 (10% IESG sample at a significance of $\rho < 0.01$) based on the testing of variables approved for correlation (see Table 52) from the 2013 survey, the theme counts from IPA analysis of 10% of the IESG minutes 1991 to 2016, and the theme counts from 100% IPA analysis of 2015 to 2016. In addition, solidarity behaviors have a positive correlation with perceived results ranging from (0.531 to 0.782) based on the data from the 2013 survey (cohort mean and IETF chair responses), the 2017 survey (all responses, cohort mean, and IETF chair cohort mean). The perceived results behavior positively correlates with actual results for the 2013 survey ranging from 0.538 (cohort μ) to 0.583 (IETF chair cohort μ).

The researcher investigated the conflict behavior only in the 2017 survey and the IPA analyses (10% 1991 to 2016 and 100% 2015 to 2016). Solidarity and conflict had a strong negative relationship in the 2017 survey results of -0.558 (2017 survey all responses) to -0.722 (2017 cohort responses) for the IESG responses at a significance of $\rho < 0.01$. Conflict and solidarity have a strong negative relationship (-0.879 at a significance of $\rho < 0.01$) based on the IETF Chair responses to the 2017 survey. Solidarity and conflict have a strong positive relationship in behavioral data from the IPA analyses (0.716 for 10% sample 1991 to 2016, 0.686 for 100% sample 2015, 0.588 for 100% sample 2016). The strong positive relationship may be due to TC themes being the majority of themes in the conflict theme count (1991 to 2016: 1768 out of 1890 conflict themes; 2015: 413 TC themes out of 433 conflict themes, 2016: 549 TC out of 564 conflict themes). Because the IESG cohorts review and approve actions, TC is a natural part of their activities. The 2017 survey data found that conflict behavioral data has a negative correlation to perceived results (-0.396 for 2017 cohort mean data from 1991 to 2016, -0.509 for 2017 all responses from 1991 to 2016, and -0.984 for IETF chair responses), and no impact on the results. The consistent negative correlation aligned with the open-ended survey qualitative comments on the negative impact of conflict. Conflict and task interdependence (TI) did have a strong negative correlation of -0.718 for the 2017 survey IESG responses cohort mean and -0.879 for the 2017 IETF chair responses at a significance of $\rho < 0.01$. Similar to the solidarity-conflict correlation, the conflict-TI relationship had a strong positive correlation in the IPA analyses (0.727 for 10% IESG minutes from 1991 to 2016, 0.679 for 2015 IESG combo minutes, and 0.592 for IESG combo minutes at a significance of $\rho < 0.01$). The TI behavior did not correlate to results (R) in the data from the surveys, but it had a strong positive correlation to results (R) in the IESG minutes (0.738 for 10% IESG minutes, 0.798 for 2015 IESG combo minutes, and 0.632 for 2016) at a significance of $\rho < 0.01$.

The solidarity scores from the surveys predicted 22% of the variance in the results and 29% to 58% of the variance in the PR at a significance level of $\rho < 0.05$. The solidarity counts from the IESG minutes predicted between 44% to 66% of the variance in the results. The detailed results show consistency except for the 2017 survey's solidarity results. Based on the group mean per IESG cohort from the 2013 survey, the solidarity behavior predicted 22.1% of the variance in the results (R) and 29.3% of the variance in the PR at a significance level of $\rho < 0.05$. The solidarity data for the IESG cohort mean from the 2017 survey predicted none of the variances in the results but predicted 58.3% of the variance in the PR at a significance of $\rho < 0.01$. Solidarity theme counts from 10% sample (1991 to 2016) predict 62.2% of the variance in the results at a significance of $\rho < 0.01$. Similarly, the solidarity theme counts in the 100% sample predicted 64.6% of the variance of results in 2015 and 43.7% in 2016 at a significance level of $\rho < 0.01$. Controlling for TI did not affect the results in the HRM modeling tests on the reduced model for data from the surveys and the IESG minutes except for one case. Controlling for TI impacted the ability of the HRM modeling tests to detect the influence of conflict in the 10% sample of the IESG (1992, 2020) minutes from 1991 to 2016. When not controlling for TI, the solidarity and conflict theme counts from 10% sample of IESG minutes (1991 to 2016 IESG cohorts) predict 77.3% of the variance in the results (62.2% solidarity and 15.1% conflict) at a significance level of $\rho < 0.01$. Strand-4's analysis supports the reduced theoretical model Hypothesis 1 for the group behavior of solidarity predicting real and perceived results.

The 2017 survey's lack of prediction of the actual results provided an interesting quandary. The solidarity and conflict data from the 2017 survey for group means per IESG cohort or all IESG member responses correlated to the perceived results, but not the actual results. The survey data from all IETF chair responses from the 2017 survey solidarity and conflict correlated to perceived results at a significance level of $\rho < 0.01$. The mean score for solidarity for each IESG cohort from 2013 survey response data positively correlated to both the results and the perceived results.

Why did the participant's opinions fail to correlate with the results? Is this a failure in the theoretical model, or does the failure to correlate with the actual results indicate something else? Three possible reasons exist. The first possibility is that the 10% estimate is inaccurate because it is a sample. However, the 10% sample aligns with the 100% sample during 2015 to 2016. A second possibility is the IESG cohort during 2013 to 2016 might believe the perceived results because the impact of their reorganization of the IETF area in 2014 to 2016 may cause them to perceive their cohorts as successful rather than measuring their IESG cohort on the standard measures (documents published, WG actions, and IETF management). A third possibility is that the IESG cohorts may not know their IESG cohorts' actual statistics since the IETF organizational statistics do not always correlate with the effective decisions per year.

Strand-4: Alternate reduced model Correlation and hierarchical regression modeling results. The researcher included the OCB construct in the two surveys (2013 and 2017) and the IPA encoding of two samples of IESG minutes (10% sample of IESG minutes 1991 to 2016 and 100% sample of IESG minutes 2015 to 2016). The researcher's purpose for this inclusion of the OCB construct is to test concurrent validity and discriminant validity of solidarity. The only OCB behavioral data suitable for multivariate analysis was the OCB behavioral data in the 2017 survey summarized by IESG cohort means for 1991 to 2016. The correlation tests found OCB positively correlated with solidarity at 0.637

(significance $\rho < 0.01$) and TI at 0.476 (significance at $\rho < 0.05$). OCB also positively correlated to perceived results at 0.522 (significance at $\rho < 0.05$). Theme counts did not have the same variability as survey responses, so even though the IPA themes did not create a reliable OCB scale, the OCB data had normality, homoscedasticity, and linearity. Therefore, the IPA data analysis ran correlation and HRM tests on the OCB behavioral data summarized per IESG cohorts for the 10% sample (1991 to 2016) or per meeting OCB behavioral data for 2015 or 2016. OCB had strong positive correlations to solidarity, conflict, TI, and the results at a significance of $\rho < 0.01$.

The researcher ran HRM modeling tests on the data from the surveys and the IESG minutes using an alternate theoretical model in which OCB replaces solidarity. The sequences for these modeling tests were HRM-4-OCB-R, HRM-4-OCB-PR, HRM-5-OCB-R, and HMR-5-OCB (see Table 48 for details). The data from the 2017 survey IESG responses summarized per group found that the OCB behavior did not predict any variance in the results but predicted 27.2% of the variance in the perceived results. Conversely, the OCB theme counts from 10% sample of the IESG minute for 1991 to 2016 predicted 61.5% of the variance in the results. The OCB theme counts from the 100% sample of the IESG minutes predicted 71.2% (2016) to 73.1% (2015) of the variance of the results. The data indicated a difference between the OCB scores from the survey and OCB themes encoded in the IPA analysis. OCB survey results indicated a difference between solidarity and OCB based on proven instruments, so it was reasonable to conclude that solidarity has concurrent and discriminant validity for solidarity.

The OCB behavioral data from the theme counts from IPA analyses of the IESG minutes (1991 to 2016 10% sample and 2015 to 2016 100% sample) show a much higher result than the survey responses even though the categories were the same. The difference between the survey’s results and the IPA analysis could have resulted from the context of the historical record, encoding rules, errors in encoding, or any combination of these three issues. Many decisions in the IESG minutes record IESG members exhibiting OCBs so that the larger OCB theme count might have been due to the social perspective of the minutes. OCB scores on the 2017 survey predicted less of the variance in the perceived results than solidarity in the 2017 survey data. This result also suggested that OCB and solidarity were unique constructs because solidarity had concurrent and discriminant construct validity. Therefore, solidarity represented a unique construct.

Table 52: Strand-4 – Survey and Interpretive Phenomenological Analysis – Comparison of Results

Reduced Model	Dataset 3a: 2013 Survey: all IESG responses	Dataset 4a: 2017 Survey: all IESG responses	Dataset 3b: 2013 Survey: Cohort Mean (group)	Dataset 4b: 2017 Survey: Cohort Mean (group)	Dataset 3c: 2013 Survey: all IETF chair responses	Dataset 4c: 2017 Survey: All IETF chair responses	Dataset 1: 10% Sample 1991 to 2016 (per year)	Dataset 2 100% Sample 2015 to 2016 (per year)
Cohort slots	1989 to 2013: 129 1991 to 2013: 125	1989 to 2016: 94 1991 to 2016: 88	1989 to 2016: 25 1991 to 2016: 23	1989 to 2016: 28 1991 to 2016: 26	1994-2013: 15	1994 to 2016: 16	Decisions: 10%: 71.3 100%: 713	Decisions: 2015: 820 2016: 785
Total cohort slots/years	1989 to 2013: 315 1991 to 2013: 289	1989 to 2016: 357 1991 to 2016: 337	1989 to 2016: 25 1991 to 2016: 23	1989 to 2016: 28 1991 to 2016: 26	1989 to 2013: 25 1991 to 2013: 23	1989 to 2016: 28 1991 to 2016: 26	1991 to 2016: 26	2015: 1 2016: 1
Behaviors [IPA:IBA]	All responses mean [meaning]	All responses mean, % of max	Cohort mean range (CMR)	Cohort mean range (CMR)	All years mean Range (R):	All years mean Range	10%: 832.4 *1 100%: 8324	2015: 8816 2016: 8721
solidarity	89-13: 5.73 91-13: 5.74 [agree]	89-16: 5.42 91-16: 5.43 [agree somewhat]	89-13: 4.95-6.18 91-13: 4.95-6.18 [agree somewhat to agree]	89-16: 4.30-6.10 91-16: 4.30-6.10 [undecided to agree somewhat]	94-13: 6.13 R: 5.70-6.65 [agree to strongly agree]	94-16: 6.18 Range: 5.7-6.9 [agree to strongly agree	10%: 257.9 100%: 2579 (31%)	2015:3244 2016: 3249 (37%)
Conflict	n/a	89-16: 2.97 91-16: 2.96 [somewhat disagree]	n/a	89-16: 2.39-3.75 91-16: 2.42-3.75 [disagree to undecided]	n/a	94-16: 2.93, Range: 1.5-3.5 [strongly disagree to somewhat disagree]	10%: 72.7 100%: 727 9%	2015: 433 2016: 564 (5, 6%)
TI	89-13: 6.14 91-13: 6.13 [agree]	89-16: 5.57 91-16: 5.58 [agree]	89-13: 5.67-6.67 91-13: 5.67-6.67 [agree to strongly agree]	89-16: 3.83-6.00 91-16: 3.83-6.00 [undecided to agree]	94-13: 6.88 R: 6.67-7.00 [strongly agree]	94-16: 7.0 [strongly agree]	10%: 295.8 100%: 2958 36%	2015: 2926 2016: 3075 (33%, 35%)
OCB	89-13: 6.27 91-13: 6.26 [agree]	89-16: 6.09 91-16: 6.11 [agree]	89-13: 5.25-6.40 91-13: 5.25-6.40	89-16: 5.50-6.53 91-16: 5.50-6.53	94-13: 6.44 R: 5.50-7.00	94-16: 6.11 Range: 5.3-7.0	10%: 773.8 100%: 7738 93%	2015:8096 2016:8195 (92%, 94%)

Reduced Model	Dataset 3a: 2013 Survey: all IESG responses	Dataset 4a: 2017 Survey: all IESG responses	Dataset 3b: 2013 Survey: Cohort Mean (group)	Dataset 4b: 2017 Survey: Cohort Mean (group)	Dataset 3c: 2013 Survey: all IETF chair responses	Dataset 4c: 2017 Survey: All IETF chair responses	Dataset 1: 10% Sample 1991 to 2016 (per year)	Dataset 2 100% Sample 2015 to 2016 (per year)
Perceived Results (PR)	89-13: 5.13 91-13: 5.13 [somewhat agree]	89-16: 5.31 91-16: 5.34 [somewhat agree]	[agree somewhat to agree] 89-13: 4.17-5.57 91-13: 4.17-5.57 Undecided to agree	[agree somewhat to agree] 89-16: 4.17-6.33 91-16: 4.17-6.33 undecided to agree strongly	[agree to strongly agree] 94-13: 5.60 Range: 4.7-7.0 [somewhat to strongly agree]	[agree somewhat to strongly agree] 94-16: 6.24 Range: 6.0-7.0 [agree to strongly agree]	n/a	n/a
Reliable scales	S, TI, PR HS, VS	S, C, TI, OCB HS, VS, RC, TC PR	S, TI, PR HS, VS	S, C, TI, OCB, HS, VS, RC, TC PR	S, OCB TI constant	S, C, OCB, PR TI constant	Per meeting S, C, TI, OCB DI Qualitative: All	Per meeting 2015: TI 2016: TI Qualitative: all
HRM Ready	Reduced Model for Survey						IPA Reduced Model	
1989 to 2016	None	S, C, PR	S, TI, PR	S, C, PR	1994 to 2016 only	1994 to 2016 only	Per year only	Per mtg/year 2015-combo
1991 to 2016	None	S, C, PR	S, TI, PR	S, C, OCB, TI, PR	S, PR	S, C, PR	S, C, TI, OCB*2, DI	TI S, C, OCB*2
Correlation	Reduced Model for Survey						IPA Reduced Model	
1989 to 2016	None	S-C** (-.627)	None	S-C** (-.627) S-TI** (.481)	n/a	n/a	n/a	n/a
1991 to 2016	None	S-C** (-.558)	None	S-C** (-.722) S-TI** (.830) S-OCB** (.637) C-TI** (-.718) TI-OCB* (.476)	1994-2013: none	1994 to 2016: S-C**(-.879)	S-C** (.716) S-TI** (.938) C-TI** (.727) S-OCB** (.902) C-OCB** (.728)	S-C** (.616) S-TI** (.991) C-TI** (.679) S-OCB** (.919) C-OCB** (.692)

Reduced Model	Dataset 3a: 2013 Survey: all IESG responses	Dataset 4a: 2017 Survey: all IESG responses	Dataset 3b: 2013 Survey: Cohort Mean (group)	Dataset 4b: 2017 Survey: Cohort Mean (group)	Dataset 3c: 2013 Survey: all IETF chair responses	Dataset 4c: 2017 Survey: All IETF chair responses	Dataset 1: 10% Sample 1991 to 2016 (per year)	Dataset 2 100% Sample 2015 to 2016 (per year)
							TI-OCB** (.938)	TI-OCB** (.918)
Correlation to Results	2013 all IESG responses	2017 all IESG responses	2013 Cohort mean	2017 Cohort mean	2013 IETF chair mean	2017 IETF chair mean	Dataset 1 per year	Dataset 2 per mtg/year
1989 to 2016	None	None	S**(.517) PR*(.451)	None	1994 to 2016	1994 to 2016	n/a	n/a
1991 to 2016	None	None	S* (.470) PR** (.538)	None	PR* (.583)	None	S** (.845)	S** (.804) TI** (.798)
1991 to 2016 OCB	-	-	-	-	-	-	C* (.409) TI** (.738) OCB** (.784)	C** (.545) OCB** (.855)
Correlation to PR	2013 all IESG responses	2017 all IESG responses	2013 Cohort mean	2017 Cohort mean	2013 IETF chair mean	2017 IETF chair mean	Dataset 1 per year	Dataset 2 per mtg/year
1989 to 2016	None	S** (.706) C** (-.521)	S** (.531)	S** (.713) C** (-.479)	1994 to 2016	1994 to 2016	n/a	n/a
1991 to 2016	None	S** (.721) C** (-.509)	S** (.541)	S** (.764) C** (-.396) TI** (.625) OCB* (.522)	S** (.790)	S** (.782) C**(-.984) (TI constant)	n/a	n/a
HRM model	2013 all IESG responses	2017 all IESG responses	2013 Cohort mean	2017 Cohort mean	2013 IETF chair mean	2017 IETF chair mean	Dataset 1 per year	Dataset 2 per mtg/year
HRM-4-R predicts % of variance of results	None	None**	89-13: S** predicts 26.7% 91-13:S* predicts 22.1%	None**	n/a	n/a	91-16: S** predicts 62.2-73.6% of results	S** predicts 2015: 64.6% 2016: 43.7% of results
HRM-4-PR predicts %	None	89-16: S** predicts 49.8%	89-13:S* predicts 28.1%	89-16: S** predicts 50.8%	n/a	n/a	n/a	n/a

Reduced Model	Dataset 3a: 2013 Survey: all IESG responses	Dataset 4a: 2017 Survey: all IESG responses	Dataset 3b: 2013 Survey: Cohort Mean (group)	Dataset 4b: 2017 Survey: Cohort Mean (group)	Dataset 3c: 2013 Survey: all IETF chair responses	Dataset 4c: 2017 Survey: All IETF chair responses	Dataset 1: 10% Sample 1991 to 2016 (per year)	Dataset 2 100% Sample 2015 to 2016 (per year)
of variance of PR		91-16: S^{**} predicts 52.0%	91-13:S^{**} predicts 29.3%	91-16: S** predicts 58.3%				
HRM-5-R (no TI) predicts % of variance of results	None	None	89-13: S^{**} predicts 26.7% 91-13:S^{*} predicts 22.1%	None	n/a	n/a	91-16: S** (62.2%) C** (15.1%) for 77.3% of results	S^{**} predicts 2015: 64.6% 2016: 43.7% of results
HRM-5-PR (no TI) predicts % of variance of results	None	89-16: S^{*}, C^{*} predicts 52.2% [S*(49.8%), C* (2.4%)] 91-16: S^{**} predicts 52.0%	89-13:S^{*} predicts 28.1% 91-13:S^{**} predicts 29.3%	89-16: S^{**} predicts 50.3% 91-16: S** predicts 58.3%	n/a	n/a	n/a	n/a
Alternate Model Hypothesis supported	Not valid	Hypothesis 1 for perceived results	Hypothesis 1 for actual and perceived results	Hypothesis 1 for actual and perceived results	None	None	Hypothesis 1 for actual results	Hypothesis 1 for actual results
HRM-4-OCB-R predicts % of variance of results (R)	Not valid	Not valid	Not valid	Not valid	not valid	not valid	91-16: OCB^{**} predicts 61.5% of results	2015: OCB^{**} predicts 73.1% results 2016: OCB^{*}, C* predict 76.3% of results with OCB 71.2%

Reduced Model	Dataset 3a: 2013 Survey: all IESG responses	Dataset 4a: 2017 Survey: all IESG responses	Dataset 3b: 2013 Survey: Cohort Mean (group)	Dataset 4b: 2017 Survey: Cohort Mean (group)	Dataset 3c: 2013 Survey: all IETF chair responses	Dataset 4c: 2017 Survey: All IETF chair responses	Dataset 1: 10% Sample 1991 to 2016 (per year)	Dataset 2 100% Sample 2015 to 2016 (per year)
HRM-4-OCB-PR predicts % of variance of perceived results (PR)	Not valid	Not valid	Not valid	91-16: OCB** predicts 27.2% of results	not valid	not valid	n/a	n/a
HRM-5-OCB-R (no TI) predicts % of variance of results (R)	Not valid	Not valid	Not valid	89-16: not valid 91-16: none	Not valid	Not valid	91-16: OCB** + C** predicts 77.8% of results with OCB 61.5%, C 16.4%	OCB** predicts 2015: 73.1% 2016: 71.2%
HRM-5-OCB-PR predicts % of variance of perceived results (PR)	Not valid	Not valid	Not valid	89-16: not valid 91-16: OCB** Predicts 27.2% of PR	Not valid	Not valid	n/a	n/a

IBA = individual behavior actions, PR = perceived results, DI = Discovered-IESG themes, *= $\rho < 0.05$ **= $\rho < 0.01$

*1 – total IBA: 21,643 from formal 1991 to 2016.

*2 – variables split between scales with reliability based on theme counts and reliability based on qualitative measures

Table 53: Strand-4 – Descriptive Data per Cohort Year

Results 10% IPA theme counts				
Type	Sum	% type	mean mtg	mean year
All	1853	100%	23.7	71.3
Docs	945	51%	12.1	35.4
WG	505	27%	6.5	19.4
I-mgt	403	22%	5.2	15.5
Result	1130	61%	14.5	43.5

Result Components				
Docs	539	29%	6.9	20.7
WG	416	22%	5.3	16
I-mgt	175	9%	2.2	6.7

2017 Survey results		
Q#	Topic (*N* = *94*)	Mean
1	Documents	5.59
1a	Standard Documents	5.41
1b	Non-Std. Documents	5.40
2	WG actions	5.06
3	IETF management	5.38
4	IETF Leader	5.48

Strand-2 decisions and results			Dataset 1: 10% Sample Theme counts (10%)					Dataset 4: 2017 Survey Cohort mean (Likert-7)					Dataset 3: 2013 Survey Cohort Mean (Likert-7)			
Year	Acts *1	Res *1	IBA	S	C	TI	OCB	PR	S	C	TI	OCB	PR	S	TI	OCB
1989	200	107 54%	-	-	-	-	-	5.89	5.37	2.39	5.33	5.73	5.00	5.43	6.59	6.00
1990	260	135 52%	-	-	-	-	-	5.89	5.37	2.39	5.33	5.73	5.00	5.43	6.50	6.00
1991	320	180 56%	394	123 31%	47 12%	120 30%	394 83%	5.33	5.24	2.92	5.50	5.65	4.60	5.30	6.07	5.52
1992	680	420 62%	779	240 31%	58 7%	239 31%	622 80%	5.33	5.24	2.92	5.5	5.65	4.58	5.20	5.83	5.25
1993	490	290 59%	578	190 33%	84 15%	183 32%	480 83%	5.00	5.25	3.17	5.47	5.72	5.11	5.95	6.67	6.13
1994	470	340 72%	563	186 33%	46 8%	180 32%	526 89%	5.00	5.24	3.33	5.67	5.90	5.00	5.35	6.67	6.20
1995	490	390 80%	545	183 34%	30 6%	177 32%	486 89%	4.83	5.78	3.04	5.33	5.75	4.83	5.51	6.58	6.05
1996	610	540 89%	725	229 32%	51 7%	227 31%	715 99%	5.17	5.16	3.25	5.42	5.90	5.11	5.73	6.67	6.20
1997	400	340 85%	471	162 34%	21 4%	160 29%	437 91%	5.67	6.07	2.94	6.00	6.53	4.78	4.95	6.44	6.00
1998	460	360 78%	592	177 33%	22 5%	174 33%	538 93%	5.50	5.83	2.42	6.00	6.40	5.17	6.08	6.33	6.40
1999	500	480 96%	703	231 33%	34 5%	231 33%	654 93%	5.50	5.83	2.42	6.00	6.40	4.93	6.01	6.73	6.40
2000	330	270 82%	452	135 30%	34 8%	134 30%	411 91%	5.11	5.07	3.17	4.89	5.93	4.44	5.69	6.17	6.23
2001	480	330 69%	672	154 23%	38 6%	275 41%	667 99%	5.17	4.85	3.75	4.33	5.90	4.33	5.66	5.87	6.00
2002	690	330 48%	727	154 27%	38 10%	275 33%	701 96%	4.78	4.92	3.17	4.56	5.93	4.33	5.66	5.87	6.00
2003	1030	520 50%	1013	299 30%	77 8%	361 36%	932 92%	4.17	4.30	3.33	3.83	5.50	4.17	5.54	5.83	6.40
2004	870	520 60%	1003	296 30%	88 9%	389 39%	1003 100%	5.17	4.75	3.58	4.33	5.70	5.50	6.18	6.33	6.40
2005	970	600 62%	1163	296 29%	91 9%	360 31%	1163 100%	6.00	5.95	2.50	6.00	6.40	5.50	6.18	6.33	5.93
2006	760	350 46%	954	272 29%	89 9%	339 36%	878 92%	5.56	5.90	2.94	6.00	6.53	5.33	5.48	6.39	6.00

Results			
2017 Survey		IPA Datasets	
All (A)	*N = 94*	DS1 themes	DS2 themes
Year (Y)	*N = 28*		
1989 to 2016 Behavior Means		Year	2015 (%IBA)
		Est.	2016 (%IBA)
Solidarity (S)	A: 5.42	2579 (31%)	3244 (37%)
	Y: 5.43		3249 (37%)
Conflict (C)	A: 2.97	727 (9%)	433 (5%)
	Y: 2.96		564 (6%)
TI	A: 5.57	2959 (36%)	2926 (33%)
	Y: 5.50		3075 (35%)
OCB	A: 6.09	7738	8096
	Y: 6.02	(93%)	8195
IBA		8324	8821
			8721

Strand-2 decisions				Dataset 1: 10% Sample				Dataset 4: 2017 Survey					Dataset 3: 2013 Survey			
Year	Act[*1]	Res[*1]	IBA	S	C	TI	OCB	PR	S	C	TI	OCB	PR	S	TI	OCB
2007	1060	420 40%	1177	392 33%	182 15%	417 35%	1078 92%	5.33	6.10	2.27	6.00	6.33	5.46	5.74	6.38	6.00
2008	840	370 44%	1114	256 23%	141 13%	345 31%	1010 91%	5.42	6.1	2.72	6.00	6.33	5.46	5.74	6.38	6.00
2009	1130	640 57%	1221	358 29%	90 7%	455 37%	1125 92%	5.41	6.13	2.75	5.92	6.35	5.57	5.86	6.37	6.22
2010	1030	520 50%	1134	380 34%	108 10%	461 41%	1109 98%	5.92	5.98	2.79	5.92	6.35	5.47	5.88	5.70	6.00
2011	869	510 59%	1099	391 36%	120 11%	474 43%	1044 95%	5.93	5.82	2.87	5.87	6.28	5.36	5.90	5.67	5.93
2012	830	500 60%	954	318 33%	133 14%	387 41%	877 92%	5.83	5.80	2.86	6.00	6.37	5.54	6.00	5.82	6.18
2013	1050	650 62%	1117	388 35%	83 7%	495 44%	1059 95%	4.87	5.00	3.10	5.73	6.24	5.44	5.73	5.89	6.07
2014	600	390 65%	744	253 34%	49 7%	303 41%	647 87%	4.87	5.00	3.10	5.73	6.24				
2015	780	560 72%	970	301 31%	47 5%	277 29%	873 90%	5.22	5.28	3.11	5.89	6.40				
2016	710	480 68%	837	297 35%	55 7%	289 35%	779 93%	6.67	5.60	3.42	5.67	6.40				

Period Mean				10% Theme mean				2017 Survey cohort means					2013 Cohort mean			
Period	Act	Res	IBA	S	C	TI	OCB	PR	S	C	TI	OCB	PR	S	TI	OCB
89-16								5.39	5.42	2.97	5.57	6.09	6.13	5.73	6.14	6.04
91-16	713	435	8324	258	73	296	774	5.31	5.43	3.01	5.58	6.11	5.13	5.74	6.13	6.05
%		61%		36%	9%	36%	93%									
2015	820	526	8821													
%		64%														
2016	785	478	8721													
%		61%														

Table 54: Strand-4 – Scale Reliability in Datasets 1 to 4 Compared With Published Scale Reliability

Cronbach Alpha on Standardized items												Cronbach Alpha on Standardized items					
Theme Counts from IPA analysis												Likert-7 Scales					
DS1: 10% Formal per year ('91-'16)						DS2: 100% Sample 2015 per Meeting						DS4: 2017 Survey (1991 to 2016)					
N	S	C	TI	OCB	DI	*N*	S	C	TI	OCB	DI	*N*	S	C	TI	OCB	PR
26	0.378	0.150	0.174	0.215	0.941	28	0.739	0.463	0.779	0.573	0.606	88	0.919	0.878	0.898	0.797	0.765
DS1: 10% Formal per meeting (1991 to 2016)						DS2: 100% Sample 2016 per Meeting						DS3: 2013 Survey (1991 to 2013)					
N	S	C	TI	OCB	DI	*N*	S	C	TI	OCB	DI	*N*	S	C	TI	OCB	PR
94	0.916	0.884	0.896	0.800	0.765	27	0.629	0.578	0.802	0.608	0.650	125	0.908	n/a	0.862	0.685	0.869
Cronbach Alphas from Published Research on Solidarity, C, TI, and OCB												DS4: 2017 Survey (1989 to 2016)					
Koster and Sanders (2006) per individual*1						Sanders and Schyns (2006a) per individual*2						*N*	S	C	TI	OCB	PR
N	HS	VS	OCBgc	OCBa		*N*	HS	VS	TI	Teams		94	0.916	0.884	0.896	0.800	0.765
674	to: 0.85	to: 0.78	0.70	0.70		193	0.77	0.85	0.77	35							
						Pearson, Ensley, & Amason (2002)*3						DS3: 2013 Survey ('89-'13					
674	From:	From:				*N*	TC	RC	Firms			*N*	S	C	TI	OCB	PR
	0.92	0.89				148	0.72	0.86	48			129	0.908	n/a	0.858	0.669	0.853

*1 – Cronbach alpha with no indication of the number of teams or status (Koster & Sanders, 2006, p. 528).

*2 – Cronbach alpha with no indication of the type of teams (Sanders & Schyns, 2006a, pp. 543–544).

*3 – Cronbach alpha with an indication that 148 people were top managers of publicly traded firms (Pearson et al., 2002, p. 122).

Strand-4: Internet Engineering Steering Groups under an Internet Engineering Task Force chair's leadership. The theoretical model suggests that the IETF chair influences the solidarity of the IESG as a TMT by vertical solidarity or support of HS between members. Solidarity influences group effective consensus decision-making as indicated by increased output. One questioned the following: Can the influence of the IETF chair as the leader be observed in the changes in the behaviors of solidarity (S), conflict (C), TI, OCB, and in the results (perceived and actual)?

The Strand-4 analysis examined this question by considering the five datasets created from the IPA analysis and the surveys per IETF chair to determine if the IETF chair's leadership explains some of the variances. Table 55 contains the mean behavioral scores for the reduced theoretical model (S, C, TI, OCB), the results, and perceived result scores during the IESG cohorts led by each IETF chair, and Table 56 contains similar scores for the full theoretical model. In addition, Table 55 provides the perceived effectiveness of the IETF chair and the leader ratio from the IPA analysis. The leader ratio is the ratio of two-person debates to group debates recorded in the IESG minutes.

Due to the lack of IESG minutes for 1989 to 1990, the researcher split entries for the first IETF chair into two periods: 1989 to 1990 and 1991 to 1993. Due to the limited number of IETF chair periods (seven IETF chairs), statistical analysis beyond the descriptive statistics reaches into flights of assumptions ungrounded by facts. Strand-4 analysis for the leadership of the IETF chair as a leader of IESG teams turns to observations based on the plots of behaviors, results (R), PR, and perceived influence of the chairs shown in Figure 30, Figure 31, and Figure 32. The researchers observed from these data grouped per IETF chair that (a) actual results did not align with PR or the perceived leader effectiveness, (b) PR and perceived leadership effectiveness aligned with solidarity, (c) conflict increases differ between the survey 2017 and the IPA analysis, (d) OCB patterns differed from solidarity patterns. This section discusses these four observations.

Determining the results for an IESG cohort was difficult because the IETF organizational results for documents published, WG actions, and IETF management per year were indirectly related to the IESG cohort's successful decision-making for that year (see Chapter 4 for the discussion). The estimate for the IESG cohort was calculated based on an IPA of 10% of the IESG minutes, which counted decisions and decisions with results. There is a potential that the 10% sample does not represent the whole year. The 100% sample of 2015 to 2016 aligned with the results predicted, but only the 2013 survey found a positive correlation between the perceived results and the actual results. The 2017 survey did not find any correlation between the perceived result scores per IESG and the actual results.

The mixed results between the two surveys and the two IPA analyses present a contradictory picture until one examines the IESG Cohort responses within the period of an IETF chair provided in Table 55. The IESG group scores (mean IESG cohort responses) averaged per IETF chair show that perceived results (PR), IETF chairs effectiveness (CQ4) scores align with mean solidarity scores and the IPA analysis of solidarity. Figure 30 shows four graphs that compare mean solidarity scores (2013 survey, 2017 survey, and 10% IPA), perceived results (PR) scores, % decisions with results scores, and score for an IETF chair impact the results of an A IESG cohort's decisions. Graph 1 within Figure 30 shows the mean scores for IESG cohorts under a single IETF chair for solidarity (S-2017), perceived results (2017-PR), and the IETF chair effectiveness (2017-CQ4) plotted versus the terms of the seven IETF chairs. Graph 1 also shows solidarity from the 10% IPA analysis (Dataset 1) plotted as % of IBA times 10 to provide a clear comparison with the survey scores. Graph 1 in Figure 30 clearly shows the alignment between the solidarity scores, the perceived results, and the effective leadership scores. Graph 2 in Figure 30 adds to graph 1 a plot of the results (represented as a % of decisions with results multiplied by 10; %Rx10) and the mean conflict scores for the IESG Cohorts from the 2017 survey conflict scale and the open-ended conflict IPA analysis. Although lower conflict may influence some of the changes in results, the real results during 1995 to 2016 are abnormally high in this diagram. Graphs 3 and 4 in Figure 30 show the same information for the 2013 survey with one exception. Because the survey in 2013 did not query for conflict, the researcher includes the mean scores on conflict from the 2017 survey in Graph 4. The HS and VS mean scores per IETF chair follow the patterns of the combined solidarity mean scores. Figure 31 in Graphs 3 and 4 shows the mean of IESG cohort scores per IETF chair period for HS and VS. Notice that solidarity data from the 2017 survey has more volatility (due to fewer participants) in these graphs than the 2013 survey or the IPA analysis.

During IETF Chair 3's tenure (1995 to 2000), 87% of the consensus decisions in the IESG cohorts had results. This rate of successful decisions could be due to a skew in IESG minutes in the 10% sample or actual historical events. During 1995 to 2000, there was tremendous growth in Internet technology firms, such as Intel, Cisco, and Microsoft. This period of growth ended with the "dot-com" downturn in the stock market on March 11, 2000 (Randewich & Krauskopf, 2020). In addition, the external pressure during this period to publish IETF standards documents and establish new work within the IETF was high. This higher percentage of successful decisions signals is a potential risk for the findings of this research. Future research should investigate Strand-2 statistics for 1995 to 2000 by doing a 100% analysis to determine the actual number of decisions with results.

Considering conflict from the vantage point of the IESG Cohort scores per IETF chair indicated differences between historical records, the survey questions, and the open-ended questions that aligned with the correlations discovered in the per IESG cohort analysis between conflict and results. Figure 31 plots the mean of the IESG cohort scores per IETF chair for solidarity and conflict from the 10% IPA analysis (Dataset 1), and survey responses (2013 and 2017 survey (datasets 3 and 4)) and the open-ended question IPA analysis (dataset 5). Graph 1 of Figure 31 shows solidarity, conflict, and results as a % of decisions. Graph 2 plots what percentage of the total theme counts for IBA that the theme counts for each behavior versus the percentage of decisions that resulted in measurable results. Graph 2 illustrates that the IESG minutes record more conflicts in 2001 to 2004 and 2007 to 2012, while fewer conflicts existed in 1996 to 2000. The conflict in the historical records is composed chiefly of TC. TC is an accepted part of the IESG group's process as reviewers of new technology, but RC is "socially unacceptable."

The historical record in the IESG minutes records few RCs, but the survey responses do include RCs. The 2017 survey responses have higher group means for conflict scores in 2001 to 2004 IESG cohorts than other periods, and lower group means for conflict in 1989 to 1990 IESG cohorts. The 2017 survey's open question indicated some internal conflict in 2001 to 2004 caused by the IESG handling of the interaction of the IETF and the ISO standards organization. The 2017 open-ended question indicated "underground conflict" that burst into fractions in the IETF during 1996 to 2000 and 2013 to 2017 caused by IESG members attempting to take on the IETF chair's role. During 1996 to 2000, the IETF's expansive growth caused it to interact with existing SDOs to determine its place within the constellation of SDOs.

During 2013 to 2017, the IANA (2020) transition caused the IETF chair and IAB chair to be drawn away to exterior focus. These periods seem to have increased the conflict within the IESG as the IETF chair's attention was distracted from the IESG by exterior events. Tracing conflict by IESG cohorts under a single IETF chair shows the dynamics of expression of conflicts in a political environment of the IETF as a change organization for IT. The Strand-4 analysis of the IESG cohort descriptive information per IETF chairs view on conflicts suggests that the 2017 survey results, which indicate a medium negative correlation (-0.479 at a significance of $\rho < 0.01$) between conflict and the perceived results, are valid as RC and TC is represented in the survey analysis. The positive medium correlation between conflict and the results (0.409 at $\rho < 0.01$ for 10% IPA 1991 to 2016 and 0.545 for the 100% IPA in 2015) is also valid since the conflict recorded in the historical record is primarily TC.

The correlation and HRM modeling results regarding the survey indicate that when OCB was measured concurrently with solidarity, the OCB construct operates differently from the solidarity behavior construct. The mean IESG cohorts per IETF chair confirms s observation. Figure 32 plots the mean IESG cohort scores per IETF Chair from the surveys and the 10% IPA analysis for OCB and solidarity. Graph 1 in Figure 32 shows the mean scores for the IESG Cohorts per IETF chair's tenure for OCB behaviors from the 2013 survey (OCB-2013) and the

2017 survey (OCB-2017). The 10% IPA analysis plots the percentage of OCB's total individual actions multiplied by 10. This multiplication allows one to consider the difference between the IESG member's self-reporting scores and the historical record. The historical record in the IESG minutes reports on actions taken by IESG, and most of these actions exhibit some form of organizational citizenship behavior. Only when an IESG member fails to perform expected duties is there a lack of OCB behaviors. Graph 2 in Figure 32 plots the average IESG cohort means per IETF chair for the full theoretical model's OCB behaviors (OCB-GC and OCB-A). This plot shows variability between the OCB-related scores on the two surveys but similar plots. Graph 3 of Figure 32 plots solidarity and OCB mean cohort scores for all IESG cohorts under each IETF chair from survey data. The researcher also plots in Graph 3 the percentage of IBAs that the solidarity and OCB behaviors exist in the 10% sample of IESG minutes per IETF chair's tenure. These percentages are multiplied by 10 to provide similar scales to the 7-point Likert scale results. Graph 3 aligns with the correlation and HRM results from the IESG cohorts.

Figure 30: Strand-4 – Solidarity and conflict versus results and IETF chair effectiveness.

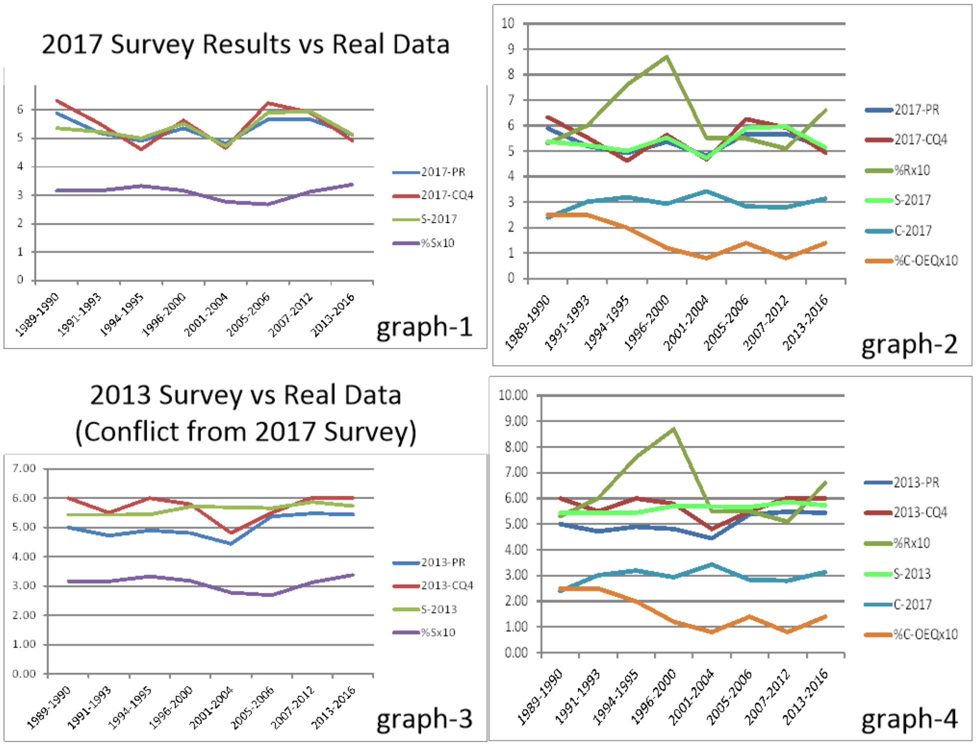


Figure 31: Strand-4 – Solidarity and conflict per IETF chair period.

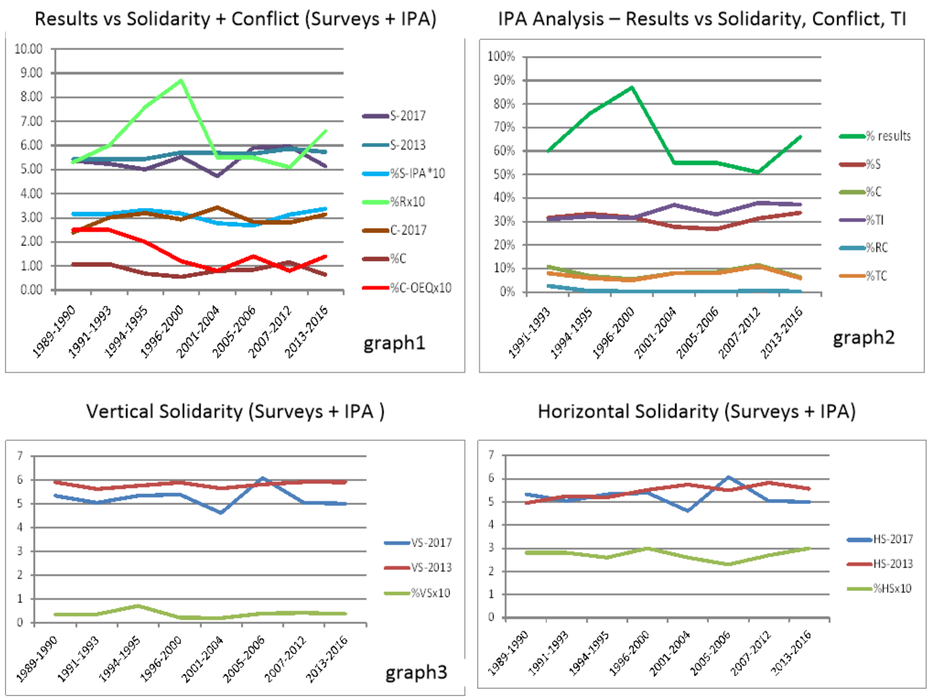
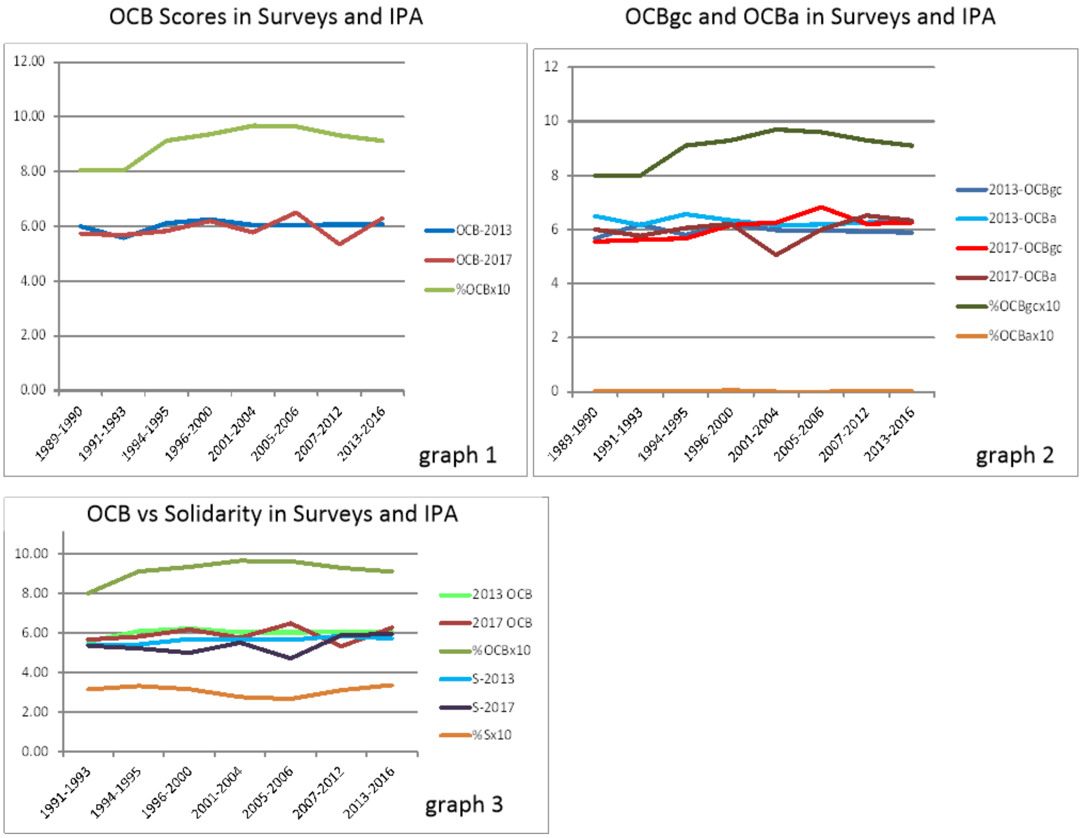


Figure 32: Strand-4 – Organizational citizenship behaviors versus solidarity.



Detailed Legend for Figures 157-159

Mean scores for IESG Cohorts during each IETF chairs time period

- 2017-PR –mean perceived results score from 2017 survey
- 2013-PR – mean perceived results score from 2013 survey
- 2017-CQ4 – mean IETF leader effectiveness score from 2017 survey
- 2013-CQ4 – mean IETF leader effectiveness score from 2013 survey
- S-2017 – mean solidarity score from 2017 survey
- S-2013 – mean solidarity score from 2013 survey
- HS-2017 – mean horizontal solidarity score from 2017 survey
- HS-2013 – mean horizontal solidarity score from 2013 survey
- VS-2017 – mean vertical solidarity score from 2017 survey
- VS-2013 – mean vertical solidarity score from 2013 survey
- C-2017 – mean conflict score from 2017 survey
- C-2013 – mean conflict score from 2013 survey
- OCB-2017 – mean OCB score from 2017 survey
- OCB-2013 - mean OCB score from 20123 survey

Strand-2 Results from IPA 10% Analysis per IETF chair time period

- % results – decisions with results as a percentage of the total decisions
- %Rx10 – % results multiplied by 10 for better display

Theme counts from IPA analysis of 10% of IESG Minutes for time period of IETF chair given as percentages of total individual behavior actions (IBA)

- %S – solidarity theme counts as a percentage of total IBA for the period
- %VS – vertical solidarity theme counts as a percentage of total IBA for the period
- %HS - horizontal solidarity theme counts as a percentage of total IBA for the period
- %TI – task interdependence theme counts as a percentage of total IBA for the period
- %C – Conflict theme counts as a percentage of total IBA for the period
- %OCB –OCB theme counts as a percentage of total IBA for the period
- %OCBgc – OCB-GC theme counts as a percentage of total IBA for the period
- %OCBa– OCB-GC theme counts as a percentage of total IBA for the period
- %TC – task conflict theme counts as a percentage of total IBA for the period
- %RC – relationship conflict theme counts as percentage of total IBA for the period
- %S-IPAx10 - %S times 10 (for clarity in display)
- %OCBx10 - %OCB times 10 (for clarity in display)
- %OCBgcx10 - %OCBgc times 10 (for clarity in display)
- %OCBax10 - %OCBa times 10 (for cliarity in display)

Theme counts from IPA analysis of 2017 survey open questions on conflict given as percentages of total individual behavior actions (IBA) per time period

- %C-OEQx10 - % of total Conflict themes in 2017 Open-Ended Questions per IETF chair period multiplied by 10

Table 55: Strand-4 – Reduced Model Behavior Totals per Internet Engineering Task Force Chair Period

Period Totals				Dataset 4: 2017 Survey Means per Period							Dataset 5: 2017 Survey Open-Ended questions theme count				Dataset 3: 2013 Survey Means Per Period					
Period	Act	Res	%	*N*	S	C	TI	OCB	PR	Ldr CQ4	OEQ RC	OEQ TC	OEQ C	% total themes	*N*	S	TI	OCB	PR	Ldr CQ4
89-90	460	242	53	6	5.37	2.39	5.33	5.73	5.89	6.33	-	-	-	-	4	5.43	6.50	6.00	5.00	6.00
91-93	1490	890	60	13	5.24	3.01	5.49	5.68	5.21	5.54	4	9	13	25%	12	5.43	6.14	5.58	4.72	5.50
94-95	960	730	76	9	5.01	3.19	5.50	5.83	4.92	4.62	3	7	10	20%	7	5.44	6.62	6.11	4.90	6.00
96-00	2300	1990	87	14	5.53	2.93	5.60	6.19	5.36	5.64	2	4	6	12%	19	5.71	6.46	6.25	4.81	5.79
01-04	3070	1700	55	9	4.73	3.43	4.30	5.78	4.82	4.67	1	3	4	8%	16	5.69	5.92	6.05	4.44	4.81
05-06	1730	950	55	4	5.91	2.83	6.00	6.50	5.67	6.25	3	4	7	14%	8	5.66	6.38	6.05	5.38	5.50
07-12	5759	2960	51	25	5.96	2.80	5.95	6.34	5.68	5.92	3	1	4	8%	60	5.87	6.01	6.06	5.48	6.02
13-16	3140	2080	66	15	5.14	3.14	5.76	6.29	5.13	4.93	3	4	7	14%	3	5.73	5.89	6.07	5.44	6.00

IESG Cohort Means averaged per Period				Dataset 1: IPA Analysis of 10% of IESG Minutes (1991 to 2016)											Dataset 2: IPA Analysis 2015 to 2016					
				Theme counts					% of IBA				Leader Ratio		2015			2016		
Period	Act	Res	%	IBA	S	C	TI	OCB	%S	%C	%TI	%OCB	Raw	adjust	%S	%C	%TI	%S	%C	%TI
89-90	230	121	53	-	-	-	-	-	-	-	-	-	-							
91-93	497	297	60	1751	553	189	542	1406	32%	11%	31%	80%	44.0	4.4						
94-95	480	365	76	1108	369	76	357	1012	33%	7%	32%	91%	32.2	3.2						
96-00	460	398	87	2943	934	162	926	2755	32%	6%	31%	94%	40.0	4.0						
01-04	768	425	55	3415	947	275	1265	3303	28%	8%	37%	97%	23.8	2.4						
05-06	865	475	55	2117	568	180	699	2041	27%	9%	33%	96%	4.1	4.1						
07-12	960	493	51	6699	2095	774	2539	6243	31%	12%	38%	93%	6.6	6.6						
13-16	785	520	66	3668	1239	234	1364	3346	34%	6%	37%	92%	8.9	8.9	37%	5%	33%	37%	6%	35%

Table 56: Strand-4 – Full Model Behavior Scores per Internet Engineering Task Force Chair

Dataset	DS: 1 IPA analysis (1991 to 2016)								DS4: 2017 Survey								DS3: 2013 Survey Scores					
Years	IBA	HS	VS	RC	TC	TI	OCBgc	OCBa	*N*	HS	VS	RC	TC	TI	OCBgc	OCBa	*N*	HS	VS	TI	OCBgc	OCBa
‘91-‘16	21643	5498	758	122	1768	7692	20060	58	88	5.40	5.46	2.78	3.24	5.58	6.11	6.11	125	5.65	5.40	6.13	5.90	6.26
’89-‘16	-	-	-	-	-	-	-	-	94	5.40	5.45	2.74	3.20	5.57	6.07	6.11	129	5.63	5.83	6.14	5.89	6.27
2015	8816	2979	273	20	413	2926	8025	71	-	SWA to A	SWA to A	SWD	SWD	A	A	A	-	SWA to A	SWA to A	A	A	A
2016	8721	3017	232	15	549	3075	8164	31														

	DS1: IPA Analysis (1991 to 2016) % of IBA								DS:4 2017 Survey								DS:3 2013 Survey Scores					
Years	S	HS	VS	RC	TC	TI	OCBgc	OCBa	*N*	HS	VS	RC	TC	TI	OCBgc	OCBa	*N*	HS	VS	TI	OCBgc	OCBa
‘89-‘90									6	5.33	5.40	2.22	2.56	5.33	5.56	6.00	4	4.95	5.90	6.50	5.67	6.50
‘91-‘93	32%	28%	3%	2.7%	8%	31%	80%	0.2%	13	5.04	5.45	2.87	3.15	5.49	5.62	5.77	12	5.25	5.61	6.14	6.19	6.17
‘94-‘95	33%	26%	7%	0.5%	6%	32%	91%	0.1%	8	5.33	4.69	3.08	3.29	5.50	5.67	6.06	7	5.20	5.69	6.62	5.81	6.57
‘96-‘00	32%	30%	2%	0.3%	5%	31%	93%	0.6%	14	5.39	5.66	2.80	3.00	5.60	6.17	6.21	19	5.52	5.89	6.46	6.19	6.34
‘01-‘04	28%	26%	2%	0.1%	8%	37%	97%	0.1%	9	4.61	4.84	3.48	3.37	4.30	6.26	5.06	16	5.75	5.64	5.92	5.98	6.16
‘05-‘06	27%	23%	4%	0.0%	8%	33%	96%	0.0%	4	6.08	5.75	2.50	3.17	6.00	6.83	6.00	8	5.50	5.81	6.38	5.96	6.19
‘07-‘12	31%	27%	4%	0.7%	11%	38%	93%	0.3%	25	5.04	5.87	2.35	3.25	5.95	6.21	6.52	60	5.83	5.91	6.01	5.93	6.26
‘13-‘16	34%	30%	4%	0.2%	6%	37%	91%	0.3%	15	5.00	5.27	2.84	3.44	4.76	6.27	6.33	3	5.57	5.90	5.89	5.89	6.33
’91-‘16	31%	27%	4%	0.6%	8%	36%	93%	0.3%	88	5.40	5.46	2.78	3.24	5.58	6.11	6.11	125	5.65	5.82	6.13	5.90	6.26

Strand-4 Quantitative Risks and Biases

The revised mixed-mode analysis used historiometric techniques for the IPA analyses for the predictive theoretical models (full and reduced) in this research to align to the recommendations Ligon et al. (2012) had for the following "sample plan formation," content-coding scheme development," "material preparation," "code logistics" and "descriptive and multivariate" analysis (p. 1118). Due to these recommendations, this researcher incorporated specific sampling plans (Appendix F) and a content-coding scheme development that created a specific coding methodology (Appendices E and G to N) on material gathered from the IETF website (IESG minutes and data referred to). Ligon et al. (2012) provided two types of performance criteria: counts of actions ("positive contributions" or "number of institutions established") and Likert-scale rating metrics (pp. 1124–1125). This researcher used a mixed-mode methodology using theme counts for Strand-1 behavioral actions and Strand-2 result data (decisions and results). The two surveys (Dataset 3: 2017 survey, Dataset 4: 2013 survey) used established scales with 7-point Likert responses. Ligon et al. (2012) stated that the goal of historiometric practices was to generate a "collection of quantitative data to which virtually any analytic technique can be applied depending on the research model given" (p. 1126) and then check for internal and external validity. Strand-4 quantitative analysis compared the results using descriptive statistics and multivariate techniques (scale reliability, normality, homoscedasticity, linearity, absence of correlated errors, and HRM) based on these historiometric best practices. This section reviews the internal and external validity risks and the potential bias in the results.

This analysis of validity risks and potential common bias began by determining the best practices in the statistical analysis used in historiometric mixed-mode research applicable to research into leadership in TMT and solidarity in groups. The best practices were described from historiometric research into leadership in TMTs and solidarity in groups that apply to this research. Based on these best practices, the risks were summarized in the IPA data analysis and the survey data analysis. Finally, validity issues were considered that would impact

Strand-4's results and common method biases in datasets 1-4. Table 57 provides a summary of the validity risks and bias. Appendix O discusses additional details on the validity risks and common method biases for the IPA analysis of Dataset 1 (10% sample 1991 to 2016), while Appendix P discusses these issues for the IPA analysis of Dataset 2 (100% sample 2015 to 2016). Appendix Q discusses additional details on the validity risks and common method biases for the Surveys datasets (Dataset 3: 2013 survey, Dataset 4: 2017 survey, and Open-Ended Questions in 2017 (Dataset 5).

Strand-4: Historiometric techniques and risk. The multivariate statistical techniques depend on the type of research. Ligon et al. (2012) reviewed over 60 historiometric research articles for "type of leaders," "research question," "conclusion," "data source," "main constructs of assessment," "construct encodings," and "statistical analysis performed in the research" (p. 1108). Within this group of 60 research articles, Ligon et al. (2012) summarized the following three articles on the leadership of teams: DeChurch, Mesmer-Magnus, and Doty's (2011) article as studying "effects of leadership important in mission-critical multiteam environments," Hunter, Johnson, and Ligon (2011) article as studying "leadership styles that influence … outstanding college and NFL [National Football League] coaches" of athletic teams, and Giambatista (2004) as investigating NFL teams whether a "coach succession influence[s] team performance differently than [a] owner succession" (pp. 1108–1110). DeChurch, Mesmer-Magnus, and Doty (2011) used the historiometric techniques combined online sources, just as this research did to augment the IESG minutes. Giambatista (2004) conducted a longitudinal study to examine data that included periods of coach and owner leadership, team performance (percentage games played that the team won), and objective indicators. Giambatista used analytical techniques, such as correlation, time series modeling, and regression. NFL coaches lead a team of assistant coaches to guide football players and trainers toward a winning season; thus, the NFL coach and his assistant coaches are similar leadership structures to the IETF chair and IESG members. Because Giambatista studied NFL coaches and owners for years, similar analytical techniques are valid for this IESG study (23 to 28 years).

The original research into solidarity examined solidarity, OCB, and task interdependence in groups and teams. Koster and Sanders (2006) used descriptive statistics, correlation, and regression modeling to look at solidarity dynamics to analyze 954 people in groups from nine different organizations. Sanders and Schyns (2006a) used descriptive statistics, correlation, and regression modeling to examine solidarity in "193 employees in 35 teams in a Dutch ministry" (p. 538). Koster et al. (2007) used a mixed-mode study that combined ethnographies with survey responses to examine solidarity and task interdependence. The ethnographies used a 5-point Likert scale for solidarity and a scale of 0 to 1 for task interdependence (1 = *task interdependence* and 0 = *no interdependence*). Based on Ligon et al. (2012), researchers have successfully used descriptive statistical and multivariate (correlations and modeling) statistical analysis in the group-level historiometric research of a TMT, ethnographic research on solidarity and task interdependence, and solidarity research. Therefore, these are valid quantitative techniques for this research.

Strand-4: Summary of risks in the Interpretive Phenomenological analysis data. The IPA analysis of the 10% sample, 100% sample, and 2017 OEQ on conflict share common concerns for validity (internal, external, statistical, and construct) and common rater bias, and a few concerns are unique to each IPA study. The shared internal validity concern for the 10% sample and 100% sample of the IESG minutes was that IESG minutes operated like pointers to online databases. The IETF continued to transfer online databases from the early years (1991 to 2006) to the IETF (2016b) Datatracker, but inconsistencies existed between online databases and missing data. This missing data might have been essential, and in some code notes, the researcher denoted where she witnessed differences from the notes.

The responses to the 2017 OEQ on conflict resembled the recorded history in the IESG minutes as a compressed discussion of the events, which assumed the researcher understood the context. The internal validity risks unique to the 10% study were the potential that the 10% sample did not represent the actual data. A unique validity risk to the 100% sample was that some IESG meetings held at IETF meetings or the IESG retreats occurred without public recording, but for 2015 to 2016, the researcher, in her duties as narrative scribe for IESG, had copies. In addition, the researcher redacted any private minutes to remove private discussions prior to being included in the IPA analysis, but the researcher included private conversations in the encodings. The external validity risks concerning the IPA analysis arose from the same issues of including in the analysis the data referred to by the compressed version of text plus the issue that RC at the IESG level was considered socially poor or unacceptable behavior. Thus, historical records actively dampened out RC questions, and this habit could have limited or encouraged RC comments on the 2017 OEQs on conflict.

The statistical validity and construct validity risks for the IPA analysis of Dataset 1 (10% sample of IESG minutes), Dataset 2 (100% sample of IESG minutes), and Dataset 5 (2017-OEQ on conflict) also had issues common to all three datasets, and some unique issues. The first common issue was statistical power due to the size of each sample that impacted group analysis per decision, per IESG meeting, and per IESG cohort (yearly). Dataset 1’s 10% sample had 1,853 decisions, 78 meetings, and 26 years. Dataset 2’s 100% sample had 2 years where the IESG minutes from 2015 recorded 820 decisions in 28 meetings, and the IESG cohort 2016 recorded 785 decisions in 27 meetings. Dataset 5’s sample had only 32 comments across 28 years. Depending on the scope of the multivariate analysis (all individual group decisions, group decisions in a meeting, or decisions for a cohort year), Datasets 1, 2, and 5 lacked sufficient statistical power. Dataset 5 statistical power only allowed descriptive statistics. The multivariate statistics needed 15 to 20 cases per correlation relationship and 74 cases for the reduced model.

Dataset 1 had sufficient statistical power to analyze behaviors for all decisions (N = 1,843) and decisions per meeting (N = 78), but 26 years only allowed one regression/correlation result to have validity. Dataset 2 had similar limits per year. Using the reduced model, the HRM modeling tests in Dataset 1 and Dataset 2 using the reduced model had enough statistical strength to provide a valid prediction that solidarity (S) predicts results. The HRM modeling tests on Dataset 1 behavioral data summarized per year ($N = 26$) found solidarity predicted 62.2% of the variance in the results and conflict (15.1%) together predicted 77.3% of the variance in the results. However, these HRM modeling test results were at risk due to being too low in statistical power. Construct validity centered on theme counts for behavioral scales (per decision, group, meeting, and IESG cohort year) rather than Likert-7 scales. The construct validity was strong as part of a historiometric study that combines Likert scales (in the surveys) with counts, but stand-alone counts might not have been as strong. Scale reliability tests based on theme counts for Datasets 1 and 2 did not have the same meaning as those based on Likert scales. The researcher ran scale reliability tests on each behavioral variable but did not exclude behaviors based on scale reliability alone.

The potential common rater bias had a similar basis in datasets 1, 2, and 5. The primary researcher did 90% of encoding for the IPA for IESG minutes for Datasets 1, 2, and 5. Even though the primary researcher used a codebook for encoding established by three independent raters, the single rater's viewpoint could impact the encodings. The primary researcher was trained in leadership (Ph.D. candidate), an expert in Internet technology, and had participated in IETF's standardization process for 30 years. The researcher's background as a trained expert allowed knowledgeable categorization, but it could have still contained unconscious biases.

Strand-4: Summary of bias risks in the survey data. The 2013 and 2017 surveys used established instruments, reducing the risk to internal, external, statistical validity, and construct validity issues. The researcher administered the 2017 survey anonymously, but it acted as a retest to the 2013 survey that was not anonymous. This scenario reduced the number of responses to the 2017 survey (25 responses for 94 cohort slots over 28 years) from the 2013 survey's level of response (41 responses for 129 cohort slots over 25 years). This situation created test-retest internal and external validity issues due to the self-selection of the participants to take or not take the 2017 survey. Section Q.5 examines "what if" the 2013 survey respondents had all responded to the survey to determine the participation would grow from 3 people per cohort to four-people cohort, the IESG cohorts from 2007 to 2012 would have a more extensive participation base (~50-

90% of the cohort), and the IETF chairs would have five chairs out of seven IETF chairs participating. The cohort means did not change significantly. The construct risks in the survey data are due to the level the survey instruments and published research describe. The survey instruments targeted an individual level and not a group. This research methodology assumed that the mean score of an IESG response per year was the group behavior, which might not have been accurate. A second construct risk was that RC at the IESG level was “socially unacceptable,” so individuals might have been used to damping any expression in public forums. This public concern could have influenced the 2017 responses to the conflict questions from Jehn’s (1995) instrument.

The researcher used Strand-2’s data on the results of IESG decisions per IESG cohort in the statistical analysis of the model. By using these data, the methodology avoided the common rater bias. The perceived results questions had a common rater bias since the IESG members, and IETF chairs responding to this information were rating behaviors and the results. However, the researcher considered the perceived results as one of the behavioral characteristics of an IESG. For example, the IESG cohort means in 2013 predicted 22.1% variance in the results (1991 to 2016) and 29.3% variance in perceived results (1991 to 2016), while perceived results had a positive correlation to actual results (0.538 at a significance level of $\rho < 0.01$). In contrast, the group means for 2017 survey behaviors did not explain any variances in the results per IESG cohort but explained 58.3% of the variance in perceived results per IESG cohort. The researcher understood from this common method bias that the IESG members who responded to the 2017 survey had common viewpoints on solidarity and the effectiveness of their IESG cohort.

Strand-4: Summary of risks in analysis. The Strand-4 analysis mixed-mode used the best practices from historiometric research that build quantitative data from historical data so that statistical analysis techniques were conducted of descriptive statistics and multivariate analysis. The researcher in the Strand-4 quantitative analysis conducted the descriptive statistics on all five datasets. However, the descriptive statistics for Dataset 5 were limited to the open-ended questions on conflict from Jehn's (1995) ICS instrument. The Strand-4 comparison of results in Table 52 compares the Likert-7 scale means for group means by providing the percentages of total individual actions (IBA theme counts) for each behavior. For example, the mean IESG cohorts score for solidarity from the 2017 survey from 1991 to 2016 on the Likert-7 point scale ranges from 4.30 to 6.10 (undecided to agree somewhat). The mean score for all IESG respondents is 5.74 on the Likert-7 point scale (agree somewhat). Solidarity theme counts per year in the 10% sample range from 123 to 392 themes or 21% to 31% of the individual behavior actions per year. The mean theme count per year is 257.9 themes out of a mean IBA count of 832.4 (31%). The total theme count for solidarity is 6705 out of 21,643 (31%). Because Ligon et al. (2012) seemed to keep the scores separate, the researcher listed these scores for side-by-side comparison for the descriptive statistics.

The multivariate statistics run on the quantitative datasets 1-4 performed the same sequence of tests. These tests included reliable scale tests, statistical tests to determine suitability for correlation and HRM tests, correlation tests, and HRM modeling sequences. The researcher ran reliability scale tests on the theme counts for behaviors, but these reliability tests do not correlate with reliability tests on Likert scale tests. Ligon et al. (2012) did not record reliability tests run with theme counts, so these reliability scale tests were exploratory. The suitability tests usually found homoscedasticity, linearity, and a lack of correlated errors in the quantitative data in datasets 1-4, but the normality varied. The theme count normality for the 10% sample was complete for the theme counts summarized per cohort for the data from the formal minutes except for RC. The theme counts for the 100% sample for 2015 to 2016 summarized per meeting had normality except for RC. The non-normality for the RC behavior data aligns with the active damping of the RC in the historical records. The original sequence for HRM modeling of the reduced model (HRM-4 modeling sequence) added one additional model (HRM-5) for the four datasets because task interdependence did not appear to operate as a control variable. HRM-5 modeling removes the control variable TI from the modeling tests as an exploratory venture to confirm whether it was a control variable or not. The researcher felt this addition useful due to theoretical issues, but it could be considered statistical method "fishing," a statistical validity risk. Although the statistical analysis of the four datasets provided a great wealth of information, the research conclusions from the Strand-4 quantitative analysis focused on the two hypotheses of the reduced model.

Table 57: Strand-4 – Summary of Bias and Risks in Datasets 1-5

Error or Bias Type	2017 Survey (Dataset 3 + 5) (2017 Survey risks and biases in section Q.3.7 with open-ended (OEQ) questions in Q.2.5)		2013 Survey (Dataset 4) (details in section Q.4.7)		10% of IESG Minutes (Dataset 1) 100% of IESG minutes (Dataset 2) 100% of Open-Ended Questions (Data set 5) (Details in sections O.7.1, P.7.1, and Q.2.5)	
Overall-Summary	low to medium	Internal validity: 2017 survey acts as a retest of the 2013 survey. External validity and statistical validity: impacted by the size of response (25 responses 94 cohort slots, 28 years) Construct validity: Survey instruments are tested per individual and not per group. Relationship conflict is "socially unacceptable" for IESG as a group. OCB and solidarity concurrent and discriminant Common rater bias: Using perceived results (PR) as a dependent variable has a common rater issue with behaviors. However, Strand-2 results do not have any common rater issues.	low to medium	Internal validity: 2013 survey was not anonymous. External validity and statistical validity: impacted by the size of response (41 responses 129 cohort slots, 25 years) Construct validity: Survey instruments are tested per individual and not per group. Relationship conflict is "socially unacceptable" for IESG as a group. OCB and solidarity concurrent and discriminant Common rater bias: Using perceived results (PR) as a dependent variable has a common rater issue with behaviors. However, Strand-2 results do not have any common rater issues. The researcher used perceived results(PR) as a	Low to medium	Internal validity: Minutes references to online databases (DBs) and early years (1991 to 2006) than 2007 to 2016. External validity: Minutes solid historical record approved 2-6 weeks after creation, but pointers to DB require hand merging. Statistical validity: The size of each dataset is reduced by summarization to meetings (78 meetings in 10% sample, 28 meetings in 2015, 27 meetings in 2016) and number of years in 10% (28) Construct validity: Theme counts used instead of Likert-7 scale for behaviors per person and groups (per decision, per meeting, per cohort year). Relationship conflict (RC) is "socially unacceptable" behavior for the IESG, so historical records reduce the number of relationship conflicts. Common rater bias: 10% analysis: Three experts did unique IPA analysis for half of the encodings in the 10% sample and jointly came to an acceptable encoding method for the codebook. The second half of the 10%

Error or Bias Type	2017 Survey (Dataset 3 + 5) (2017 Survey risks and biases in section Q.3.7 with open-ended (OEQ) questions in Q.2.5)	2013 Survey (Dataset 4) (details in section Q.4.7)	10% of IESG Minutes (Dataset 1) 100% of IESG minutes (Dataset 2) 100% of Open-Ended Questions (Data set 5) (Details in sections O.7.1, P.7.1, and Q.2.5)
	The researcher used perceived results(PR) as a group perception rather than a reality.	group perception rather than a reality.	sample was encoded by the primary research and reviewed by two experts. 100% sample (2015 to 2016) was encoded using the codebook set by three experts as raters, but the primary researcher encoded all the ratings.

Strand-4 Qualitative Analysis of Quantitative Data

The quantitative results from Strand-4's combination of the five datasets and the potential risks examined qualitatively from the perspective of mixed-mode research that combined historiometric research with survey research showed that the research supported Hypothesis 1 but not Hypothesis 2 at the group level for the IESG as a TMT in a technology change organization. Conflict did have a negative correlation in survey results and a positive correlation in historical records due to the damping of RC in meetings and meeting records. The researcher found the mixed-mode research on the historical records in the IESG minutes, and the snapshots of opinions from IESG members in two surveys strengthened this conclusion. This section examines Strand-4's quantitative results using the following three questions used in qualitative analysis of the quantitative data in Strands-1–3:

1. "Were the outputs qualitative effective?"
2. "Was the solidarity real?"
3. "Was conflict open or hidden?"

These questions helped provide insights on the Hares reduced model ability to explain the variance in perceived results and the variance in actual results.

The difference in Internet technology in 1986 and 2016 demonstrated the effectiveness of the IETF as a change organization. Internet technology in 1986 only existed in a few research institutions predominately located in North America. In 2016, Internet technology was critical for IT in business, education, and personal lifestyles worldwide. The question was whether the IESG cohorts as a TMT of the IETF led the IETF as a SDO from historical records and snapshots of the opinions. SDOs were organizations that develop standards for emerging technology for organizations in industries in the process of change. Internet technology had changed many other organizations, so as the IESG as a TMT led the technology change, it had resistance from existing SDOs that existed before the (ITU, IEEE; 3GPP, 2011), application-based consorts (W3C), or industry-based standards. Over time, the IETF has established a firm place among SDOs in the IT industry in networks, security, network management, and Internet interconnection of applications. However, some IESG members had been concerned that IETF keeps pace with the rapidly changing landscape of the applications and datacenter topologies.

Solid support within the IESG group from other IESG members and the IETF chair helps the individual IESG member handle the change resistance. Solidarity for an individual IESG member or IETF chair measures the extra efforts to live up to agreements and support others during unexpected failures, mistakes, and a mix of pleasant and unpleasant tasks. During periods of change within the organization, the IESG members who precipitate change encounter resistance within and from forces outside the IETF. The IESG minutes capture pointers to online records on the decisions of the group and the behaviors in those decisions. The IESG minutes as pointers to historical databases and records in this study are accurately captured and confirmed within 2-8 weeks of recording. The consistency between theme counts for solidarity in the IPA analysis (10% sample [1991 to 2016] and 100% sample [2015 to 2016]) and the survey responses from the two surveys (2013 and 2017) confirm that the solidarity detected is real and varies per IESG cohort.

The solidarity behaviors measured in the survey and the historical records were individual behaviors within a group context. The survey was used to measure individual opinions on the group per IESG cohort year. The IPA analysis considered behaviors indicated by a person's actions per IESG decision denoted the behaviors in an IBA. Each behavior existed (1 = *exist*) or did not exist (0 = *does not exist*). The assumption was that the group behavior for solidarity per decision was the sum of the individual behaviors exhibiting solidarity in their behavioral actions regarding a decision. The solidarity theme count for a meeting was the sum of the solidarity counts for all decisions in the meeting. Similarly, the sum of solidarity theme counts for the entire IESG cohort year was the sum of solidarity theme counts for all decisions through the year. The theme count approach considered total actions recorded in historical records per year versus the average perception of IESG member responses for an IESG cohort. IESG members may exhibit group behaviors that are not recorded (e.g., RC), and individual perceptions may not reflect reality. If perceptions of IESG members on the survey and historical records agree, it suggests substantial qualitative support for the theory. The one weakness in the historical record was that the longitudinal analysis was a 10% sample that estimated and did not define reality per year.

The mixed-mode research finds that the research supported Hypothesis 1 based on solidarity explaining different mounts of variance in actual results and the perceived results of consensus decision-making from 1991 to 2016. Solidarity explained at a significance of $\rho < 0.01$ showed that the variance in the actual results occurred at different levels depending on the dataset (10% sample of IESG minutes [1991 to 2016]: 62.2%, 100% sample of IESG minutes in 2015: 64.6%, 100% of IESG minutes in 2016: 43.7%, 2013 survey yearly cohort mean: 29.3% [1991 to 2016]. The historical record may explain more of the results than the survey because of the reduced participation in the survey or the type of interactions the historical record presents. The solidarity predicted 52% to 58% of the variance in perceived result scores in the surveys. The 2013 survey found a positive correlation between perceived results (0.538 for 1991 to 2016 at a significance of $\rho < 0.01$). The variance in the 2013 and 2017 surveys may be due to the smaller participation in 2013 or 2013 to 2017 team reorganization actions. The IESG cohorts in 2013 to 2015 completed reorganization and revitalization of all areas within the IETF, so these cohorts may feel the results of their efforts were beyond the IETF records. Because historical records and surveys confirmed Hypothesis 1, it had strong support, even though the 2017 survey's limited response only linked to perceived results.

RC was less socially acceptable in the IETF than TC during 1989 to 2016, and RC continued to be less socially acceptable than TC from 2017 to 2021. The small number of incidents of RC recorded in the IESG minutes and the 2017 survey responses that ranged from *disagree* to *undecided* were measurable outcomes of this social behavior in the IESG. In contrast, the open-ended questions captured reports of significant RC, such as factions in the IESG working actively or passively against the majority of the IESG or the IETF chair. The historical records varied based on the IESG cohort in amount RCs were not recorded in the minutes.

For example, the minutes for the IESG cohort in 2016 had incidents of less RC than the IESG cohort minutes in 2015. However, the responses to the open-ended questions suggested that more RC occurred. The researcher found RC behaviors in the historical records in meeting discussions or online forum comments. The IESG minutes and the online forum recorded, stored, and verified these incidents of group discussion within 2 to 6 weeks of the initial event; thus, these data did not degrade over time. Although records on early IESG discussions (1991 to 2000) might have suffered some error in transfer, the historical records did not degrade based on time. Unfortunately, the survey questions captured the view of a survey respondent at a particular moment in time regarding their IESG cohort. For example, an individual who took both 2013 and 2017 surveys might have changed their view of history. The open-ended questions on the 2017 survey indicated that IESG members actively reduced the expressions of RC recorded in the minutes, and some cohorts had underlying or hidden conflict within the group.

Hypothesis 2 was not supported, but conflict correlated to results in the historical record and perceived results in the 2017 survey. The theme counts for conflict per IESG cohort positively correlated to results in the 10% sample, but the conflict theme counts (task and relationship) occurred in 5% to 15% of the individual behaviors. The researcher found the RC themes occurred in 0.01% of the IBAs (20 themes) found in the IESG minutes from the 2015 IESG cohort and 0.2% of the IBAs (15 themes) in the minutes from the 2016 IESG cohort. TC themes occurred in 5% of the IBAs in the 2015 IESG cohort's minutes and 6% of the IBAs in the 2016 minutes. The conflict theme count totals from IPA of historical records (10% and 100% sample) contained mostly TC themes, so the positive correlation fit with group theories on the benefits of TC. The group means per IESG cohort for the conflict scale on the 2017 survey negatively correlated to perceived results at a value of -0.479 at a significance level of $\rho < 0.01$. The surveys queried for a combination of RC and TC, so the negative correlation with perceived results suggested conflict moderates the perception of IESG members on their cohort's effectiveness in decision-making. The issues with the 2017 survey's statistical validity due to the response size caused the researcher to wonder whether conflict correlated in the IESG response to actual results. Conflict was a complex construct in groups. The qualitative analysis of Strand-4 results confirmed Strand-4's results that Hypothesis1 was supported, and Hypothesis 2 was not supported.

Strand-5: Qualitative Results

Strand-5 triangulated the qualitative information detected in Strands-1–4 to determine if the results on the reduced model hypotheses are valid from a qualitative point of view. The methods in Strands-1–4 contained both quantitative data analysis and qualitative analysis of the quantitative data. The complete qualitative analyses for Strands-1–3 are found in Appendices O to Q: Strand-1 in O.7.2, Strand-2 in P.7.4, Strand-3's 2017 survey in Q.3.8, Strand-3's 2013 survey in Q.4.8, and the open-ended conflict questions in Q.2. The qualitative analysis was provided of Strand-4's quantitative data. Strand-2 collected qualitative information on IETF areas and WGs within these areas from 1989 to 2016 and changes in IETF organizations from 1989 to 2016. Appendix R summarizes qualitative information on the progression of technology standards with IETF areas and IETF organizational changes in Sections R.2 to R.5. Strand-5's codebook in Appendix M describes a five-step sequence for analyzing this qualitative data to create master theme charts and weighted node diagrams. Table 58 lists the six master themes found in Strand-5's qualitative analysis of the qualitative data from Strands-1–4, followed by a discussion of these master themes. Appendix R contains the weighted node diagrams, theme grids, and master theme charts with quotes for Strand-5 in Section R.1.

The master themes provided vital points that needed to be evaluated by qualitative analysis of the theoretical reduced model. Master Themes 1–3 were the subject of the following three questions in the qualitative analysis of the quantitative Datasets 1 to 5 (IPA analysis 10% [1991 to 2016], IPA 100% analysis [2015 to 2016], 2013 survey, 2017 survey, 2017 OEQ):

1. Were decisions with results effective leadership decisions?
2. Was the solidarity of the IESG cohort real?
3. Was IESG conflict open or hidden?

Effective consensus decision-making was stated qualitatively as consensus decisions that fulfilled the IESG leadership duties as a TMT to help the IETF fulfill its mission. The IETF (2021b) states its mission was “to make the Internet work better by producing high quality, relevant technical documents that influence the way people design, use, and manage the Internet” (para. 2). The IETF (2021b) also states that the IESG was “responsible for technical management of IETF activities and the Internet standards process” (para. 2). Restating this mission using leadership terms, the IETF (2021b) is a SDO leading change in Internet technology in the IT industry by creating high-quality relevant RFCs that influence the way people design, use, and manage the Internet. The IESG, as a TMT, leads the technical changes targeted for deployment within a decade and manages the IETF activities and the IETF standards process. The IESG reaches decisions using consensus decision-making. Effective consensus decisions in the IESG result in the publication of documents that influence Internet technologies toward a “better Internet” that is deployed for applications, networks, security, and management of Internet technologies. The IETF's original vision of the Internet is described in this section, as well as how the IETF standards have migrated toward that vision. The qualitative analysis on whether the solidarity in IESG TMTs was “real solidarity” uses Koster and Sanders’s (2006) research that gives a qualitative definition of solidarity. Jehn (1997, 1999) developed a qualitative description of TC and RC before creating the 6-item ICS scale. The researcher uses qualitative description and the qualitative results from the open-ended questions on conflict from survey 2017 to help judge if the conflict was hidden or real.

Feedback from IESG members and IETF chairs during the 2013 and 2017 surveys was that specific circumstances caused variation in the ability of an IESG cohort to review and publish standards, create and manage WGs, and handle IETF management. Master Themes 3 to 5 showed the unique pressures, focus, and conflicts due to change resistance for IESG cohorts under the leadership of a single IETF chair. The theoretical basis for the qualitative analysis of effective organizational leadership came from the comprehensive model for diagnosing organizational systems by Cummings and Worley (2005). Cummings and Worley provided a three-level (individual, group, and organizational) model for diagnosing organization systems based on the computer processing model that inputs are processed to produce outputs. At the organizational level, Cummings and Worley's (2005) model posits that an organization takes inputs from the general environment and industry structure and processes these inputs via organizational design components to produce effective organization outputs. The design components are the "technology," "strategy," "structure," "human resource systems," and "measurement systems" (p. 89) filtered through an organization's culture. Yukl (2010) pointed out that leaders influenced the culture of a group through "espoused values and visions, role modeling and attention, and reactions to crisis" (p. 306). The researcher applied Cummings and Worley (2005) to diagnose whether the IESG decisions were effective results by reframing the question on effectiveness as the following: "Did the decisions of an IESG cohort help the IETF be an effective change organization leading the global IT industry toward the IETF's vision of a 'better internet' through the influence of high quality, relevant standards for that period?" (p. 20).

The IETF chair leads the IESG within the global IT environment to continually manage the organizational components and culture to create influential, high-quality, relevant standards. One way to summarize the organizational diagnostic information from Cummings and Worley's (2005) model is a SWOT (strengths, weakness, opportunities, and threats) chart. The SWOT of the IETF's organization is examined during the tenure of each IETF chair. Next, the section examines how each IETF chair's vision/strategy for IETF technology standards and organizational changes influence the IETF culture. The organizational changes include organizational structure, human resources, measurement systems, and tools. In addition, the IESG culture the IETF chair's leadership behaviors (espoused values, vision, role modeling, attention, and reaction to crisis) create influences the solidarity and conflict in the culture of the IESG. Finally, the qualitative analysis concludes with two hypotheses of the reduced model.

One potential theory for evaluating the IETF as a change organization in IT is the organizational learning theories. The IETF, dedicated to creating standards for a “better Internet,” must continually learn how to create better standards. The IETF organization could have single-loop (adaptive) learning, double-loop learning (generative), or “dueterolearning,” which Cummings and Worley (2005) defined as “learning how to learn” (p. 20). The Cummings and Worley (2005) model and Yukl’s (2010) leadership recommendation aligned with this researcher’s solidarity-conflict model, but the learning organization theory showed an additional complexity not warranted for the qualitative analysis for this research project.

Master Themes

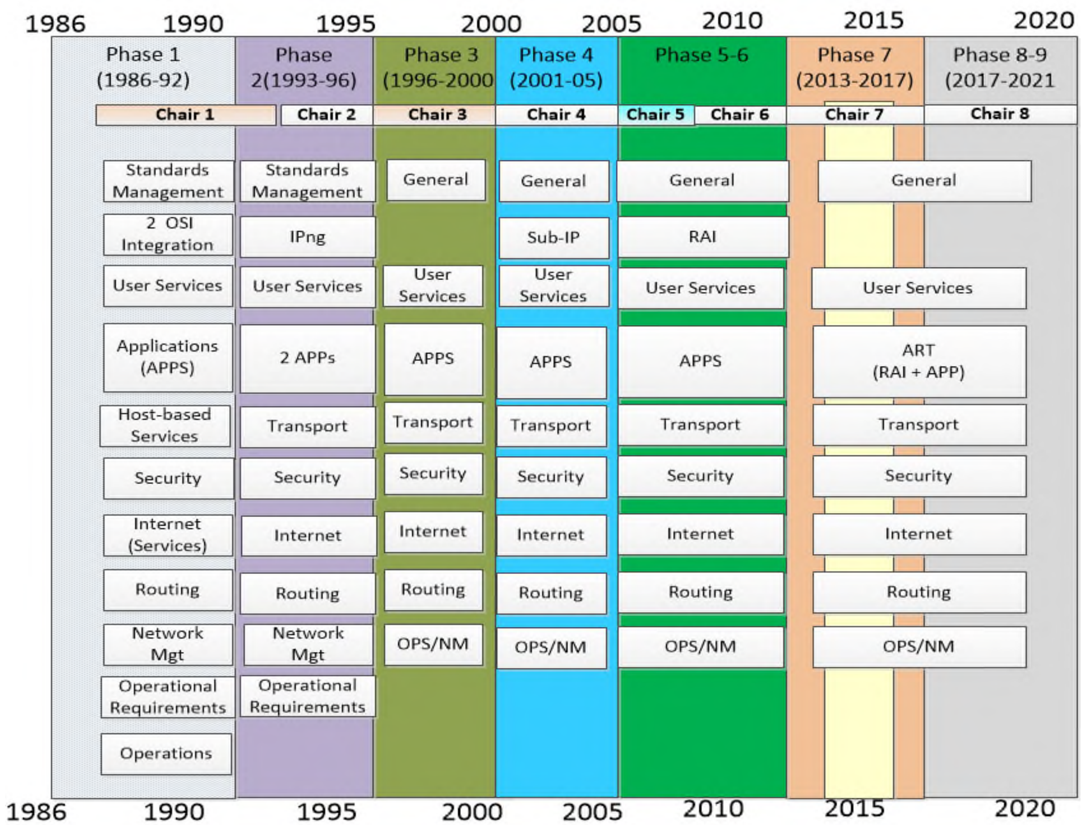
Based on the IETF mission and the IESG function of the IESG as a TMT, the Strand-5 theme analysis found six master themes. The theme chart brought up themes from the three questions in Strands-1–4 regarding qualitative validation of the effectiveness of IESG decisions, the existence of solidarity, and the nature of the conflict (task or relationship and open or hidden). Three other themes arose from Strand-5’s qualitative analysis of Strand-2 data regarding IETF areas and IETF organizational changes, Strand-2’s data on the challenges, plans, and accomplishments during the tenure of each IETF chair, Strand-3’s qualitative analysis of open-ended conflict questions (OE-C-Qs), Strand-4’s quantitative analysis. The fourth theme was “unique set of pressures per IETF chair’s tenure.” During each IETF chair tenure, these pressures were a unique set of internal and external pressures that the researcher summarized using a SWOT (strength, weakness, opportunity, and threat) analysis. The fifth master theme was a “unique focus per IETF chair.” Each IETF chair had a unique focus on what changes to Internet technology needed to progress to advance the Internet and how to adapt the IETF organization to progress those changes better. The IETF chair’s focus led the IESG cohorts to change specific technology standards and change or maintain the IETF organization. The sixth master theme was “variances in conflict arises out of the resistance to change moderated by the leadership of the IETF chair.”

Table 58: Strand-5 – Master Themes Detected

Master theme	Where themes detected
1) Effectiveness of consensus decisions 2) Real Solidarity 3) Conflict varies between open and hidden expression	Strands 1-4
4) Unique set of pressures per IETF chair tenure 5) Unique focus per IETF chair tenure 6) Conflict variances arise out of resistance to change moderated by the leadership of the IETF chair	Strand-2 data and Strand-3 OE-C-Qs Strand-4

Themes 4, 5, and 6 showed the necessity for the qualitative analysis to examine IESG cohorts grouped per IETF chair. The IETF nomcom group uniquely selects an IETF chair for his/her ability to nurture existing WGs or create new WGs that produce high-quality, relevant standards for the upcoming four years. Relevant standards may need to enhance old standards or address upcoming needs due to changes in the network physical connection technologies, changes in the needs of the applications for Internet connections (e.g., real-time video over phones), or management protocols. The set of unique pressures during an IETF Chair's tenure came from the changes in the general environment, such as business downturns (e.g., in the "dot.com" bubble burst in 2000), dramatic increases in bandwidth requirements from the IT industry (e.g., the explosion or the world-wide-web or mobile phone Internet usage), or pressures from other SDOs (ITU, IEEE, 3GPP [2011], or W3C). Strand-3's analysis of the open-ended conflict found that relationship and TC aligned with external pressures, internal change resistance, or internal factions within the IESG during an IETF chair's tenure. The IETF cohorts each of the seven chairs led from 1989 to 2016 are IETF Chair 1 (1989 to 1993), IETF Chair 2 (1994 to 1995), IETF Chair 3 (1996 to 2000), IETF Chair 4 (2001 to 2004), IETF Chair 5 (2005 to 2006), IETF Chair 6 (2007 to 2012), and IETF Chair 7 (2013 to 2016). Two other chairs had served since 2016 (IETF Chair 8 [2017 to 2020] cohorts and IETF Chair 9 [2021 to current]). Figure 33 shows the diagram of IETF areas under IETF Chairs 1-9.

Figure 33: Structure of IETF Areas of Work (1986 to 2020)



Internet Engineering Task Force's Founding Vision of Internet

The vision for the Internet as worldwide multimedia communication between individuals and their computers arose out of early work at the Massachusetts Institute of Technology and Defense Advanced Research Projects Agency (DARPA; Leiner et al., 2009). Licklider and Clark (1962) described the "Galactic Network" concept in a series of memos as "a globally interconnected set of computers through which everyone could quickly access data and programs from any site" (p. 23). Licklider headed DARPA in October 1962, and his association with Leonard Kleinrock, Lawrence G. Roberts, Steve Crocker, Vint Cerf, and Jon Postel helped form the architecture of packet-switched Advanced Research Projects Agency Network (ARPANET) in 1968 and the TCP/IP protocols, such as IP from RFC 791 (Postel, 1981a) and TCP from RFC 793 (Postel, 1981b).

During the 1980s, the widespread deployment of LANS, PC, and Ethernet technology allowed these nascent TCP/IP protocols to flourish beyond research networks, so the Internet required a “scalable distributed mechanism for resolving hierarchical hostnames (e.g., www.acm.org) into an Internet address” (Leiner et al., 2009, p. 26). Mockapetris (1983a, 1983b) created the Domain Name System (DNS) (RFC882 and RFC883), which became an Internet standard in 1987 (RFC1034 and RFC1035; see Mockapetris, 1987a, 1987b). Researchers found that the ARPANET allowed them to send files, email rapidly, or chat on blackboards.

By 1985, DARPA could no longer use it for networking research, so the NSFNet decided in 1985 to 1986 to create an NSFNET backbone that could carry research traffic and commercial traffic offered by regional networks. The NSFNet decided that network standardization would occur in an open fashion using the same vision as DARPA. The NSFNET and DARPA discontinued the DARPA Gateway Algorithms and Data Structures task force and formed two new task forces: the IETF and the Internet architecture task force (see IETF, 1986, p. 3). The IETF retained many of the members of the DARPA task force who became IESG members and IETF chairs.

The vision of a “better Internet” expressed at the first IETF meeting had specific requirements for the network connectivity, types of applications that needed to be supported, and management requirements (IETF, 1986). The network needed a robust, rich topology connected hierarchically (e.g., local network, regional network, and national network) that enables virtual connections between users and their computers. The vision suggested the Internet needs to support applications for email and database transfers between two users and multi-destination delivery of email, database transfers, video conferencing, and parallel computing. Applications should also be able to query the network for locations of generic resources. The network should be able to prioritize certain types of traffic and services. Management needed tools to quickly determine problems or loads on the network. NSF empowered the IETF to set the requirements for hosts, routers, and gateway routers that connect to the NSF-funded regional networks and backbone in open standards environments. Setting these requirements and enabling management of the hosts and networks via the network was the chief focus of the IETF Meetings 1 to 15.

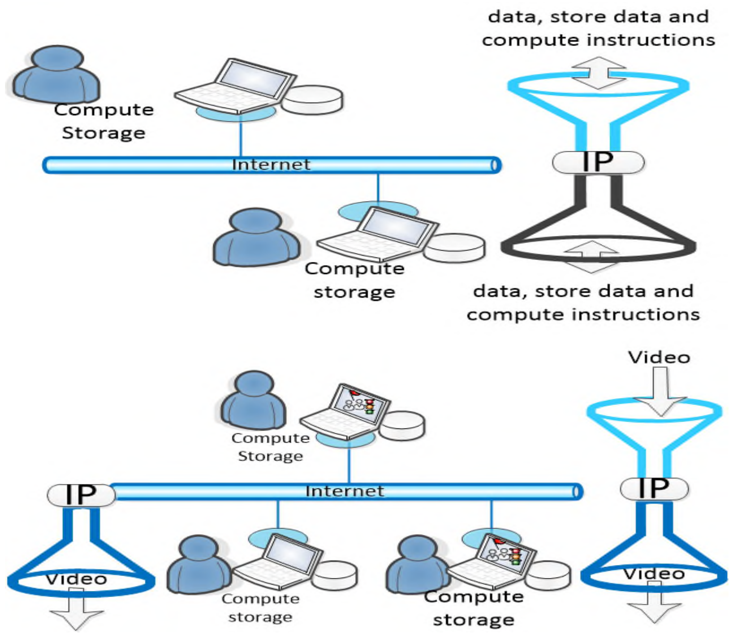
IETF members discussed email security at IETF 6 (1986) during a discussion augmenting an Internet host requirement to include secure email for military use. Security for management protocols was discussed in IETFs (1991, 1992) 8 to 14 during 1987 to 1989 in the Authentication (auth) WG prior to forming the IESG. After the IESG's formation with Security Area, guidelines for secure operation of the Internet were written into a standard by Internet Security Policy (spwg) from 1989 to 1991, published as RFC1281 (Pethia, Crocker, & Fraiser, 1991). These guidelines delineated what security was the responsibility of users, computer hardware and software vendors, and network service providers. RFC1281 recommends that "users, service providers, computer hardware and software vendors are responsible for cooperating to provide security" (Pethia et al., 1991, p. 1). The IETF's vision on security is that "technical improvements in Internet security" would be sought on "a continuing basis" and noted that security extensions to the protocol suits should include host security (passwords), applications ("file transfers, telnet, and [e]mail"; Pethia et al., 1991, p. 1), routing, and network management. These improvements moved Internet protocols from 1989 to 2016 from no security to authentication to secure protocol additions to privacy additions.

Migration Toward the Internet Engineering Task Force's Vision of the Internet (1989 to 2016)

The qualitative analysis first examined whether the leadership of the IESG cohorts from 1991 to 2016 fulfilled the mission of the IETF, creating a progression of high quality and relevant standards which influenced the way people design, use, and manage the Internet. The IESG members in these cohorts had to work together to create WGs, approve documents as standards, approve IETF management of registry information and organization structure. Individual IESG members led areas of work within the IETF to produce the standards documents that the IESG approved. The first step in the Strand-5 analysis was to determine the progress of WGs and standards within an area. Appendix R contains summary charts per Area on technical and organizational work progression. During specific periods, the IESG changed the IETF's areas, so simply using per area summaries of standard progressions did not provide a consistent analysis (see Figure 33) of the progression of technical and organizational standards. Therefore, the researcher grouped areas in the IETF based on a theoretical model of network architecture.

Strand-5's qualitative analysis used a network architecture model from Day (2008), denoted as the network inter-process communication (IPC) model, as a theory by which to group technical areas in the IETF into consistent large groups. These larger groups allowed a consistent analysis of effective progression toward the IETF vision of a "better" Internet. Day's (2008) network IPC model posits that networks are distributed IPC facilities running over physical networks. The virtual connections between computers carry data and instructions on what to do with the data (e.g., store, compute, and show video). Day's network IPC model closely aligns with Licklider and Clark's (1962) description of the functions of a galactic network that allows people and computers to communicate. Figure 34 shows an example of the logical application of Day's (2008) IPC model to the IP Internet for applications that process data and video applications across the Internet. Based on this network IPC model, Strand-5's analysis grouped the IETF areas into an application group, a net WG, a management group, and a security group. In addition, the researcher included IETF areas regarding standard management or general organizational standards into a single group.

Figure 34: Day's Network inter-process communication model applied to Internet protocol (IP) networks.



Strand-5's qualitative analysis examined these four technology groups and the one organizational group to determine the progression toward the IETF's vision of a better Internet. This section considers the progression of the four technology groups. Figure 33 shows the IETF areas that have existed from 1989 to 2017. The IETF areas grouped into the application area included applications (APPs), host-based services, real-time applications and infrastructure (RAI), and applications and real-time (ART). Day's (2008) network IP model indicates that creating logical groups of pipes over physical connectivity augments physical connectivity. Based on this concept, the researcher included the IETF areas for Internet services, Internet, OSI Integration, Routing, operational requirements, IPng, and sub-IP into the network group. The IETF areas included in the management group are network management, operations, user services, and operations/network management (OPS/NM). The security group simply contains the security area plus some general area work in 2013 to 2017 on pervasive monitoring as a technical attack. Finally, the IETF areas standardizing organizational changes are standards management and general area.

Strand-5's analysis arranged the data from the summaries of the IETF areas into a progression of relevant standards in each of the four technology groups (application, network, management, and security) and the organization group. The researcher summarizes the progression of the technology standards produced by the IETF toward the vision of the Internet in a series of diagrams that show technologies and standards versus time (Figures 35 to 42). Based on the Network IPC theoretical model, the first thing to consider is the growth of the physical bandwidth from 1989 to 2021.

Figure 35 shows the growth of physical bandwidth. Relevant technology specifications lead deployed technology by 1 to 5 years, so the researcher tracked technology 5 years after 2016. The available physical bandwidth in 1989 was a composite of local area networks (LANs) comprised of Ethernet (10 Mbps), Token Ring (16 Mbps) and FDDI (100MB) technology, 1G phone network (2Kpbs), and 1.5 Mbps NSFNet (a national network funded by the US Government). The available physical bandwidth in 2021 consisted of LANs with 400 Gigabits per second (Gbps), 5G phone networks (80-100 Mbps, or 1–3Gbps), Wi-Fi connections (1Gbps), and global WAN networks with 100-200 terabits per second (Tbps) (Cisco, 2021). By 2018, 3.9 billion users accessed the Internet connecting to 1.7 billion sites on the W3C, which provided over 220,000 unique third-party applications running in cloud data centers. Most data traffic travels in the high-speed LANs in cloud data centers (e.g., Google's or Amazon Web Services), and 39% of these data centers are hosted in hyper-scale data centers (Cisco, 2018). Cisco (2021) experts estimated the number of Internet users would grow to 4.5 billion in 2020, but this number was lower than actual numbers (e.g., 4.6 billion; Johnson, 2021) due to the COVID-19 pandemic. The IEEE standardizes most physical device standards (Ethernet, Wi-Fi), and the W3C standardized many web technologies.

The IETF (2021a, 2021b) standards in the application group showed a clear progression from 1989 to 2016 of high quality, relevant standards that advanced the Internet communication applications to support 3.9 million users, 1.7 websites, and 220,000 applications running in cloud data centers. The types of applications deployed before 1989 included faxes, video conferences, phone calls, company email, data transfer to data center programs from mini-computers, access to data center programs via phone lines. Early Internet standards created open standards for email, VOIP, W3C, online chat (Jabber), and video conferencing. The IETF published a progression of open standards that continually enhance the application protocols and define supporting functions such as email addresses, URN/URLs for websites, interconnections to phone numbers (mobile and fixed). For example, the initial applications for email were adaptations of X.400 (ISO standard) for the IP Internet, IMAP, and POP email with X.500 for directory services. By 1992, Internet standards stopped considering the OSI applications as relevant standards. The IETF cohorts from 1993 to 2016 guided a sequence of work to enhance email protocols (IMAP, POP, SMTP, MIME), secure email, and added contacts (vCard), calendar (iCalendar) protocols, and instant message features. Figure 43 shows a graphical summary of this progression.

During 2008 to 2015, new standards (webrtc, AVT) allowed applications to stream data between web-user applications and IP content delivery applications to send movies to users. Also, during this period, the IETF worked on standards to allow Emergency Warning messages via IP networks. In 2015 to 2016, vehicular networks were examining protocol to provide both emergency messages and infotainment data over vehicular networks. The IETF also began standardizing Internet of Things (IoT) applications and protocols for communication and management. The application group of IETF areas has a clear progression of successful applications, except for high-speed data storage applications. During the IESG meetings during 2013 to 2017, some IESG members raised concerns about the IETF was staying relevant for Internet protocols for web applications with a very short life cycle (months). During 2014 to 2017 work was begun to enhance the speed of transport (QUIC WG). The IETF leadership selection committee (nomcom) selected an IETF chair from that ART area (IETF Chair 8) to focus on applications and improve organizational processes and tools.

The progression of standards in the IETF (2020e, 2020f) standards in areas associated with the network group has two threads: stable connectivity over physical topologies and virtual network connectivity over the stable logical topologies. Figure 38 shows the progression of network protocol WGs. Standards work started in the network area during 1989 to 2000 focused on stable connectivity over physical technologies (LANs, Phone networks [1G-2G], satellite, Carrier [e.g., AT&T] links or virtual links [SMDS or IP over ATM]) and a revision to the original IP protocol (denoted as IPv4) with more address space (denoted as IPv6). From 1997 to 2017, the IETF network-based areas added virtual network technologies for wide-area networks and data center networks (e.g., mpls, lsvr, and rift WGs). The IETF (2020e, 2020f) progression of relevant standards in the network group of areas (IP and routing) tends to occur in long-lived WGs (15 to 25 years). These areas contrast with the application group of Areas that use short-lived protocols for most work and a few long-lived WGs. Two exceptions to long-lived areas were the IPng and the Sub-IP areas. The IPng area focused on making decisions on what went into the IPv6 protocol and associated changes for the Transport and Application Area. The IESG (2020) formed the IPng area to encourage the cooperation of experts from the protocols for applications, end-to-end transport of data, IP protocols, routing, and management to work on the revision to the IPv4 protocol. After the initial definition of IPv6, the IESG disbanded the IPng area. Similarly, the IESG created the sub-IP Area to help IP network, routing, operations, and network management (NM/OPS) experts to define initial relevant standards for the virtual network topologies jointly. After a period of joint work, IESG members transferred the work back into these existing IETF areas.

During 2010 to 2017, the IETF (2020k) network area focused on network protocol changes to enable transitions of data from corporate data centers to cloud data centers. These cloud data centers were operating in large and hyper-scale data center facilities. The protocols within the network technology group of Areas have a clear progression of relevant protocols. During 2013 to 2017, IESG cohorts believed that a reorganization of the WGs on the Internet, routing, and NM/OPS areas were needed to address the changing needs for data centers, wireless and mobile networks, machine to machine networks (denoted as the IoT), and increased need for virtual topologies to support applications.

The progression of standards in the management group of IETF (2020k) areas augments the standards in the network and application technology by refining requirements for network management and designing protocols to configure and monitor the IP network and application protocols. For example, Day's (2008) discussion on the Network IP model indicates the need for an IPC management task that manages IPC resources, network resources, and systems access control (p. 205). Figure 38 overlays the IETF standards work for operational requirements, network management protocols, and network management data definitions on the progression of IETF network standards, with the WGs for this appearing in red. The IETF has discussed access control in WGs in the security and management protocol group. The IETF standards work elected to use a simplified model for network management protocol (SNMP) with specific formats for management information databases (MIBs) rather than the ITU-T CMIP (X.700) protocol. The IESG cohorts in 1989 to 1991 agreed to standardize SNMP Version 2 with a small amount of security could monitor network and application technologies. During initial deployments of IP technology (1986 to 2000), these simplified MIBs allowed the quick definition of management extensions for the Internet protocols in the network and application areas of work. However, the simplifications to the MIBs and security of the SNMP protocol caused the protocol to remain unused for configuring networks. The IETF management WG attempted to work these problems in Version 3 of SNMP but switched in 2001 to a data-driven paradigm. The IETF management groups developed secure management protocols for these data-driven paradigm for data monitoring (IPfix) and configuration (NETCONF/RESTCONF) operating on Yang data modules. The IESG cohorts in 2013 to 2017 focused on standardizing Yang modules for all areas of the IETF by chartering WGs (e.g., i2rs) or design teams within each IETF area to progress the work quickly.

Integrating the data-driven management paradigm into Internet technology initially met resistance since management systems were considered a differentiator by network equipment vendors (e.g., Cisco), network operators, and data center operators. However, as these management systems age, the new paradigms are a core for the next generation of network management. Deployments of proprietary management tools mix the use of SNMP with MIBs and the newer protocols (IPfix, NETCONF, and RESTCONF) Yang modules. The IESG cohorts from 2001 to 2016 have handled the difficult task of maintaining the SNMP MIB paradigm and developing the data-driven paradigm protocols with Yang models. This dual thread consumed considerable efforts by specific IESG ADs leading the NM/OPS area and the IESG reviewing specifications. The decisions by IETF Chair 7 and the IESG cohorts during 2013 to 2016 pushed the management to IETF protocols toward the data-driven paradigm. As a result, the IETF management protocols and the Yang models have become "the prototypes for new proprietary management models" for IETF protocols (Haas & Patel, 2021). The management area shows a steady progression amidst challenges due to the initial choice for SNMP and external threats from proprietary work.

The progression of IETF standards in the security group of IP Areas and associated WGs in other areas treated security as an enhancement rather than a requirement for application and network technologies. These enhancements progressed from authentication to security to encryption for privacy and concerns regarding the invasion of privacy through restricting pervasive monitoring. The security technology shows a clear progression in the IETF application technology, as Figure 40 illustrates. Security has long-lived strands of work for the following significant efforts in the application area groups: (a) secure transport layer (TLS in the TLS WG), (b) authentication protocols (OTP, Kerberos [KRB-WG], ABFAB, ACME, OAuth), (c) securing of DNS (DNSSEC, IPseckey, dane, DNSSD, DPRIVE), (d) private key infrastructure (PKIX, pk4ipsec, hokey, ipsecme, curdle), and (e) enhancement for web and IoT security. The security area in the network area includes (a) securing of the IP protocol (IPSEC, mobile ILE, IPSECME WGs), (b) point to point link security (PPPEXT), (c) routing security requirements (rpsec), (d) key and authentication for routing protocols (karp), and (e) securing the BGP protocol and infrastructure (SIDR and SIDRops).

Figure 41 illustrates this progression of security in the network protocol. The NM/OPS and security area split the standardization work on the access protocol WGs (Radius, AAA, Radius extensions [RADEXT], Kerberos, EAP, and EMU). During the 2013 to 2017 focus on NM protocols, the security ADs managed WG working on access protocols. The security technology area shows a steady migration of security features from authentication to security, to encryption, and to privacy. The privacy work in 2013 to 2016 pushed back on protocol allowing an invasive level of monitoring of the user data passed via IETF protocols as needing to be restricted except under government-mandated lawful intercept (an equivalent of court-ordered wire-tapping). This work included reviews of the application, network, and management technologies to determine how to balance privacy and monitoring. One criticism of the IETF protocols was the lack of security in the original protocols and the slow migration toward secure protocols. Other parties consider the slow inclusion of security technology a benefit since security technologies have improved over time. The IESG cohorts from 1989 to 2016 balanced the need for high-quality security with the ability to operationally deploy security in IP protocols within research networks, commercial networks, and data centers.

The organizational areas of the IETF focused on writing standards which defined the open process for creating and publishing standards on the IETF as an organization. These areas published standards on the IETF process, the WG process, the nomination committee (nomcom), intellectual property, and registry policy for IETF documents. Figure 42 shows the sequence of these published standards and the WGs that created these standards from 1989 to 2020. The WGs in this group of IETF areas created the initial revision of these documents from 1990 to 2004 except for the nomination committee. The nominations committee selects the leaders for positions in the IESG, the IETF chair, the IAB, and administrative groups (IETF’s administrative director [denoted IAOC director] from 1991 to 2016 and IETF trust) that handle meeting venue selection and IPR issues. This area collected input from the IETF community and authored a document about IETF needs and concerns regarding the U.S. government's transition of the IANA (2020) registry from funding by the U.S. government to non-government funding. The WGs groups and BOFs from this area worked on recommendations for changes to the IETF tools and the RFC format. The organizational standards area shows a progression process documents from 1989 to 2016 regarding IETF standards production.

The Strand-5 qualitative analysis concludes that the technology and organizational changes managed by the 1989 to 2016 IESG cohort had effective decisions to create a progression of standards that a level of quality and relevance to be deployed and grow the internet. This qualitative conclusion confirms the quantitative results that while each IESG cohort percentage of decisions with measurable results could vary between 40% to 96% of the IESG cohort decisions, the mean percentage was 61%. The mean rate of decisions resulting in a measurable result for the cohort under a single IETF chair varied less from 51% to 87%, with a mean value per IETF chair of 63%. This sequence also matches varying degrees of the progression of technologies during the leadership of different IETF chairs.

As the master themes showed, the IETF efforts to change the Internet continually toward a “better Internet” faced unique external pressures from the IT industry, the SDOs, and the IETF members per IETF chair tenure due to conflicts due to resistance to changing the technology. Part of the internal pressures each IESG cohort experienced was the IETF chair’s focus on changes in technology and IETF organizational structure and conflict within the IESG (hidden and open conflict). The following three sections summarize the unique pressures for the IESG cohorts under a particular chair, the unique focus established by each IETF chair, and the variances in conflict due to TC, RC, and change resistance. After examining these facets, the Strand-5 qualitative analysis concludes with an analysis of whether the results from the qualitative analysis support Hypothesis 1 and Hypothesis 2 of the reduced model.

Unique Pressures for Internet Engineering Steering Group per Internet Engineering Task Force Chair Tenure (Theme 4)

Feedback from several IETF chairs during the 2013 and 2017 surveys indicated different external and internal pressures during their tenure as an IETF chair. The qualitative analysis found that these reports by IETF chairs were accurate, causing variations in solidarity and conflict. The IESG cohorts led by an IETF chair had unique external pressures from the global IT industry and web of SDOs and internal pressures caused by internal forces within the IETF. The researcher summarizes these pressures using a SWOT (strength, weakness, opportunity, and threats) analysis of the IETF organization shown in Table 59. The researcher classified the external pressures on the IETF as opportunities or threats and internal pressures as strengths and weaknesses of the IESG at the beginning of the IETF Chair's tenure. If these changed mid-term, there is a dating on the changes.

The SWOT analysis demonstrated that the IETF, as an SDO for Internet technology, operated as a change agent in the IT industry during the 1989 to 2016 cohorts. Internet technology originated in research networks funded by governments in the United States, Europe, Middle East, and Asia. The success of early adopters of Internet technology, such as email and websites, caused the IESG during 1989 to 2000 to have tremendous pressures to publish standards for new Internet technology quickly. This growth is often denoted as the "dot com" bubble where overhead demands for websites caused tremendous market growth for Internet technology and services in the IT market. At the same time, existing SDOs either found the IT market for their technology growing with the IETF or shrinking. For example, the IEEE and W3C (consortium) found that IETF growth encouraged growth in their sectors. On the other hand, the ITU found the OSI 7-layer technology market shrinking as the U.S. government abandoned OSI deployments requirements. The internal pressures were a balance of the strengths of the founding of the IETF and the weakness of being a new SDO in the IT industry filled government SDOs (e.g., ITU) and professional organizations (e.g., IEEE). The IETF was founded on a shared vision held by the leaders with a clearly defined practical mission set by the NSF as part of the efforts to launch a national research backbone with regional networks. The NSF's plan called for commercialization of the internet technology and regional networks within 3 to 5 years, so it actively funded efforts to transfer technology to the commercial sector. The NSF created the IETF without organizational documents, so during the first 15 years, the organizational processes were documented as IETF standards. The tremendous growth in the use and commercialization of Internet technologies helped foster relationships with IEEE and new SDOs such as the W3C Consortium and caused conflict with the ITU. IESG cohorts from 1989 to 2000 dealt with different mixes of these pressures per year.

The second decade of the IETF (2016b) started with a tremendous slow down as the "dot com bubble" burst in 2000, causing the IT market for Internet technology to slow down as companies failed or consolidated or reduced offerings. The IESG from 2001 to 2012, led by three IETF chairs, dealt with economic changes due to the dot.com failure and resulting consolidations. The consolidations aided the growth of some cloud data center companies, such as Google. New technologies in 3G mobile networks, 10G Ethernets, and Wi-Fi network access aided growth from 2001 to 2007. Sensor networks based on Wi-Fi technology also increased. YouTube started streaming video services in 2005 (Arthurs, Drakopoulou, & Gandini, 2018), with Netflix following in 2007 (Lobato, 2019). The 2008 stock market crash originated in the banking sector, so this financial crisis, while it slowed down IT market expenditures on new IT technology, did not stop the growth of technology by IT companies. Secure payment (PayPal), Cloud data services (e.g., Google and Amazon Web Services), and mobile phone applications grew during this period. Increased deployments of applications outside North America caused pressures for the internalization of these applications. The SDOs conflict between IETF and the ITU erupted in public disagreements over MPLS virtual technology requiring IESG in 2005 to 2012 to address these issues.

The strength of the IETF (2016a, 2016b, 2016c) during the second decade was its continuing ability to provide a progression of high-quality standards for the existing application and network technologies, which increased functionality, security, and management interfaces. Network management work became challenging as deployments found SNMP and MIB technology had issues with large-scale secure deployments. The experts in the network management area decided to switch to data-focused NETCONF and yang models. Conflicts within IESG occurred as IESG members pushed hard to have the documents from their IETF area prioritized for review by the IESG so publication could meet external deployment deadlines. For example, the application group of IETF areas pushed hard to get necessary IETF reviews and approval for changes out to mail, web, streaming video, and real-time video to encourage the SDOs who were part of 3GPP (2011) to adopt relevant Internet technology for 3G and 4G phones. The network group of areas worked hard to publish specifications relevant for growing features in Wi-Fi, Ethernet, and data center networks that could be secured and managed. This conflict for review resources within the IESG caused conflicts within the IESG. IESG document reviews also had conflicts over IPR claimed in the document since an IPR review was part of the IESG document review process, and different Areas in the IETF allowed tolerated different levels of IPR in their standards.

The third decade (2010 to 2020) began with a widespread acknowledgment that the work areas in the IETF (2020e, 2020f, 2020g, 2020h, 2020i, 2020j, 2020k, 2020l) needed to be revitalized and reorganized. IETF Chair 6 led the IESG cohorts from 2010 to 2012, so the researcher considered these IESG cohorts with the second decade. The IETF Chair 7 led the IESG cohorts from 2013 to 2016, so the researcher includes these IESG cohorts in this qualitative analysis. The IESG cohorts during 2017 to 2020 were led by IETF Chair 8 and are outside the period from 1991 to 2016 this research examines. The research 5G technology caused new opportunities in machine-to-machine applications for sensors, cloud applications, and mobile phones. The Internet protocols for machine-to-machine applications referred to as the IoT devices add to the complexity of home networks. The Internet technology in the home supports the complex mixture of office applications (email, web browsing, accounting programs) as well streaming services, sensor networks, firewalls for security, and data storage. During the third decade, the growth in

Cloud traffic also aided the growth of hyper-scale datacenters but increased concerns over user data privacy. The IETF chair and IESG members felt the changing IT landscape required the IESG to reorganize the IETF Areas to effectively match the IT industry opportunities checking and revitalizing all WGs.

During 2005 to 2012, the IESG cohorts under the leadership of the IETF Chair 5 and IETF Chair 6 encouraged the increasing the capabilities of the IETF (2016b) Datatracker database and online tools. The improved tools increased the ability of WGs and IESG members to handle more reviews. IETF Chair 5 and IETF Chair 6 worked to resolve conflicts within the IESG, resulting in lower conflicts in the IESG cohorts during 2005 to 2012 than during the IESG Cohorts during 2001 to 2004. The IESG cohort in 2013 to 2014 retained many members from the 2012 cohort, so the increased solidarity and reduced conflict aided decision making. The IESG cohorts from 2013 to 2015 undertook efforts to revitalize each area and reorganize the APP group of IETF areas into the ART and transport areas. During the IESG cohorts for 2014 to 2016, the U.S. government decided to transition the IANA (2020) registry to private funding, and the retirement of the IETF's administrative director (denoted IAOC director) caused the IETF chair and IESG cohorts in 2014 to 2016 to examine organization issues. From 2014 to 2016, IETF Chair 7 needed to focus on organizing the IETF's response to the U.S. government and the future of the IAOC administrative function rather than the IESG leadership. The extra hours during the reorganization efforts and the reduced support from the IETF chair saw growing conflict in 2014 to 2016. The researcher's evaluation of the approximately three decades of work (1989 to 2016) showed that external and internal pressures on the IESG cohorts impacted the level of group conflict and solidarity.

Unique Focus (Theme 5) and Change Resistance (Theme 6)

The IAB established the IESG as a TMT in July of 1989 during a restructuring the IAB into IAB, IETF, and Internet Research Task Force by the IETF Chair 1 (Gross, 1989, pp. 7–8) to allow for growth within the IETF. The IETF nominations committee selects both the IETF chair and IESG members (a portion of the IESG each year). The IETF nomination committees consider individuals for the position of IETF chair who have an immediate vision of how the IESG can help the IETF fulfill the IETF mission. Each IETF chair demonstrates by words, decisions, and deeds the technology standards (new and revisions) the IETF needs to publish to continue to have high quality and relevant specifications that influence the IT industry. The IETF chair has a vision for the IETF for the next 4 years to produce these standards. This vision may include what areas of new technology work, IETF area structures, ideas for new measurement systems and tools, and ways to encourage broad participation in the standard and its deployment. Each IETF chair leads each IESG cohort toward a vision for a 1 to 4 year time frame that accomplishes the IETF vision through the culture of the IETF chair's vision. Table 60 lists the technology and organizational focus for the IESG cohorts under each IETF chair, along with the key technologies standardized and organization changes made during the IETF chair's tenure. The qualitative analysis of the open-ended questions from the 2017 survey used the same per IETF analysis of technology and organizational focus to analyze the reported conflict during each IETF chair's tenure.

The researcher's Strand-5 analysis showed that each IETF chair accomplished most goals outlined in their visions for technology standards and organizational change while navigating the internal and external pressures to accomplish their goals. The open-ended questions on conflict reported increased conflict during the following IESG cohorts: 1994 to 1995, 2001 to 2004, 2005 to 2006, and 2015 to 2016. The perception of some members of the 1994 and 1995 IESG cohorts was that IETF Chair 2 failed to help create a productive inter-SDO relationship between the ISOC and the IETF. These members also felt IETF Chair 2 allowed delays in review and conflict over decisions for publishing some standards, so some IESG members used this to delay the publication of some standards. Members of the IESG cohorts from 2001 to 2004 IESG cohorts reported their perceptions that IETF Chair 4 allowed certain IESG members to claim expertise that the IESG member did not have. In addition, this culture within the IESG caused conflicts over the IETF resolutions to IETF-ITU technical conflicts caused RCs. Inter-SDO policies of the IETF govern the IETF-ISOC legal relationship and IETF-ITU technical relationship. The conflict during this period may be due to resistance to change these policies. Arrogance in an individual was one defense against technology change so rapid that individuals might lose the ability completely understand it. One question was how much of this conflict was related to IESG members resisting change in the IETF.

Members of 2005 to 2006 IESG cohorts reported their perception that the IESG members disagreed on Internet level threats, and IESG members from different areas had strong opinions, and they clashed frequently. As discussed in the SWOT analysis, exterior market influences caused a portion of this dichotomy by pulling application areas in different directions than network and management areas. These members reported that conflict simmered below “a polite surface,” and conflict avoidance delayed consensus decisions. IETF chair reported struggling with conflicts between IESG members from different areas to try to get members to resolve conflicts. The 2005 to 2006 conflict arose from pressures from exterior forces for publications of specifications to meet deadlines and an unwillingness to engage in consensus decision-making for all parties.

The 2015 to 2016 conflicts came from a leadership vacuum caused by the unexpected requirement for the IETF chair to focus on the IANA (2020) transition and IAOC retirement immediately after a reorganization of the IETF areas. Conflicts within the IESG increased during the 2015 to 2016 IESG cohorts as members tried to step into the leadership vacuum by over-stepping their roles to take on the functions of an IETF chair. In addition, the burnout from the reorganization caused some members to refuse to take on tasks and have an unwillingness to engage in the lengthy discussions needed for consensus. IESG members reported that TC from the leadership vacuum combined with burnout led to RCs that impacted long-established relationships.

The analysis showed that the perception of conflict during an IESG cohort linked to change resistance within the IESG or a perception of a leadership vacuum. The comments on conflicts from IESG cohorts in 1994 to 1995 arose from a perception that IETF Chair 2 did not establish the IETF-ISOC SDO relationship or manage the culture correctly. However, the IESG cohorts under IETF Chair 2 made effective decisions (decisions with measurable results) in 76% of the decisions, according to the 10% IPA analysis. These IESG cohorts had a higher percentage of effective decisions than the IESG cohorts under the IETF Chair 4 (2001 to 2004), IETF Chair 5 (2005 to 2006), and IETF Chair 6 (2007 to 2008). The cohorts under ranges between 51% to 55% of the decisions considered these four chairs (IETF Chair 2, IETF Chair 4, IETF Chair 5, and IETF Chair 6). The conflicts in the IESG cohort during 2014 to 2016 seemed to be caused by burnout, changes, and leadership vacuums with cumulative impacts of the conflict. Given this background, the HRM modeling for the 100% sample from 2016 (Dataset 2) makes sense. The IESG appeared to respond to an IETF chair exhibiting (or failing to exhibit) the leadership behaviors Yukl (2010) recommended to change a group's culture. The IETF chair's vision, values, behaviors, reactions to crisis, and attention to the group and member needs influenced the perceptions of the IESG.

The leadership of a virtual TMT takes a great deal of time and effort until the group finishes its phases of norming, storming, and forming before it performs effectively. Some IETF chairs reported spending time and effort on group dynamics. However, other IETF chairs reported that other factors pulled them away from such efforts to improve IESG group dynamics. Because the IETF chair's position represented a volunteer job, how many of the IETF chairs, such as the IETF chair during 2014 to 2016, did not have enough time to lead the IESG effectively? The hidden cost was that the IETF's chair's lack of time for leadership led to conflict within the IESG, which qualitatively led to less effective decision-making.

Conclusion of Qualitative Analysis

Based on the above qualitative analysis data and the relationships in the reduced theoretical model, the qualitative analysis finds the quantitative results are correct. Table 61 summarizes the quantitative analysis from Strands-1–4 and the quantitative analysis from Strands-1–5 per IETF chair. The qualitative results aligned with quantitative results and explained why the HRM modeling tests for data from the 100% sample in 2016 explained less of the variance in results than the 100% sample in 2015. Based on the Strand-5 qualitative analysis, both solidarity and conflict correlated to effective results. Solidarity had a positive effect on consensus decision making and conflict has a negative correlation. Solidarity predicted the effectiveness of consensus decision-making in the IESG TMT for actual and perceived results in historical records, survey results, and qualitative data. OCB predicted results in the historical record due to the nature of the IESG minutes to record events tightly bound to required or recommended citizenship behaviors. The OCB behavioral scores from the survey only aligned with the only perceived results rather than actual results. Therefore, solidarity was a better predictor of the effectiveness of consensus decision-making than OCB.

The quantitative analysis found conflict challenging to measure in an open standards SDO staffed by volunteers because RC is socially unacceptable. Because it was socially unacceptable, the minute-takers actively dampened (via a “pens-down” action) incidents of RC in historical records and survey responses. The quantitative results showed that conflict existed in both IESG interactions as expressed conflict and hidden conflicts expressed in delaying or refusing to engage in consensus decision making. Conflict also might have resulted from a leadership vacuum and resistance to change. Therefore, the quantitative and qualitative results both supported Hypothesis 1 of the reduced model that an increase in solidarity would increase the effectiveness of the consensus decisions made in consensus decision-making and did not support Hypothesis 2.

Hypothesis 2 in the reduced model posited that conflict had a moderating effect on solidarity increase in the effectiveness of consensus decision-making in team consensus decision making. Conflict scores on the survey negatively correlated with results because the survey instrument considered TC and RC. In contrast, the conflict theme counts from the historical records had a positive correlation in the IPA analysis of the historical record where expression of RC is damped. The qualitative analysis also indicated that conflict was caused by other factors, such as external and internal pressures, the focus for technology and organizational change set by an IETF chair, and the pressures of change resistance in the IETF. The IESG must handle continual change as a TMT of the IETF, whose organizational mission is to “make the Internet work better” (IETF, 2021, para. 2).

The IETF operated in the IT industry as a change agent for Internet technology, and the IESG led these efforts to change Internet technology. Some of these changes might include technology refinements (e.g., enhancements to email or an existing routing protocol; single-loop learning). Other technology changes might change the paradigms of Internet technology (double-loop learning). Examples of paradigms changes included how SIP changed phone technology or the data-driven paradigm of Yang in network management. The IESG, as change leaders, need to inspire WGs to learn how to refine protocols (single-loop learning or adaptive learning), create paradigms for new protocols (double-loop learning or generative learning), and teach new leaders how to learn to do the same (dueterolearning). Conflict is complex behavior in the IESG cohorts because of its function as a virtual TMT of the IETF—an IT SDO whose mission calls it to act as a change agent.

Table 59: Strand-5 – Internet Engineering Task Force SWOT Analysis for Internet Engineering Steering Group Cohorts (per Internet Engineering Task Force Chair Period)

IETF Chair Period		IESG Cohorts	IETF SWOT Analysis			
#	Identifier		Strengths	Weakness	Opportunities	Threats
1	Chair 1	1989 to 1993	1) NSF funded IETF to be SDO for NSFnet, 2) Internet growth, and 3) IANA (2020) registry for IETF protocols supported by U.S. government	1) IETF started without organization documents, so organization documents created 2) DNS operators of root servers small, trusted community	1) Commercial traffic allowed by NSF under ANS network creation	1) Existing SDOs for Networking (ITU, IEEE) 2) US GOSIP support for OSI instead of TCP/IP 3) Vendors have IT technology with OSI, SNA or DECNET
2	Chair 2	1994 to 1995	1) commercial vendors, regional networks with commercial traffic 2) stable network protocols and key applications (email and web technology)	1) no existing relationship with IEEE and ITU, and 2) Lacks legal entity to hold IPR or answer IPR claims	1) W3C (www) based on TCP/IP technologies 2) Early adopters TCP/IP mail and websites were financially successful	1) Transition of the Internet from NSF funded to commercial funding 2) Internal conflicts over IPng 3) IETF-ITU SDO level conflicts
3	Chair 3	1996 to 2000	1) IP over mobile networks 2) Virtual networks (SMDS, early MPLS, IP-VPNs) 3) Voice over IP concepts	1) IETF lacks a relationship with ITU 2) Lack of security impacts commercial deployment 3) IPv6 slow deployment creates the need for NAT boxes	1) dot.com websites increased demand for IP technology 2) Ethernet and Wi-Fi technologies aided Internet Group	1) dot.com websites overheated demand for Internet 2) ITU SDO conflicts cause internal IESG conflicts

IETF Chair Period		IESG Cohorts	IETF SWOT Analysis			
#	Identifier		Strengths	Weakness	Opportunities	Threats
4	Chair 4	2001 to 2004	1) security for TCP/IP, mail, HTTP, XML deployed 2) Virtual networks technology deployed (MPLS, L2VPN, L3VPN) 3) Voice over IP (SIP) technology 4) CDN technology	1) NM hits problems with SNMP + MIBs and starts NM work with Netconf + Yang modules 2) IETF tools need an upgrade to support a growing number of WGs and standards 3) Specification review light during WG process in 1996 to 2001, so efforts needed to improve review	1) 3G mobile technology uses IP Technology 2) VOIP in Skype connects to PSTNs 3) Secure payments deployed (PayPal) 4) Technologies below IP (ATM, cable, Bluetooth) 5) Growth of Google and Amazon Cloud services	1) dot.com bubble burst causes the Internet market to slow down 2) Carrier networks consolidation further reduces the market for Internet vendors 3) Reduced needs for IPv6 technology except in Asia
5	Chair 5	2005 to 2006	1) Existing mail, HTTP, video, VOIP allowed revisions 2) Existing IP multicast allows mobile revisions 3) RAI IETF area consolidated work on real-time video 4) IPv6 for Asian markets	1) Focus for 3G work leads to IESG conflicts on what constitutes "internet level threats" and publication delays. 2) NM must handle develop new protocols (NETCONF with Yang data models) and maintain SNMP with MIB modules 3) Scaling for VPN overlay and underlay technology	1) renewed growth in Internet Market 2) growth in real-time video applications (web and mobile) 3) YouTube streaming service 4) growth in the cloud data center market	1) ITU and IETF SDO disagree on MPLS 2) Delays in IETF protocols could cause the inclusion of fewer Internet features in 4G phone networks
6	Chair 6	2007 to 2012	1) Growth in RAI technology (VOIP, Video) 2) Growth in security in application, network, and management	3) WG documents need detailed IESG reviews 1) NM must handle and develop new protocols (NETCONF + Yang data models) plus maintain SNMP + MIBs	1) Netflix starts stream service 2) 4G deployments 3) growth in sensor networks and new machine	1) 2008 stock market crash 2) Push for internationalization of applications

IETF Chair Period		IESG Cohorts	IETF SWOT Analysis			
#	Identifier		Strengths	Weakness	Opportunities	Threats
			3) NM NETCONF and Yang modules standards 4) IETF tooling improves WG document and review tracking	2) IPv6 deployment in North America minimal 3) Some WG documents lack sufficient reviews	to machine (IoT) applications 4) growth in home network technology 5) growth in cloud data center market with hyper-scale data centers	3) ITU-IETF disagreements increase 4) Datacenters need new protocol configuration features
7	Chair 7	2013 to 2017	2013 to 2015 1) machine to machine (IoT) application support protocols (e.g., CBOR) 2) Ability to revise application, network, management, and security protocols 2015 to 2016 1) Reorganization of IETF areas to create ART area, revitalization of all areas	2013 to 2015 1) IETF Areas need revitalizing and reorganization for the third decade of IETF 2014 to 2016 1) increased stress and conflict in IESG relationships due to the IETF chair's focus on IANA (2020) transition 2) Concerns regarding supporting applications with a short product life cycle	1) Continued Growth in IoT network technology 2) research for 5G networks that include increased mobile apps and senor networks 3) Growth in Cloud data center market growth with hyper data centers 4) growth in home networking complexity	1) pervasive monitoring of IT protocols causes concern over the privacy of user data 2) Transition of IANA (2020) from U.S. government to private funding 3) IAOC retirement causes a call for reorganization of administrative functions

Table 60: Strand-5 – Focus and Accomplishments for Internet Engineering Steering Group Cohorts (by Internet Engineering Task Force Chair Period)

IETF Chair Period		IESG Cohorts	Focus		Key Accomplishments	
#	Name		Technology	Organization	Technology	Organization
1	Chair 1	1989 to 1993	1) Routing and NM protocols 2) NSFNet goes from OSI+IP to IP-only 3) Transport Area creation for multi-media transport	1) Creation of IESG 2) IETF organization documents (process, WG, nomcom)	1. Stable TCP/IP and routing protocols 2. Operational requirements for hosts and routers 3. Transport area creation	1. Launch IETF and established as open-standard organization 2. Created IESG 3. IETF organizational documents
2	Chair 2	1994 to 1995	1. Web technology 2. Commercial growth of IP protocols 3. Revision of IP Protocol (IPv6)	1. ISOC is formed to be the legal entity for IETF 2. Revisions to IETF organizational documents	1. W3C technology 2. Scaling protocols for Commercial growth 3. Choosing IPv6 as IPng protocol	1. Legal protection for IETF in ISOC 2. revision to IETF organizational documents
3	Chair 3	1996 to 2000	1. enhancements for email, web technology, and internet databases 2. online trading applications 3. network access policy and security 4. IP for mobile networks (phones) [mobile-ip] 5. Virtual Networks (VPN) technology 6. Security and management 7. SIP for Voice over IP (VOIP)	1. Revision of IETF organizational documents 2. ITU relationship	1. good progression of standards for applications, network, and management areas 2. online trading applications, 3. IP network technology for mobile (phone) networks 4. MPLS technology (controlled flows) 5. Initial SIP protocol	1. Revision of IETF organizational document 2. handled tremendous growth in participants in the IETF standards process (meetings and mail lists)
4	Chair 4	2001 to 2004	1. Content Delivery Networks (CDN) 2. VoIP enhancements (SIP) 3. IP mobility additions	1. Quantifying perceived problems in the IETF process and revisions to Intellectual Property Rights	1. New technologies aided adoption of IP technology by 3G phone networks,	1. Worked with the IETF community to establish what organizational process

IETF Chair Period		IESG	Focus		Key Accomplishments	
#	Name	Cohorts	Technology	Organization	Technology	Organization
			4. Data Center additions 5. new NM protocols 6. security additions 7. Transport aids for video streaming 8. auto-configuration 9. internationalization of names	(IPR) policy and leaders selection policy (nomcom) 3. IETF policy for early IANA (2020) allocations	2. Voice over IP (SIP) technology 3. Transport layers to aid video streaming, 4. Selection of new NM protocol (NETCONF) 5. Good progression of standards for applications, network, and management areas 6. Internationalization of names	documents needed updating. 2. Revised IETF standards for IPR and leaders selection methods (nomcom) 3. Early IANA (2020) allocation of protocol numbers to aid implementations before standardization
5	Chair 5	2005 to 2006	1. Revisions to existing applications for email, VOIP, and video applications 2. Revision to overlay and underlay VPN technology 3. Mobile IP multicast (e.g., video conference for cell phones). 4. Revisions to security protocols 5. Continued work on internationalization of names	1. Creation of RAI area 2. Specified IETF Data tracker tools for WGs 3. Continued discussions on IPR issues for IETF	1. Good progression of standards for application, network, management, and security areas 2. Technology enables VOIP, video streaming, and video conferences over IP on cell phones. 3GPP (2011) adoption of this technology. 3. Continued work on internationalization	1. Created RAI area to foster 3GPP (2011) and web technologies using VOIP and Video 2. Continual improvement on IETF (2016b) Datatracker tools 3. Continued to handle questions on IPR
6	Chair 6	2007 to 2012	1. Revisions to existing applications for email, VOIP, and video applications 2. Revision to overlay and underlay VPN technology.	1. changes to IETF data tracker, 2. Changes to RFC formats 3. Continued discussions on IPR issues for IETF	1. Security technology progression reaches into all areas of IETF 2. Continued progression on RAI area technology (VOIP, video streaming,	1. Good consensus process established in the IETF that fostered high quality, relevant documents 3. Continual improvement on IETF (2016b) Datatracker tools

IETF Chair Period		IESG	Focus		Key Accomplishments	
#	Name	Cohorts	Technology	Organization	Technology	Organization
			3. Multicast data transmission for mobile IP (e.g., phones) 4. Revisions to security protocols 5. Continued work on internationalization of names 6. Authority to citizen messages 7. NM management changes		and video conferences over IP on cell phones) 3. Authority to Citizen messages over IP	4. Continued to handle questions on IPR
7	Chair 7	2013 to 2017	1. Continued revisions to applications, network, security, and management protocols to protocols for web, email, video (streaming and conferences) for data centers and webrtc deployments 2. Internet of things support in application and network 3. Work to minimize pervasive monitoring in protocols	1. Reorganization of IETF areas and revitalization within areas 2. Response to U.S. government on the transition of IANA (2020) 3. Meeting location selection 4. anti-harassment policy	1. Continued progression of existing work to keep improving protocols for application, network, security, and management standards for the changing Internet IT landscape, 2. Added work for the Internet of things (IoT) technology to the IETF suite of protocols 3. Reviewed protocols in all IETF technology areas to reduce protocol monitoring that invaded privacy.	1. Reorganization and revitalization of WGs in all IETF areas. 2. Creation of ART area and mechanisms within IETF areas to speed up approval of work. 3. IETF community gave input on IANA (2020) translation. 4. meeting location selection policy created. 5. Antiharassment policy established.

Note. NM – network management, OSI+IP – OSI 7-layer protocols + TCP/IP protocols, WG – Working group, Nomcom – IETF nomination committee that selects IETF leaders.

Table 61: Strand-5 – Qualitative Evaluation of Reduced Theoretical Model (by Internet Engineering Task Force Chair Tenure)

IETF Chair Period		IESG cohorts	Solidarity	Conflict			Effective Decisions
			5 Quantitative datasets + qualitative	5 Quantitative datasets	2017 Survey Open-ended Questions		
#	Name				Task (TC)	Relationship (RC)	Results
1	Chair 1	89to 93 89to 90	Qualitative: 1989 to 1990: strong solidarity. 1991 to 1993: mixed solidarity Quantitative: Surveys: 5.37 (2017), 5.43 (2013) Somewhat agree	Qualitative: 89 to 90: lower conflict 91 to 93: higher conflict Quantitative: Surveys: 2.39 (2017) n/a (2013) (disagree)	1) Routing and NM technology in early years, 2) Need for IP protocol revision (IPng) 3) Role of IAB versus IESG 4) autonomy of AD in Area	1) AD protective of "turf," 2) Some ADs are unwilling to compromise or negotiate during the consensus process, 3) Factions in IESG conflict with the IETF chair and each other,	Qualitative: good 1) IETF established with organizational documents, 2) TCP/IP protocols stable 3) IESG was created to help manage IETF
1	Chair 1	91 to 93	Quantitative: Cohort mean Surveys: 5.24 (2017) 5.43 (2013) Somewhat agree IPA (10%): 32% of IBA	Quantitative: Cohort mean Surveys: 3.01 (2017) n/a (2013) Somewhat disagree IPA (10%): 11% of IBA			Quantitative: Effective decisions as a percentage of total decisions: 91-93: 60% 91-93: Average decisions per year: 297
2	Chair 2	1994 to 1995	Qualitative: Medium solidarity and factions Quantitative: Surveys: 5.01 (2017) 5.44 (2013)	Qualitative: higher conflict Quantitative Surveys: 3.19 (2017) n/a (2013) Somewhat disagree	1. Conflict over different priorities for documents on technology change 2. Conflict over ISOC creation and IETF-ISOC relationship	1. indirect means of delays during the consensus process 2. IETF chair failed to help SDO relationship formation	Qualitative: good consensus supported standards for W3C and commercial Internet growth supported, next to IP version selected (IPv6), ISOC-IETF relationship formed.

IETF Chair Period		IESG cohorts	Solidarity		Conflict		Effective Decisions
			5 Quantitative datasets + qualitative	5 Quantitative datasets	2017 Survey Open-ended Questions		
#	Name				Task (TC)	Relationship (RC)	Results
			Somewhat agree IPA (10%): 31% of IBA	IPA (10%): 7% of IBA			Quantitative: Effective decisions as a percentage of total decisions: 94-95: 76% Average decisions per year: 480 Qualitative: strong enabling tremendous growth
3	Chair 3	1996 to 2000	Qualitative: Medium to good Quantitative: Surveys: 5.53 (2017) 5.71 (2013) Somewhat agree IPA (10%): 31% of IBA	Qualitative: medium Quantitative: Surveys: 2.93 (2017) n/a (2013) Somewhat disagree IPA (10%): 6% of IBA	1. Technology conflict on IP allocation guidelines 2. Technology boundaries in ITU/IETF relationship	1. Relationship conflicts over revision to IP allocation guidelines 2. Relationship conflicts IETF-ITU conflicts	Quantitative: Effective decisions as a percentage of total decisions: 1996 to 2000: 87% average decisions per year: 460
4	Chair 4	2001 to 2004	Qualitative: Medium to good solidarity Quantitative: Surveys: 4.73 (2017) 5.69 (2013) Somewhat agree IPA (10%): 28% of IBA	Qualitative: conflict higher than normal Quantitative: Surveys: 3.43 (2017) n/a (2013) Somewhat disagree to uncertain IPA (10%): 8% of IBA	1. Technology boundaries in ITU/IETF relationship	1. IETF-ITU relationship got tense, and different IESG members reacted to conflict in changing the IETF-ITU relationship 2. Arrogance in ADs assuming expertise they did not have during reviews	Qualitative: medium Areas of consensus: 1. VOIP and video stream for 3 G networks 2. CDN networks 3. IETF process updates Quantitative: Effective decisions as a percentage of total decisions: 2001-2004: 55%

IETF Chair Period		IESG cohorts	Solidarity	Conflict			Effective Decisions
			5 Quantitative datasets + qualitative	5 Quantitative datasets	2017 Survey Open-ended Questions		
#	Name				Task (TC)	Relationship (RC)	Results
							Average decisions per year:768
5	Chair 5	2005 to 2006	Qualitative: good solidarity Quantitative: Surveys: 5.91 (2017) 5.66 (2013) Somewhat agree IPA (10%): 27% of IBA	Qualitative: medium with some hidden conflict Quantitative: Surveys: 2.83 (2017) n/a (2013) Somewhat disagree IPA (10%): 9% of IBA	1) Conflict avoidance caused the consensus decision process to stall on some decisions. 2) IESG members disagreed on what constituted Internet-level threats.	1) Interpersonal conflicts simmered below a polite surface, 2) Conflicts over the working style of IESG and the "turf" of some IESG members, 3) IESG members had strong opinions, and ADs from different areas clashed frequently	Qualitative: medium with strong decisions to: 1) Create RAI Area 2) Progress NM Quantitative: Effective decisions as a percentage of total decisions: 2005 to 2006: 55% Average decisions per year: 865
6	Chair 6	2007 to 2012	Qualitative: Strong Solidarity Quantitative: Surveys: 5.96 (2017) 5.87 (2013) Agree IPA (10%): 31% of IBA	Qualitative: medium with some hidden conflict Quantitative: Surveys: 2.80 (2017) n/a (2013) Somewhat disagree IPA (10%): 12% of IBA	1. Conflict avoidance caused the consensus decision process to stall on some decisions. 2. IESG members disagreed on what constituted Internet-level threats. 3. Technology boundaries in ITU/IETF relationship	Early years: replicated the conflicts from 2005 to 2006 Later years: Discussions in later years had strong views, but the discussions were open and collaborative.	Qualitative: medium to good Focus increasing quality of IETF standards and good consensus decisions for an increased number of standards Quantitative: Effective decisions as a percentage of total decisions: 2007 to 2012: 51% Average decisions per year: 960

IETF Chair Period		IESG cohorts	Solidarity	Conflict			Effective Decisions
			5 Quantitative datasets + qualitative	5 Quantitative datasets	2017 Survey Open-ended Questions		
#	Name				Task (TC)	Relationship (RC)	Results
7	Chair 7	2013 to 2017	Qualitative: medium 2013-2014 solidarity better than 2015 to 2016 solidarity Quantitative: Surveys: 5.14 (2017) 5.73 (2013) Agree IPA (10%): 34% of IBA	Qualitative: Increasing conflict as efforts in 2013 to 2014 caused burnout and 2015 to 2016 caused leadership gap due to IETF chair's focus on IANA (2020) transition Quantitative: Surveys: 3.14 (2017) n/a (2013) Somewhat disagree IPA (10%): 6% of IBA	1. Members of the group attempted to overstep their roles and take on IETF chair functions. 2. Simple decisions could suddenly and unexpectedly lead to conflict. 3. Unwillingness to take on additional tasks. 4. Unwillingness to engage in consensus discussions for specific topics (e.g., the special use names for another SDO)	1. IESG members attempted to micro-manage other members. 2. ADs refused to engage in consensus decisions retaining conflicts. These behaviors led to personal relationship conflicts that remained unresolved over time. 3. Unresolved tasks left a lasting impact on interpersonal resolutions.	Qualitative: good to medium IESG cohorts for 2013 and 2014 made better decisions than 2015 to 2016 (medium) Quantitative: Effective decisions as a percentage of total decisions: 2007 to 2012: 66% Average decisions per year: 785

Figure 35: Changes in information technology networks from 1975 to 2021

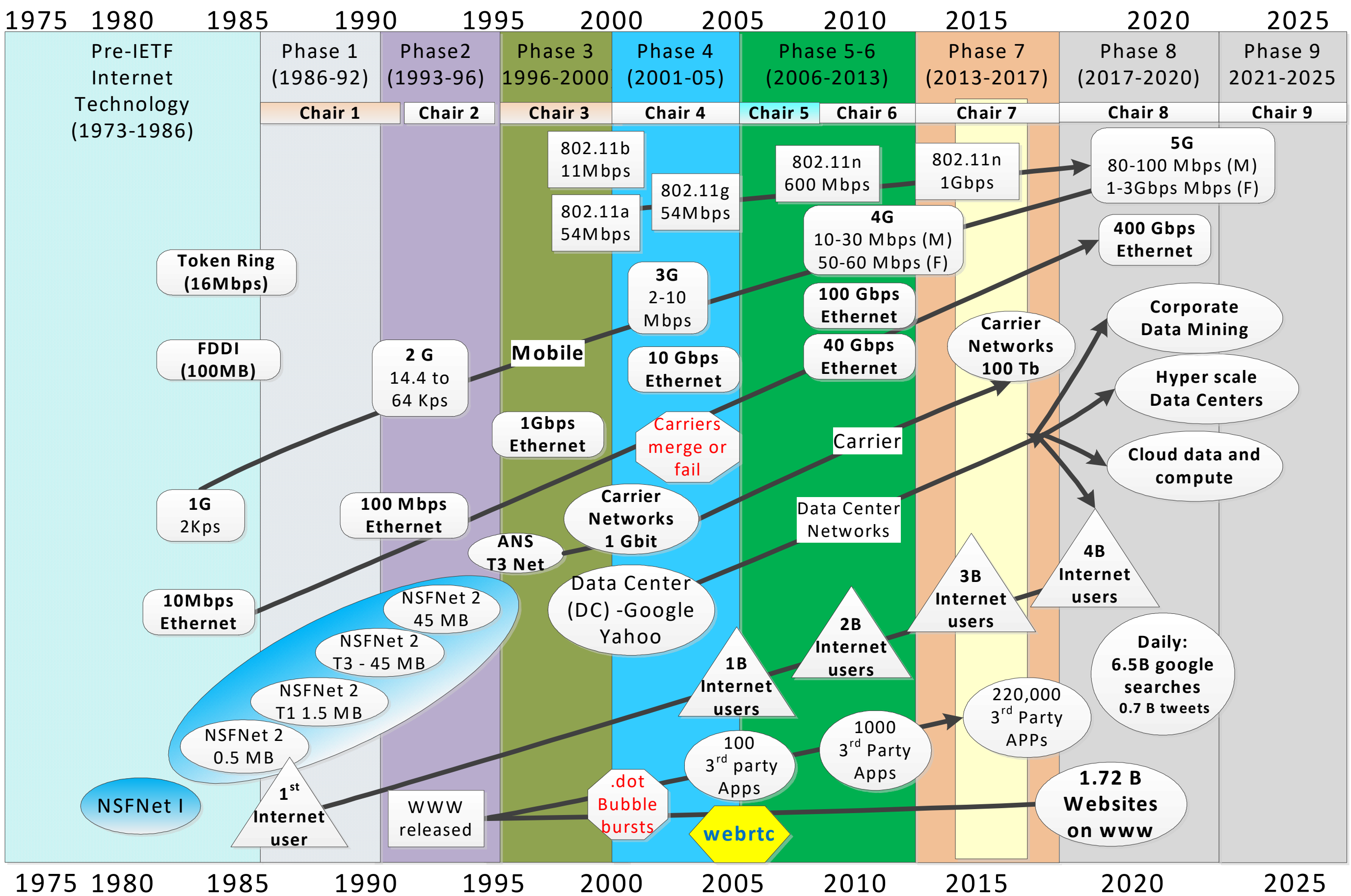


Figure 36: Network application from 1970 to 2021 enabled by Internet Engineering Task Force protocols (1985 to 2020).

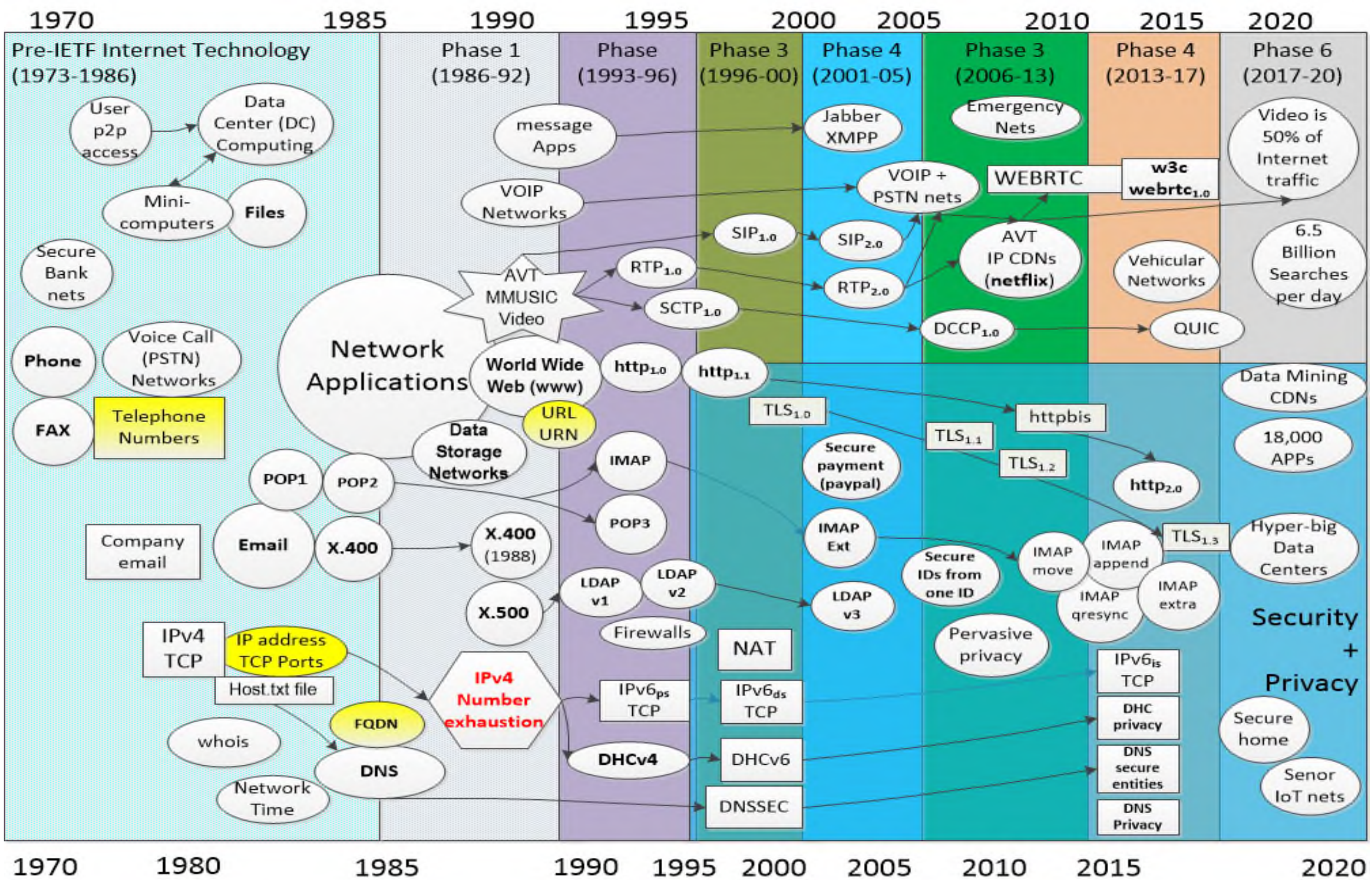


Figure 37: Internet protocol network protocols.

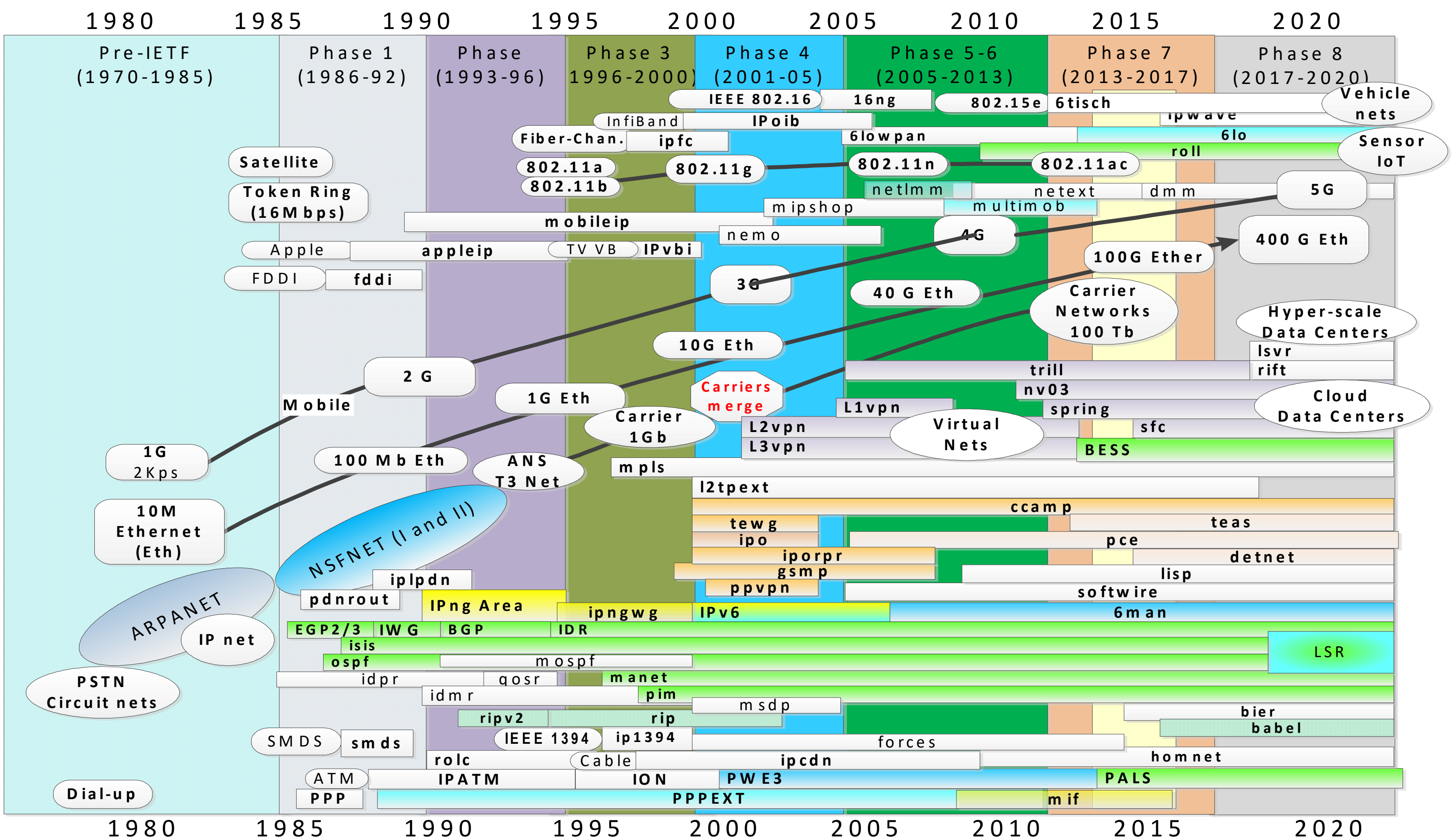


Figure 38: Strand-5 – Internet protocol network protocols plus operations working groups.

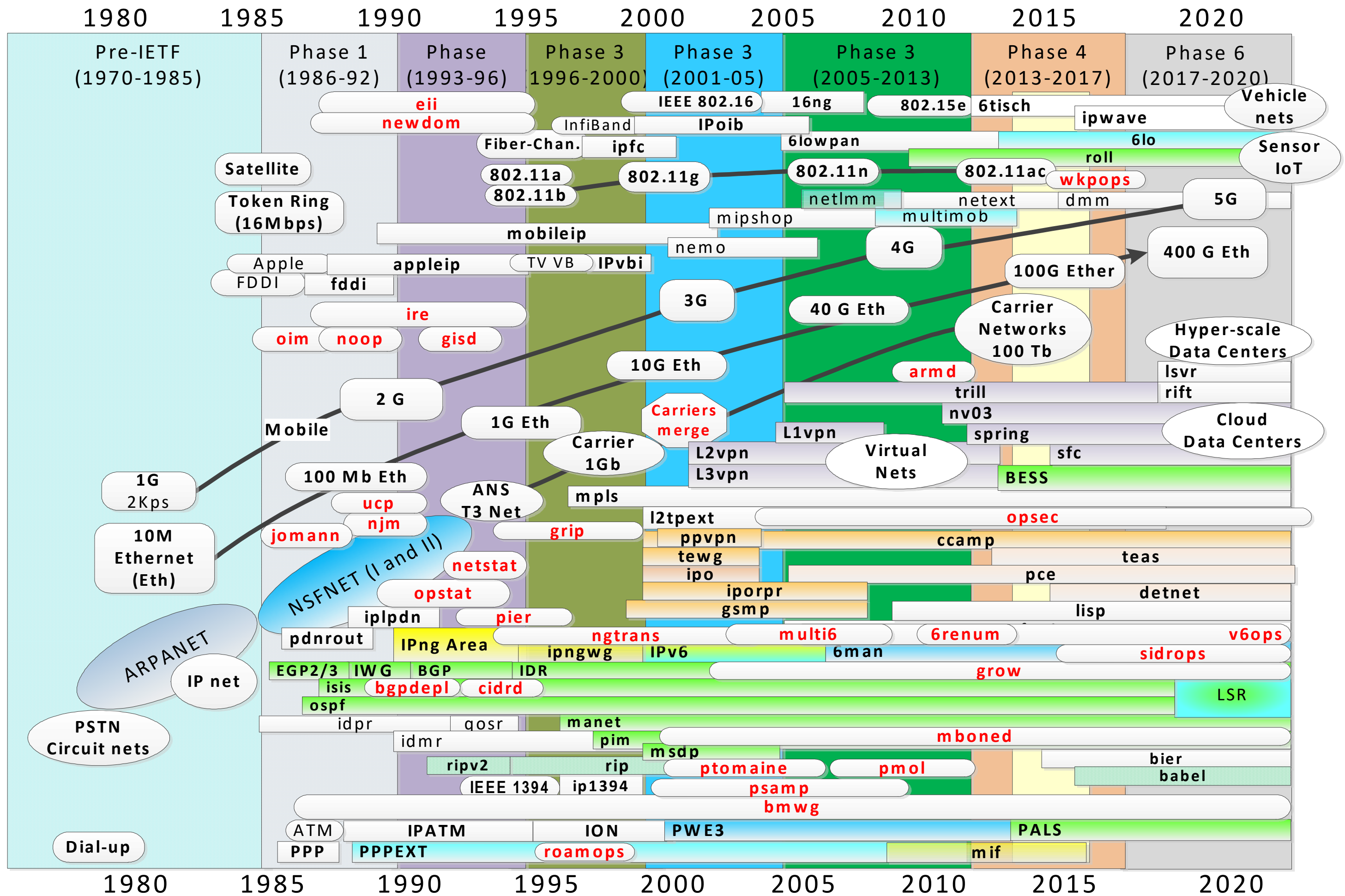


Figure 39: Strand-5 – Network protocols plus management protocols.

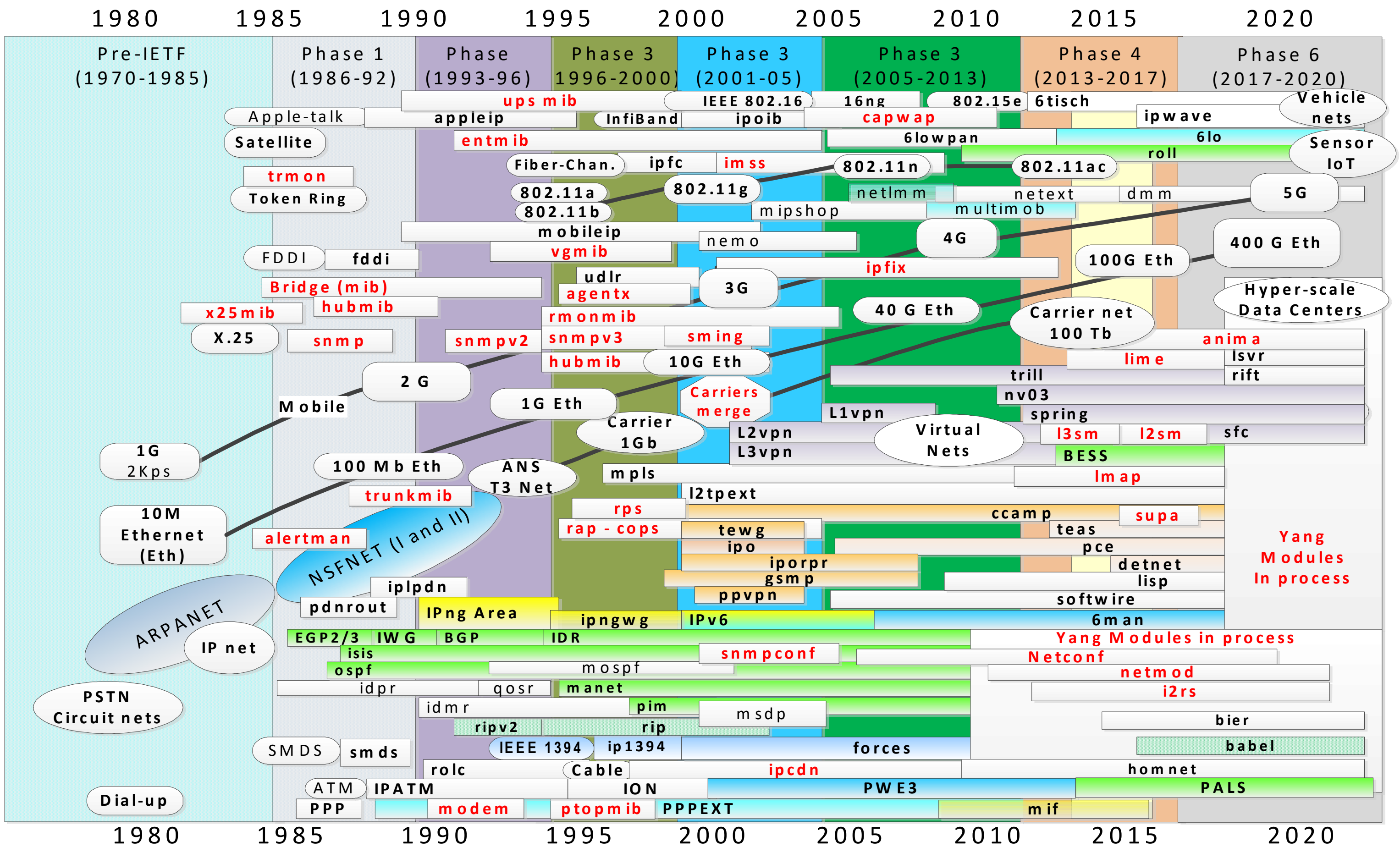


Figure 40: Strand-5 – Security additions to Internet Engineering Task Force application protocols for authentication, security, and privacy

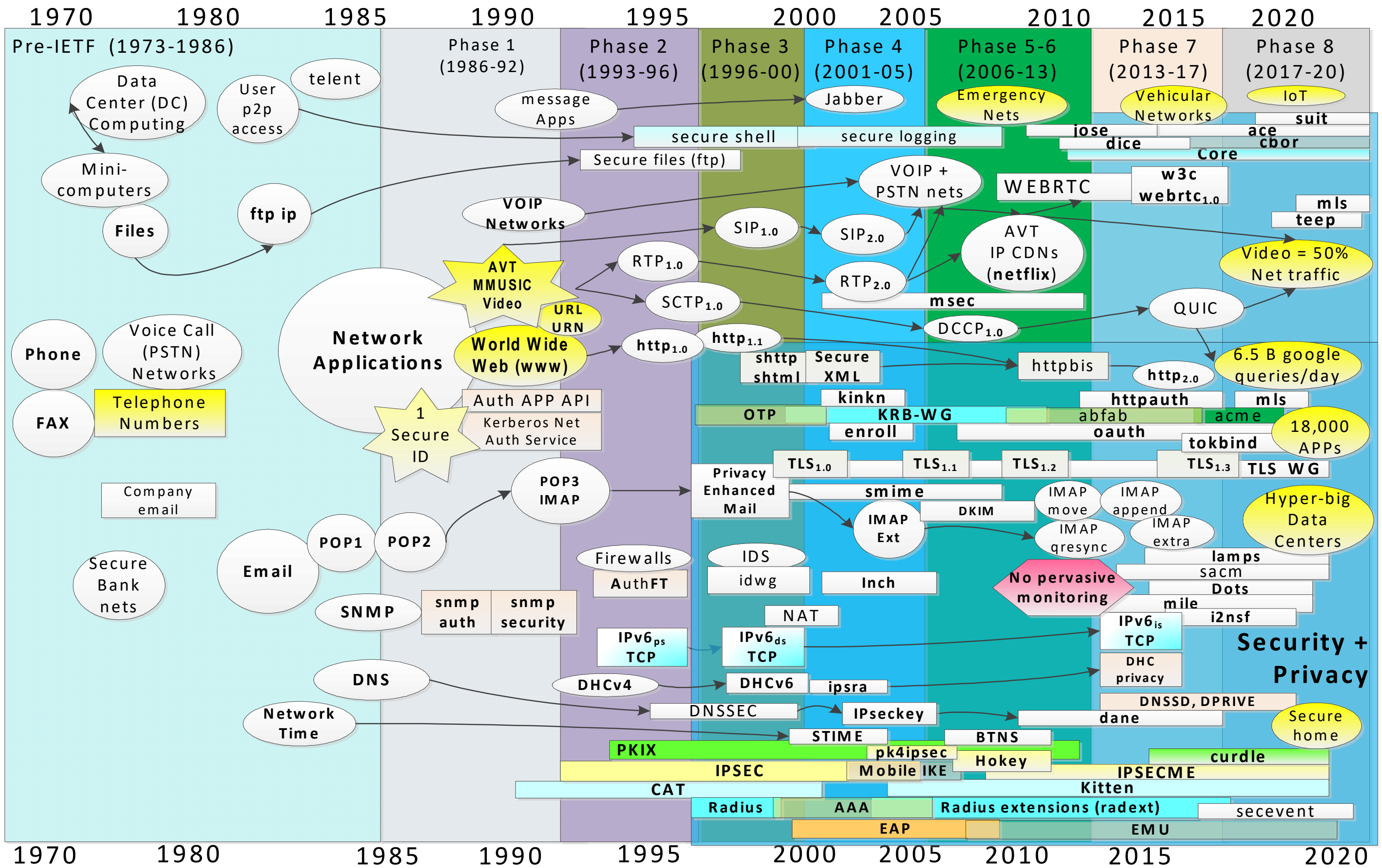


Figure 41: Strand-5 – Internet protocol network progression with security additions.

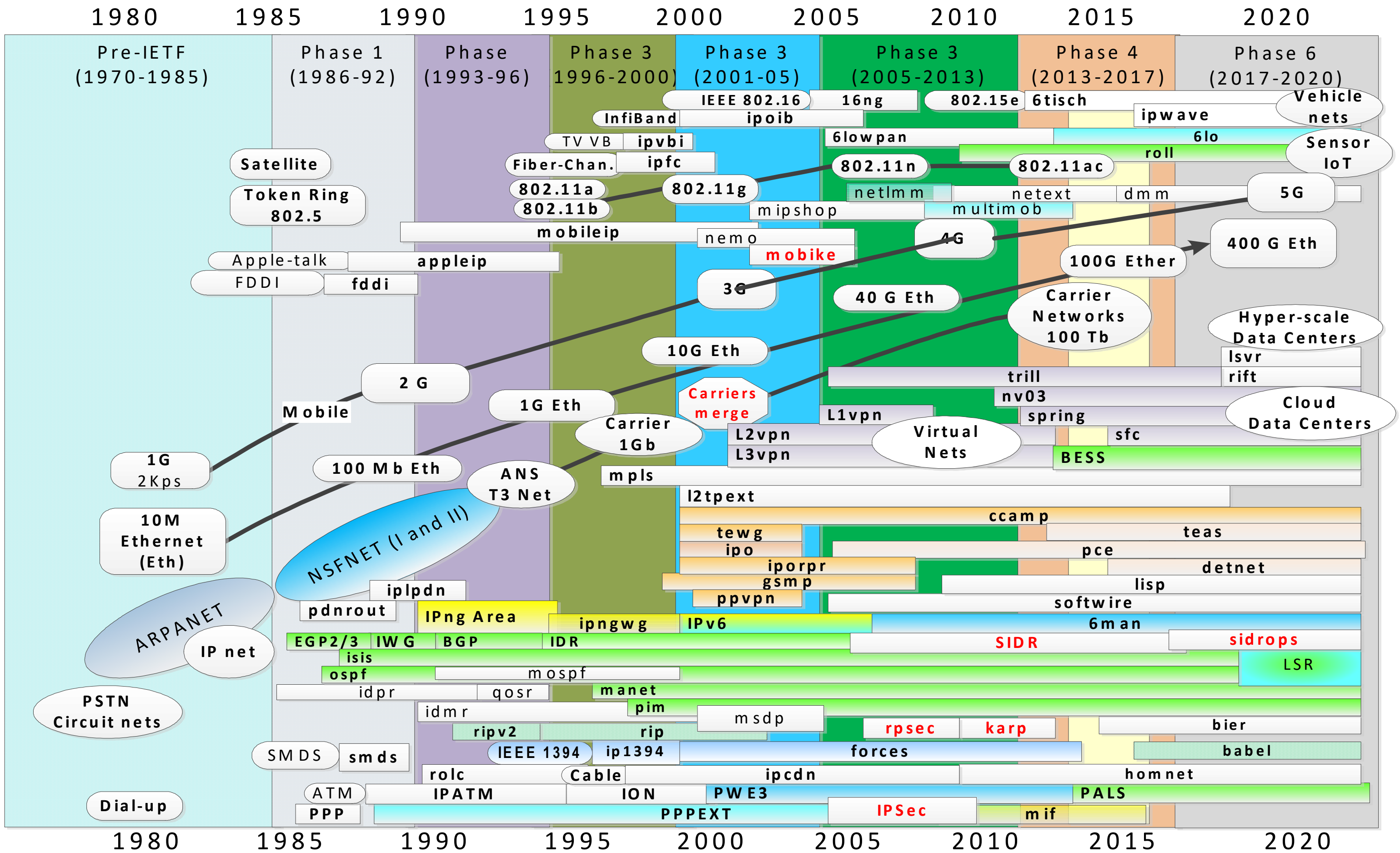


Figure 42: Strand-5: Organization standards and working groups from 1986 to 2020.

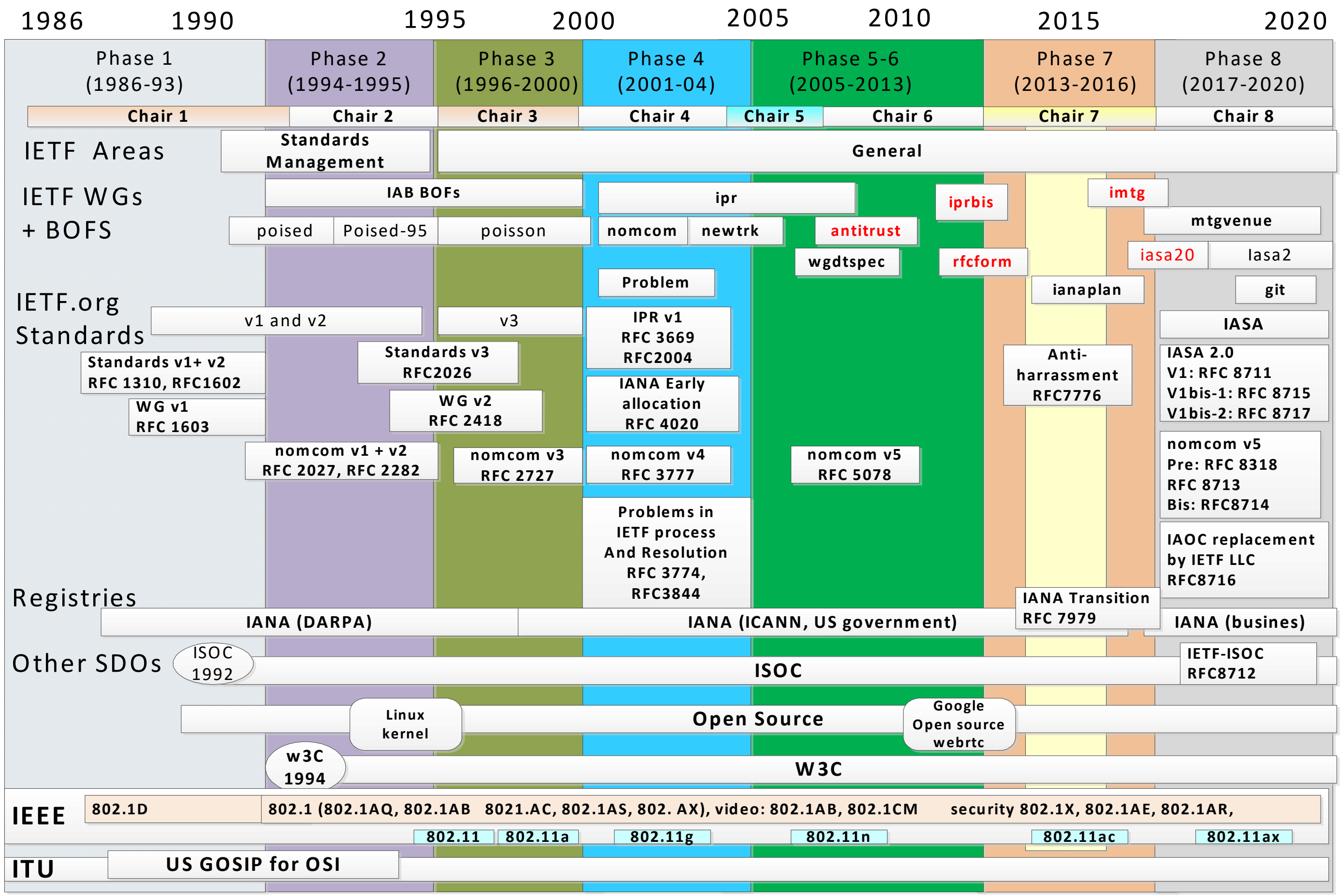
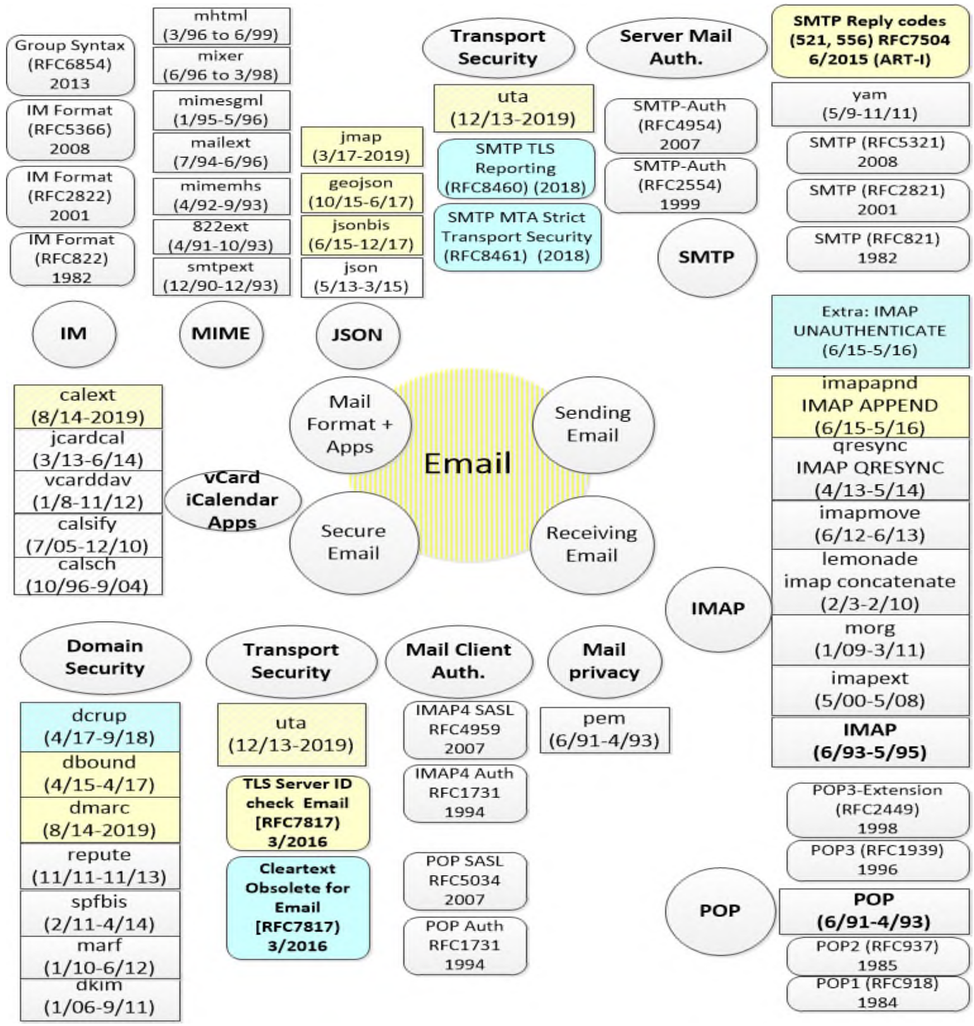


Figure 43: Migration of email protocols (1989 to 2020)



Chapter 5 – Discussion

The quantity of quality data mattered in each of the five strands of this 28 year (1989 to 2016) longitudinal mixed-mode study of a TMT that used virtual consensus decision-making. In addition, the quantity of quality data mattered during data collection, analysis, and triangulation to the validity, reliability, and real-world credibility of this research's conclusions. This longitudinal mixed-mode research collected qualitative data from various online sources guided by historiometric best practices, survey data from two separate surveys, and received individual feedback via email and in-person discussions. During the initial creation of the alternate methodology, this researcher collected qualitative and quantitative data from online sources and developed the specific encoding practices using three inter-raters (primary researcher and two other experts in behavioral analysis). Research articles on the best practices in mixed-mode methodologies underscored the need for quality data collection and analysis. However, the quantity of data was critical because the data analysis acted as a filter to remove invalid or unreliable data.

Data analysis filters act like noise filters on a cell phone call. Cell phones use electronic hardware and software to separate noise and interfaces from the valid radio signal to the cell phone to create a clear sound even when thunderstorms or electronic interference impact radio signals. The communication theory (signal-to-interference plus noise [SINR] behind cell phone filters requires a high enough level of cell phone signal for filtering algorithms to work to reduce noise (Haenggi, Andrews, Baccelli, Dousse, & Franceschetti, 2009).[11] Cell phones display how much signal level, which aids customer acceptance of more noise due to low signal rate. Like the cell phone, if the quantity of quality historical or survey data is insufficient, data analysis has gaps like the dead zones in cell phone reception. This chapter discusses the lessons learned about methods to get enough quality historical and survey data regarding a TMT in an open standards SDO operating with a

---

[11] The SINR ratio for a point x in space is the following: SINR(x) = P/ (I+N), where P is the power of the signals, I is the interference, and N is the noise (expressed in decibels [db]).

mission to “make the Internet work better by producing high quality, relevant standards that influence people design, use, and manage the Internet” (IETF, 2021b, para. 2).

Triangulation filtered additional noise away from the truth learned in this research. Data triangulation occurred in this study because multiple datasets were used to consider events of the IESG from the viewpoint of the historical data in the IESG minutes (Datasets 1 to 2 in Strands-1–5), the viewpoint of WG data (Strand-2 and Strand-5), and the survey response data (Strands-3–5). The mixed-mode methods used within-method triangulation in Strands-1–3 during the comparison of the HRM modeling, and between-method triangulation occurred in the quantitative and qualitative analysis for Strands-1–5. Strand-5 qualitative analysis methods used theory triangulation to evaluate the IESG cohorts per IETF chair by using the reduced theoretical and organizational development models to explain change leadership. These triangulations in Strands-1–5 created a pattern of filters for quantitative and qualitative data that separated the truth from the noise.

Cell phones may use a pattern of filters to separate the signal from the noise. Research into better noise reduction may examine the pattern of filters so that the cell phone delivers clear sounds for speech, music, or videos. Research into better cell phone technology examines the output of each filter before looking into how multiple filters in a cell phone interact, so an improvement in the filter for noise in speech does not reduce the quality of the sound of music or videos. The output of the data was exposed during analysis for Strands-1–3 (see Appendices O to R) and the Strands-4–5 analysis in Chapter 4. This chapter discusses the pattern of triangulation filters used in qualitative and quantitative analysis in each of these strands. This paradigm of signal/noise theory may help other researchers consider new combinations of triangulation in mixed mode longitudinal studies.

Based on the described quantity of quality data and the triangulation pattern, this research found an increase in solidarity behaviors in a team predicts the effectiveness of consensus decisions made in team consensus decision-making better than OCB for the context of the IESG TMT in the IETF. Strand-5 concludes that this modest conclusion has scholarly validity and real-world credibility within that context. Additional longitudinal historical data (1991 to 2014 and 2017 to 2021) or additional survey data on the IESG could expand the context of this conclusion. However, the knowledge gained about conflict and the reduced theoretical model suggests that an alternate model for solidarity, conflict, and task-interdependence might replace the full or reduced models. This chapter describes the knowledge gained on solidarity and conflict, the Hares reduced model, and potential changes to the Hares model. Future research plans include expansions to the IESG study and expansion of this line of research into longitudinal studies of TMTs in other volunteer organizations in the 21st century or first century.

Quantity of Quality Data Matters

The researcher found that the quantity of quality data matters in mixed-mode research during the data collection and analysis phases. Therefore, the original methodology set the surveys sampling plans and data collection methods based on the best practices described by Creswell (2009). An error hid a survey retest error in this research’s original description of the 2013 survey that made it seem like published research and hid the issue during the review of the expert reviewers on the Ph.D. committee. Dr. Emily Cabanda’s review comments pointed out the error in the 2017 survey. As in all survey research, if the researcher had carefully listened to her reviewers, the survey research methodology could have been revised. Perhaps the low response to the 2017 survey might have been avoided. This methodology issue caused a reduction in the quantity of quality data from the 2017 survey. The low response on the 2017 survey data and 2013 survey model reduced the quantitative analysis to the Hares reduced model in Strands-3–5.

The historiometric best practices helped keep historical data collection focused on qualitative data for Strand-1, Strand-2, and the qualitative analysis of open-ended conflict responses in Strand-3. Ligon et al. (2012) and Parry et al. (2014) indicated that the sample plan for historical sources must select source material containing the research studies' behaviors. This discussion examines how these historiometric best practices enabled quality data collection of appropriate sources for the IPA analysis (10% and 100%) for the Strand-1 analysis of behaviors, the Strand-2 analysis of decisions and validated results, the Strand-2 collection of WG data per area, and the data needed to provide qualitative analysis of Strand-3's open-ended questions on conflict. A discussion of the benefits of the data collection for Strand-1, Strand-2, and Strand-3 will help demonstrate the power of the historiometric best practices for data collection.

The researcher based the sample plan for Strand-1 for the IPA analysis on obtaining IESG minutes which recorded the individual and group behaviors for 10% of the IESG (1992, 2020) minutes from 1991 to 2016 and 100% of the minutes for Dataset 2 (2015 to 2016). Because the IESG operates a continuous virtual consensus decision-making process aided by IETF tools, the IETF (2016b) Datatracker database, and the IETF website, the formal and narrative minutes function as a collection of references pointers into these online data. Due to IETF's open standards process, the online data are available to all who register for an account on the IETF (2016b) Datatracker. Inspired by the historiometric best practices, the researcher found the hand-merge of the formal and narrative minutes per meeting contained pointers to the appropriate online data per meeting that constitute a full credit of the decisions and the individual and group behaviors that occurred in decisions.

The online data were stored in over 5,000 notes for the IESG minutes for 133 IESG meetings (78 for 10% sample and 55 for 100%). The IPA analysis of this quantitate of quality data (10% and 100%) found 39,186 individual behavior acts in 3,458 decisions. The researcher examined each individual's actions during an IESG decision to determine if behaviors studied existed in the behavioral actions. The IPA analysis examined each individual's behavioral actions per decision for 34 individual survey questions for the five full model variables (HS, VS, TC, RC, and TI), two alternate model variables (OCB-GC, OCB-A), 10 queries for the two discovered variables (ThankAid and FlagIssue) and one query on the state of collaboration that existed in the individual behavior patterns (45 total questions),

Historiometric principles guided data collection in Strand-2 and stand-3. The data collection for Strand-2 data collection ensured the IPA analysis found and validated data related to IESG decisions. The revised methodology gathered

Strand-2 data on the decisions and results from the merged IESG minutes and validated this data against online sources. The researcher stored this data in code notes attached to the text in the IESG minutes describing the IESG decision. Due to the historiometric methodology, the research ignored all online decisions except the data referenced by IESG minutes. In addition, the researcher gathered WG information during Strand-2's data collection phase. The data collection for the WG data used the online list of WGs per IETF Area to guide data collection of the WG data and summarization of the data. The online WG information gathered by the researcher in Strand-2 filled ~40 notebooks. However, the researcher reduced this information to the summaries of the progression of WGs found in Appendix R. The Strand-5 data used this qualitative data from Strand-2 to create the progression of standards in four technology groupings and one organization grouping that validated the IESG effectiveness in consensus decision making over time. Historiometric best practices provided a focusing lens on data collection to rapidly get a quantity of quality data.

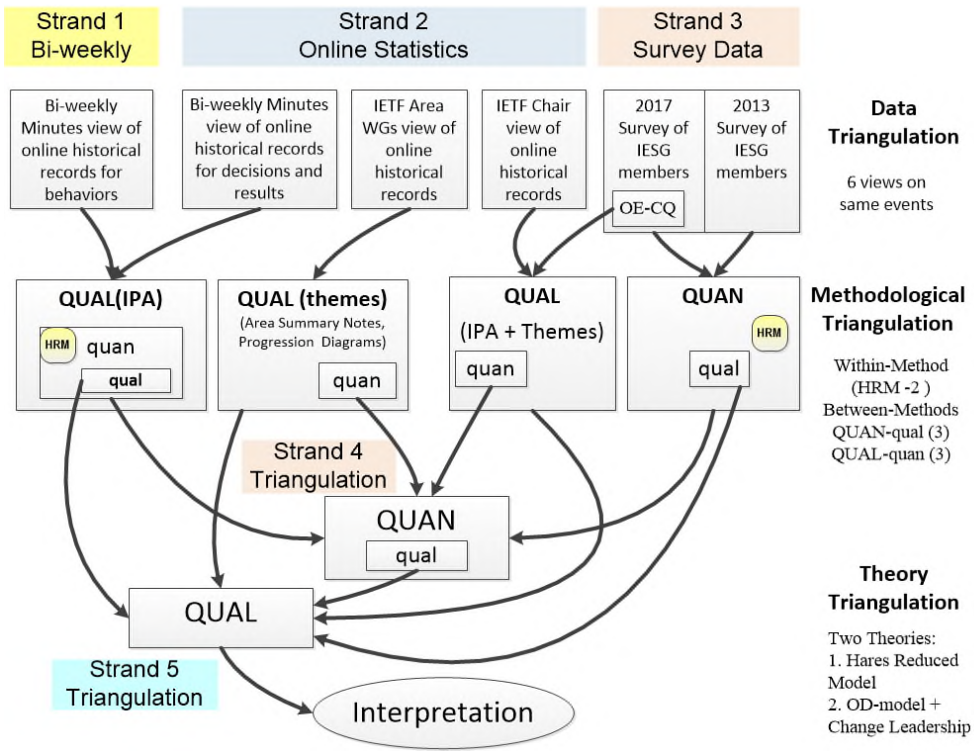
The Strand-3 IPA qualitative analysis of the open-ended conflict responses indicated that external pressures and internal pressures within the IETF or the IESG caused conflict specific to particular periods in the IETF related to IESG cohorts per chair. Qualitative data was collected and summarized per the IETF chair's tenure to validate this data on conflict against the external and internal pressures on the IESG and the existence of factions within the IESG. Best practices in historiometric methodology suggested collecting data based on the IETF chair's report per IETF meeting, the WG data for the organizational areas, and the operations requirement area. The quantity of quality data in the online sources regarding the IETF chair and the open-ended questions in Strand-3 helped provide validity and context checking for the conflict responses. Qualitative analysis in Strand-5 that compiled the qualitative results data from Strands-1–4 had strong foundations due to commonalities of the historiometric practices that focus on obtaining a quantity of quality data.

Triangulation Patterns

This research used data triangulation, method triangulation, and theory triangulation in the pattern shown in Figure 44 using a majority strategy in each triangulation. Fusch, Fusch, and Ness (2018), in their review of four categories of triangulation (data, investigator, theory, and methodological), recommended that mixed-mode research should use between-method correlation to “account for flaws and weaknesses” (p. 23) but did not indicate what type triangulation should be used. This researcher used data triangulation to present six views on each IESG Cohort’s tenure with three historical viewpoints (IESG minutes, IETF area WGs, and IETF chair view), and three perspectives from IESG members (2017 survey, 2017 open-ended conflict questions, and 2013 survey). This data triangulation sought the truth behind historical records and perceptions of history.

The methodological triangulation used between-method triangulation to determine if HRM modeling supports either of the two hypotheses in the reduced theoretical model. The research methods used between-methods methodology triangulation three times when quantitative data were the majority analysis (QUAN-qual) and three times when qualitative data were the majority analysis (QUAL-quan). Methodological triangulation sought to increase validity and reliability through these triangulation filters. The theory triangulation was used in Strand-5’s triangulation to provide a final filter on the researcher’s conclusions.

Figure 44: Web of triangulation.



The triangulation algorithm used in the triangulation was to take the opinion of the majority view. The triangulation algorithms perform functions, like the complex filters in a cell phone, to adjust the filters that tune these filters to keep up signal quality for different types of data (text, voice, music, and video). Kern (2018) examined the following five strategies for triangulating different sources of data regarding “status quo of political power and the preferred political power of traditional leaders”: “random selection,” “arithmetic mean,” “majority strategy,” “weighted average,” and “winner takes all” (pp. 166–170) in an empirical study of leaders in Uganda and Tanzania. Kern (2018) suggested that majority triangulation facilitated “triangulation of true value and comparison” (p. 166) but noted that this method did not allow weighting if some sources are of inferior quality. Ashour (2018) evaluated published research articles that examined the service markets based on high-tech telecommunication services and recommends using triangulation in all stages of the research on these markets due to rapid changes in this market. Ashour (2018) pointed out that triangulation increases “the validity and creditability of evaluation findings" (p. 205). Abdalla, Oliveira, Azevedo, and Gonzalez (2018) reviewed mixed-mode research that used triangulation, but they do not describe the methods to choose between alternatives in methodological triangulation as Kern (2018).

This research provided a complex methodology involving triangulation in data collection, data methods, and theories of interpretation. This complex methodology evaluated the IESG, a change agent that continued to push the pace on the high-tech telecommunication service and equipment markets. The current researcher considered individual triangulation as a filter that would improve the quality of the validity of the research truth. However, research on leadership and social science had not yet adopted theories that handle tuning a web of triangulation processes. The data from this mixed-mode study could be used in an experiment with different algorithms for triangulation within a web of triangulation in leadership research.

Clear Knowledge Advancement for Solidarity

This research advances the knowledge about the ability of solidarity to predict the effectiveness of team consensus decisions made in team consensus decision-making in a TMT using virtual consensus decision-making in a volunteer-based SDO, creating open standards for the Internet. In this environment, solidarity predicts the effectiveness of team consensus decision-making better than OCB. Conflict behaviors correlate to consensus decision-making in this environment but do not moderate the influence of solidarity on consensus decision-making. Beyond that fact, this research can only provide a few observations. First, conflict measured in historical records as the sum of TC recorded positively impacts the discussions of emerging standards. Second, conflict measured a combination of TC and RC in the cumulative reflections of a TMT member (IESG member) regarding a specific year has a negative impact where TC and RC exist. Third, conflict within a TMT that leads a change agent organization, such as the IETF, can be caused by forces that resist or push for change that are exterior to the organization or within the organization. When the IETF chair exhibited the change leadership behaviors recommended by the best leadership practices, it helped the IESG, as a TMT, retain solidarity and reduce conflict. This research data may help leaders of other virtual TMTs to understand the value of change leadership behaviors. The transition of these observations on conflict to research conclusions requires more quantities of quality data and evaluation of the effects of the web of triangulation. Finally, the applicability of the theoretical model of team consensus to virtual consensus decision-making in TMTs in other volunteer organizations requires additional longitudinal studies in other volunteer organizations.

These simple conclusions are significant for research into IETF and other SDOs working in Internet technology (W3C and IEEE). These conclusions come from a 28-year longitudinal study that examines historical records collected using historiometric best practices and participants' perceptions gathered by survey woven together by a web of triangulations to ensure the conclusions' validity. These conclusions also explain mixed results in earlier solidarity research and comparing solidarity with OCB. This evidence for solidarity required a web of data triangulation, methodological triangulation, and theory triangulation to filter the data into truth, just as playing streaming videos over the Internet requires a quantity of quality signal (e.g., 5 bars in 4G networks) and phone technology with complex filters to tune signal quality.

Future Work

Future directions regarding solidarity research involve expanding on the research of the IESG and expanding this research to other volunteer or commercial organizations with a TMT and longitudinal data. Expanding the research on the IESG will take three avenues: expanding the data analyzed, examining a revision to the Hares full and reduced models, and examining different triangulation methods. Future researchers should seek to expand the quantity of quality data by expanding the IPA analysis of historical data pointed by the IESG minutes to 100% of the data from 1991 to 2020 from the current state of 10% from 1991 to 2014 and 100% from 2015 to 2016. Such future researchers may to seek to gain different perceptions of the IESG members based on another survey. The expansion of the

IPA requires investing time in automation of data collection and transfer theme counts generated IPA encodings at different levels of summarizations (individual decision, group, meeting, and IESG cohort year or IESG Cohort years for an IETF chair) for quick transfer to the SPSS statistical package. The automated tools may reduce the time for the remaining research. This research has validated historical datasets that can be used to test these automated tools. The expansion of the survey data involves collaboration with other experts if a third resurvey of the IESG is possible using survey best practices or if the only expansion of survey data is to survey IESG members from 2017 to 2021 IESG cohorts. Another potential future direction for the IESG research is to create a second reduced theoretical model on solidarity that considers conflict simply as a negative correlation and test this model in parallel with the original model. A third potential direction for future research is to use the IESG research data to consider the impact of different types of triangulation filters in the web of filters.

Expanding this research to other volunteer organizations with a TMT and longitudinal data will determine if the conclusions about the predictive relationships between solidarity and the effectiveness of team consensus decisions are valid outside the original environment. Modern volunteer organizations that keep long-term records include civil service organizations (e.g., Red Cross), international religious organizations (e.g., Catholic Church or the United Methodist Church), sports organizations (e.g., Little League baseball), organizations that care for graveyards, and a variety of other organizations. The researcher plans a 25-year longitudinal study of an individual United Methodist church in Michigan.

Volunteer organizations for civil or religious organizations have existed for over 2,000 years. First-century historical writings record civil and religious leaders encouraging group harmony or unity and advocating avoiding conflict. Act 15 contains one example of religious leaders in TMT encouraging unity on a consensus decision and avoiding conflict. One challenge in examining first-century studies leadership is that a researcher needs to examine the historical records in the original language (e.g., Greek or Latin) to identify the solidarity and conflict behaviors. An example of a study that examines these records in the original language is Story's (2010) paper on the leadership paradigms in Acts 15 can identify terms that imply solidarity as mechanisms that resolve conflict in consensus decisions. The researcher also plans to expand this study to the Christian church in Corinth (70 to 120 CE) using historiometric research to examine three letters written to the Corinthian Church (1Cor, 2Cor, and 1 Clement) plus other historical documents. As these few examples of potential and planned research have shown, this initial research opens the door to many different pathways to advance the knowledge of how solidarity and conflict impact top leadership teams.

References

Abdalla, M. M., Oliveira, L. G., Azevedo, C. E., & Gonzalez, R. K. (2018). Quality in qualitative organizational research: Types of triangulation as a methodological alternative. *Administração: Ensino E Pesquisa*, *19*(1), 66–98. doi:10.13058/raep.2018.v19n1.578

Allen, R. H., & Sriram, R. D. (2000). The role of standards in innovation. *Technological Forecasting and Social Change*, *64*, 171–181. doi:10.1016/S0040-1625(99)00104-3

Amabile, T. M. (1997). Motivating creativity in organizations: On doing what you love and loving what you do. *California Management Review*, *40*, 39–58. doi:10.2307/41165921

Amabile, T. M., Schatzela, E. A., Moneta, G. B., & Kramer, S. J. (2004). Leader Behaviors and the work environment for creativity: Perceived leader support. *The Leadership Quarterly*, *15*(1), 5–32. doi:10.1016/j.leaqua.2003.12.003

Amason, A. C. (1996). Distinguishing the effects of functional and dysfunctional conflict on strategic decision making: Resolving a paradox for top management teams. *Academy of Management Journal*, *39*(1), 123–148. doi:10.5465/256633

Amason, A. C., & Mooney, A. C. (1999). The effects of past performance on top management team strategic direction. *The International Journal of Conflict Management*, *10*(4), 340–359. doi:10.1108/eb022829

Amason, A. C., & Sapienza, H. J. (1997). The effects of top management team size and interaction norms on cognitive and affective conflict. *Journal of Management*, *23*(4), 495–516. doi:10.1016/S0149-2063(97)90045-3

Arkko, J. (2012a, April 5). *Distribution of processing times*. Retrieved May 15, 2016, from Internet Engineering Task Force website: http://www.arkko.com/tools/lifecycle/wgdistr.html

Arkko, J. (2012b, June 07). *IESG and AD measurements*. Retrieved January 5, 2016, from Internet Engineering Task Force website: http://www.arkko.com/tools/admeasurements/stat/base.html

Arkko, J. (2016a, May 15). *Document statistics*. Retrieved May 15, 2016, from Internet Engineering Task Force website: http://www.arkko.com/tools/docstats

Arkko, J. (2016b, July 20). *IETF chair and IESG report, IETF 96*. Retrieved from Internet Engineering Task Force website: https://www.ietf.org/proceedings/96/slides/slides-96-ietf-plenary-0.pdf

Arthurs, J., Drakopoulou, S., & Gandini, A. (2018). Researching YouTube. *Convergence*, *24*(1), 3–15. doi:10.1177/1354856517737222

Ashour, M. L. (2018). Triangulation as a powerful methodological research technique in technology-based services. *Business & Management Studies: An International Journal*, *6*(1), 193–2018. doi:10.15295/bmij.v6i1.209

Ayub, N., & Jehn, K. (2014). When diversity helps performance: Effects of diversity on performance in workgroups. *International Journal of Conflict Management*, *25*(2), 189–212. doi:10.1108/IJCMA-04-2013-0023

Balkundi, P., & Kildruf, M. (2006). The ties that lead: A social network approach to leadership. *Leadership Quarterly*, *17*, 419–439. doi:10.1016/j.leaqua.2006.01.001

Ballardie, A. (1997, September). *RFC2201: CBT multicast routing architecture*. San Francisco, CA: Internet Engineering Task Force. doi:10.17487/RFC2201

Barker, L. L., Wahlers, K., & Watson, K. (2001). *Groups in process* (6th ed.). Needham Heights, MA: Allyn & Bacon.

Barron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173–1182. doi:10.1.1.917.2326

Bass, B. M., & Riggio, R. E. (2006). The transformational model of leadership. In G. R. Hickman (Ed.), *Leading organizations: Perspectives for a new era* (2nd ed., pp. 76–86). Loss Angeles, CA: Sage.

Besnahan, T. F., & Yin, P. L. (2007). Standard setting in markets: The browser war. In S. Greenstein, & V. Stango (Eds.), *Standards and public policy* (pp. 18–59). New York, NY: Cambridge University Press.

Bezrukuva, K., Jehn, K. A., Zanutto, E. L., & Thatcher, S. M. (2009). Do workgroup faultlines help or hurt? A moderated model of faultlines, team identification, and group performance. *Organization Science, 20*, 35–50. doi:10.1287/orsc.1080.0379

Bjorklund, R., & Holt, S. (2011). Overcoming barriers to participation in diverse strategic decision-making groups: A leadership perspective. *International Journal of Business and Management, 7,* 49–57. doi:10.1.1.475.5748

Bradner, S. (1996). *RFC2026: The Internet standards - Revision 3*. Retrieved from http://www.rfc-editor.org/rfc/rfc2026.txt

Bryman, A. (2006). Integrating quantitative and qualitative. *Qualitative Research*, *6*(1), 97–113. doi:10.1177/1468794106058877

Bucholtz, A. K., Amason, A. C., & Rutherford, M. A. (2005). The impact of board monitoring and involvement in top management team affective conflict. *Journal of Managerial Issues, 17*(4), 405–422. Retrieved from https://www.pittstate.edu/

Cabanda, E., Fields, D., & Winston, B. (2011). *Organizational leadership Ph.D.: Quantitative research methods*. New York, NY: McGraw-Hill Higher.

Carmeli, A., & Schaubroeck, J. (2006). Top management behavioral integration, decision quality, and organizational decline. *The Leadership Quarterly, 17*(3), 441–453. doi:10.1016/j.leaqua.2006.06.001

Chiao, B., Lerner, J., & Tirole, J. (2005). *The rules of standard setting organizations: An empirical analysis*. Retrieved from http://www.nber.org/papers/w11156

Choi, D. G., Lee, H., & Sung, T. (2011). Research profiling for standardization and innovation. *Scientometrics*, *88*(1), 259–278. doi:10.1007/s11192-011-0344-7

Cisco. (2018). *Cisco Global Cloud Index: Forecast and methodology, 2016–2021*. Retrieved November 20, 2021, from Cisco website: https://virtualization.network/Resources/Whitepapers/0b75cf2e-0c53-4891-918e-b542a5d364c5_white-paper-c11-738085.pdf

Cisco. (2021, March 9). *Cisco annual Internet report (2018–2023)* [White paper]. Retrieved November 20, 2021, from https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.pdf

Cole, M. S., & Bedeian, A. G. (2007). Leadership consensus as a cross-level contextual moderator of the emotional exhaustion–work commitment relationship. *The Leadership Quarterly, 18*(5), 447–462. doi:10.1016/j.leaqua.2007.07.002

Cole, M. S., Bedeian, A. G., & Bruch, H. (2011). Linking leader behavior and leadership consensus to team performance: Integrating direct consensus and dispersion models of group composition. *The Leadership Quarterly, 22*, 383–398. doi:10.1016/j.leaqua.2011.02.012

Cole, M. S., Schaninger, W. J., & Harris, S. (2002). The workplace social network: A multi-level conceptual examination. *Group & Organizational Management, 27*, 142–167. doi:10.1177/1059601102027001008

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches*. Los Angeles, CA: Sage.

Creswell, J. W., & Plano-Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). Los Angeles, CA: Sage.

Cummings, T. G., & Worley, C. G. (2005). *Organizational development and change* (8th ed.). Mason City, IA: Thomson South-Western.

Day, J. (2008). *Patterns in network architecture - A return to fundamentals*. Boston, MA: Prentice Hall - Pearson Education.

DeChurch, L. A. (2011). A historiometric analysis of leadership in mission critical multiteam environments. *The Leadership Quarterly, 22*, 152–169. doi:10.1016/j.leaqua.2010.12.013

DeChurch, L. A., Mesmer-Magnus, J. R., & Doty, D. (2013). Moving beyond relationship and task conflict toward a process-state perspective. *Journal of Applied Psychology*, *98*(4), 559–578. doi:10.1037/a0032896

DeDreu, C. K., & Weingart, L. R. (2003). Task versus relationship conflict, team performance, and team member satisfaction. *Journal of Applied Psychology*, *88*(4), 741–749. doi:10.1037/0021-9010.88.4.741

Dess, G. (1987). Consensus on strategy formation and organizational performance: Competitors in a fragmented industry. *Strategic Management Journal*, *8*, 259–277. doi:10.1002/smj.4250080305

Dess, G., & Origer, N. (1987). Environment, structure, and consensus in strategy formation: A conceptual integration. *The Academy of Management Review*, *12*, 637–647. doi:10.5465/amr.1987.4307943

de Tocqueville, A. (1945). *Democracy in America* (Vol. 2). (H. Reeve, Trans.) New York, NY: Colonial. (Original work published 1899)

deWit, F. R., Greer, L. L., & Jehn, K. A. (2012). The paradox of intragroup conflict: A meta-analysis. *Journal of Applied Psychology*, *97*(2), 360–390. doi:10.1037/a0024844

Ding, A. Y., Korhonen, J., Savolainen, T., Kojo, M., Ott, J., Tarkoma, S., & Crowcroft, J. (2013). Bridging the gap between internet standardization and networking research. *ACM SIGCOMM Computer Communication Review*, *44*(1), 56–62. doi:10.1145/2567561.2567572

Dooley, R. S., & Fryxell, G. (1999). Attaining decision quality and commitment from dissent: The moderating effects of competency and loyalty-based trust in strategic decision-making. *Academy of Management Journal*, *42*(4), 389–402. doi:10.5465/257010

Edwards, B. D., Day, E. A., Arthur, W., & Bell, S. T. (2006). Relationships among team ability composition, team mental models, and team performance. *Journal of Applied Psychology*, *91*(3), 727–736. doi:10.1037/0021-9010.91.3.727

Egyedi, T. M. (2003). Consortium problem redefined: Negotiating 'democracy' in the actor network on standardization. *International Journal of IT Standards and Standardization Research*, *1*(2), 22–38. doi:10.4018/jitsr.2003070102

Egyedi, T. M., Vrancken, J., & Ubacht, J. (2012). Inverse infrastructures: Coordination in self-organizing systems. InT. M. Egyedi & D. C. Mehos (Eds.), *Inverse infrastructures* (pp 291–312). doi:10.4337/9781781952290.00026

Ensley, M. D., & Pearce, C. L. (2001). Shared cognition in top management teams: Implications for new venture performance. *Journal of Organizational Behavior*, *22*, 145–160. doi:10.1002/job.83

Fielder, F. (1964). A contingency model of leadership effectiveness. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 1, pp. 149–190). New York, NY: Academic Press.

Fielder, F. (1967). *A theory of leadership effectiveness*. New York, NY: McGraw-Hill.

Fielding, A. T. (1999). The shared leadership in the Apache project. *Communications of the ACM*, *42*(4), 42–43. doi:10.1145/299157.299167

Fields, D. L. (2002). *Taking the measure of work*. Thousand Oaks, CA: Sage.

Fusch, P., Fusch, G. E., & Ness, L. R. (2018). Denzin's paradigm shift: Revisiting triangulation in qualitative research. *Journal of Social Change*, *10*(1), 19–32. doi:10.5590/JOSC.2018.10.1.02

Gençer, M. (2012). The evolution of IETF standards and their production. *International Journal of IT Standards and Standardization Research*, *10*(1), 17–33. doi:10.4018/jitsr.2012010102

Giambatista, R. C. (2004). Jumping through hoops: A longitudinal study of leader life cycles in the NBA. *The Leadership Quarterly*, *15*(5), 607–624. doi:10.1016/j.leaqua.2004.07.002

Girden, E. R., & Kabacoff, R. (Eds.). (2010). *Evaluating research articles from start to finish*. Thousand Oaks, CA: Sage.

Gouldner, A. (1960). The norm of reciprocity: A preliminary statement. *American Sociological Review*, *25*(2), 161–178. doi:10.2307/2092623

Graen, G., & Uhl-Bien, M. (1995). Development of leader-member exchange (LMX) theory of leadership over 25 years: Applying a multi-level multi-domain perspective. *Leadership Quarterly*, *6*, 219–247. doi:10.1016/1048-9843(95)90036-5

Greene, J. (2007). *Mixed modes in social inquiry*. San Franciso, CA: Wiley & Sons.

Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Education Evaluation and Policy Analysis*, *11*(3), 255–274. doi:10.3102/01623737011003255

Gross, P. (1989). Chairman's message. In *Proceedings of the 14th Internet Engineering Task Force (July 24–28) Stanford University* (pp. 7–8). Reston, VA: Corporation for National Research Initiatives. Retrieved November 20, 2021, from https://www.ietf.org/proceedings/14.pdf

Guinot, J., Chiva, R., & Mallen, F. (2015). The effects of altruism and relationship conflict on organizational learning. *Organizational Learning*, *26*(1), 86–112. doi:10.1108/IJCMA-12-2013-0100

Haas, J., & Patel, K. (2021). *Discussion on the value of BGP MIB* (S. Hares, Interviewer).

Haenggi, M., Andrews, J. G., Baccelli, F., Dousse, O., & Franceschetti, M. (2009). Stochastic geometry and random graphs for the analysis and design of wireless networks. *Institute of Electrical and Electronics Engineers Journal on Selected Areas in Communications*, *27*(7), 1029–1046. doi:10.1109/JSAC.2009.090902

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Upper Saddle River, NJ: Prentice-Hall.

Hambrik, D., Cho, T., & Chen, M.-J. (1996). The influence of top management teams on firms' competitive moves. *Administrative Science Quarterly*, *41*, 659–684. doi:10.2307/2393871

Hares, S. (2012, November 2). *Initial report on IETF leadership*. Retrieved January 19, 2016, from Hares IESG Research website: https://www.ndzh.com/report-Nov-2012

Hares, S. (2013, June 27). *What the IESG minutes tell us: Initial report*. Retrieved January 29, 2016, from Hares IESG Research website: www.ndzh.com/iesg-research/Minutes/Initial-Report-June-2013

Harrington, R., Lemak, D., & Kendall, K. (2002). The thread-rigidity thesis in newly formed teams: An empirical test. *Journal of Business and Management*, *8*, 127–145. Retrieved from https://home.csulb.edu/

Hechter, M. (1987). *Principles of Group Solidarity*. Berkely: University of California Press.

Hinds, P. J., & Cramton, C. D. (2012). Studying global work groups in the field. In A. B. Hollingshead & M. S. Poole (Eds.), *Research methods for studying groups and teams* (pp. 105–120). Thousand Oaks, CA: Sage.

Hogg, M. A. (2001). A social identity theory of leadership. *Personality & Social Psychology Review, 5*(3), 184–200. doi:10.1207/S15327957PSPR0503_1

Hogg, M. A., Fielding, K., Johnson, D., Masser, B., Russell, E., & Svensson, A. (2006). Demographic category membership and leadership in small groups: A social identity analysis. *The Leadership Quarterly*, *17,* 335–350. doi:10.1016/j.leaqua.2006.04.007

Hogg, M. A., & Terry, D. I. (2000). Social identity and self-categorization in small groups: A social identity analysis. *Academy of Management Review*, *25*, 121–140. doi:10.2307/259266

Hollingshead, A., & Poole, M. S. (Eds.). (2012). *Research methods for studying groups and teams: A guide to approaches, tools, and technologies*. Thousand Oaks, CA: Routledge.

House, R. (1971). A path-goal theory of leader effectiveness. *Administrative Science Quarterly, 16*, 321–329. doi:10.2307/2391905

Hunter, S. T., Johnson, J. E., & Ligon, G. S. (2011). First and 10 leadership: A historiometric investigation of the CIP leadership model. *The Leadership Quarterly*, *22*, 70–91. doi:10.1016/j.leaqua.2010.12.008

Internet Assigned Numbers Authority. (2020, February 1). *IANA protocol registries*. Retrieved from https://www.iana.org/protocols

Internet Engineering Steering Group. (1992). *IESG Steering Group minutes - December 21, 1991*. Retrieved 12 15, 2016, from https://www6.ietf.org/iesg/minutes/1992/minutes-1992-12-21.txt

Internet Engineering Steering Group. (2000, August 29). *IESG guidance on the moderation of IETF working group mailing lists*. Retrieved from https://www.ietf.org/about/groups/iesg/statements/mailing-lists-moderation/

Internet Engineering Steering Group. (2019, December 1). *IP security maintenance and extensions - Charter ballots for external review*. Retrieved December 1, 2019, from https://datatracker.ietf.org/doc/charter-ietf-ipsecme/ballot/295567/

Internet Engineering Steering Group. (2020, March 30). *Sunset4 Charter - History*. Retrieved from https://datatracker.ietf.org/doc/charter-ietf-sunset4/history/

Internet Engineering Task Force. (1986). *IETF 1 proceedings*. Retrieved November 20, 2021, from https://www.ietf.org/proceedings/01.pdf

Internet Engineering Task Force. (1987). Proceedings of the April 22-24, 1987 meeting of Internet Engineering Task Force. In P. Gross (Ed.), *Internet Engineering Task Force 6* (p. 335). Retrieved November 20, 2021, from https://www.ietf.org/proceedings/06.pdf

Internet Engineering Task Force. (1991). *IETF 22 proceedings*. Retrieved February 1, 2020, from https://www.ietf.org/proceedings/22.pdf

Internet Engineering Task Force. (1992). *IETF 23 proceedings*. Retrieved February 1, 2020, from https://www.ietf.org/proceedings/23.pdf

Internet Engineering Task Force. (2016a, May 15). *IESG members*. Retrieved May 15, 2016, from http://www.ietf.org/iesg/members.html

Internet Engineering Task Force. (2016b, May 6). *IETF Datatracker, 6.20.0*. Retrieved May 15, 2016, from https://datatracker.ietf.org/

Internet Engineering Task Force. (2016c, April 10). *IETF 95 meeting agenda - Buenos Aires, April 3-8, 2016*. Retrieved from https://datatracker.ietf.org/meeting/95/agenda.html

Internet Engineering Task Force. (2019, December 6). *Core based trees (CBT) multicast routing architecture - History*. Retrieved from https://datatracker.ietf.org/doc/rfc2201/history/

Internet Engineering Task Force. (2020a, March 1). *Active IETF working groups*. Retrieved from https://datatracker.ietf.org/wg/

Internet Engineering Task Force. (2020b, April 1). *BOFs - Wiki*. Retrieved from https://wiki.tools.ietf.org/bof/trac/

Internet Engineering Task Force. (2020c, March 1). *Concluded groups*. Retrieved from https://datatracker.ietf.org/group/concluded/

Internet Engineering Task Force. (2020d, April 1). *IETF BOF Wiki*. Retrieved from https://trac.tools.ietf.org/bof/trac/wiki

Internet Engineering Task Force. (2020e, April 1). *Interdomain routing - Group history*. Retrieved from https://datatracker.ietf.org/wg/idr/history

Internet Engineering Task Force. (2020f, April 1). *Interface to the routing system (i2rs) - Group history*. Retrieved from https://datatracker.ietf.org/wg/i2rs/history/

Internet Engineering Task Force. (2020g, April 1). *IP over asynchronous transfer mode (IPatm) - Group history*. Retrieved from https://datatracker.ietf.org/wg/ipatm/history/

Internet Engineering Task Force. (2020h, April 1). *Limited additional mechanisms for PKIX and SMIME (LAMPS) - History*. Retrieved from https://datatracker.ietf.org/wg/lamps/history/

Internet Engineering Task Force. (2020i, March 1). *Locator/ID separation protocol - Charter-ietf-LISP-04 - History*. Retrieved from https://datatracker.ietf.org/doc/charter-ietf-lisp/history/

Internet Engineering Task Force. (2020j, April 1). *Mobile ad-hoc networks*. Retrieved from https://datatracker.ietf.org/doc/charter-ietf-manet/history/

Internet Engineering Task Force. (2020k, March 2). *Multiparty multimedia session control (mmusic) - Group history*. Retrieved from https://datatracker.ietf.org/wg/mmusic/history/

Internet Engineering Task Force. (2020l, April 1). *Multiple interfaces (mif)*. Retrieved from https://datatracker.ietf.org/wg/mif/history/

Internet Engineering Task Force. (2020m, January 20). *RFC Index*. Retrieved from https://www.ietf.org/download/rfc-index.txt

Internet Engineering Task Force. (2020n, February 1). *RFC 3760*. Retrieved from https://datatracker.ietf.org/doc/rfc3760/ballot/

Internet Engineering Task Force. (2020o, February 1). *RFC 2201 history*. Retrieved from https://datatracker.ietf.org/doc/rfc2201/history/

Internet Engineering Task Force. (2021a, November 20). *Groups*. Retrieved from https://www.ietf.org/about/groups/

Internet Engineering Task Force. (2021b, November 20). *Mission and principles*. Retrieved from https://www.ietf.org/about/mission/

Internet Society. (2020, February 1). *Internet Society (home page)*. Retrieved from https://www.internetsociety.org/

Jehn, K. A. (1994). Enhancing effectiveness: An investigation into the advantages and disadvantages of value-based intragroup conflict. *International Journal of Conflict Management*, *5*(3), 223–238. doi:10.1108/eb022744

Jehn, K. A. (1995). A multimethod examination of the benefits and detriments of intragroup conflict. *Administrative Science Quarterly*, *40*(2), 256–282. doi:10.2307/2393638

Jehn, K. A. (1997). A qualitative analysis of conflict types and dimensions in organizational groups. *Administrative Science Quarterly*, *42*, 530–667. doi:10.2307/2393737

Jehn, K. A. (1999). Why differences make a difference: A field study of diversity, conflict, and performance in workgroups. *Administrative Science Quarterly*, *44*(4), 741–764. doi:10.2307/2667054

Jehn, K. A., & Bezrukuva, K. (2010). The faultline activation process and effects of activated faultlines on coalition formation, conflict, and group outcomes. *Organizational Behavior and Human Decision Processes, 112*, 24–42. doi:10.1016/j.obhdp.2009.11.008

Jehn, K. A., Chardwick, C., & Thatcher, S. (1997). To agree or not to agree: The effects of value congruence, individual demographic dissimilarity, and conflict on workgroup outcomes. *International Journal of Conflict Management*, *8*(4), 287–305. Retrieved from http://o-search.proquest.com/libary.regent.edu/docview/199041613?accountid=13479

Jehn, K. A., & Chatman, J. A. (2000). The influence of proportional and perceptual conflict composition on team performance. *The International Journal of Conflict Management*, *11*(1), 56–73. doi:10.1108/eb022835

Jehn, K. A., Greer, L., Levine, S., & Szulanski, G. (2008). The effects of conflict types, dimensions, and emergent states on group outcomes. *Group Decision and Negotiation*, *17*(6), 465–495. doi:10.1007/s10726-008-9107-0

Jehn, K. A., Northcraft, G. B., & Neale, M. A. (1999). Why differences make a difference: A field study of diversity, conflict and performance in workgroups. *Administrative Science Quarterly*, *44*(4), 741–763. doi:10.2307/2667054

Johnson, J. (2021, November 20). *Global digital population as of January 2021*. Retrieved from Statista.com website: https://www.statista.com/statistics/617136/digital-population-worldwide/

Kashy, D. A., & Hagiwara, N. (2012). Analyzing group data. In A. B. Hollingshead, & M. S. Poole (Eds.), *Research methods for studying groups and teams* (pp. 311–328). New York, NY: Routledge.

Kern, F. G. (2018). The trials and tribulations of applied triangulation: Weighing different data sources. *Journal of Mixed Methods Research*, *12*(2), 166–181. doi:10.1177/1558689816651032

Klimoski, R., & Mohammed, S. (1994). Team mental model: Construct or metaphor. *Journal of Management*, *20*(2), 403–437. doi:10.1016/0149-2063(94)90021-3

Korsgaard, M. A., Schweiger, D. M., & Sapienza, H. J. (1995). Building commitment, attachment, and trust in strategic decisions: The role of procedural justice. *Academy of Management Journal*, *38*(1), 60–84. doi:10.5465/256728

Koster, F., & Sanders, K. (2006). Organisational citizens or reciprocal relationships? An empirical comparison. *Personnel Review*, *35*(5), 519–537. doi:10.1108/00483480610682271

Koster, F., Stokman, F., Hodson, R., & Sanders, K. (2007). Solidarity through networks. *Employee Relations*, *29*(2), 117–137. doi:10.1108/01425450710719978

Kotlyar, I., Karakowsky, L., & Ng, P. (2011). Leader behaviors, conflict, and member commitment to team-generated decisions. *The Leadership Quarterly*, *22*, 666–679. doi:10.1016/j.leaqua.2011.05.007

Lambooij, M., Flache, A., & Siegers, J. (2009). Shadow of the future, risk aversion, and employee cooperation. *Rationality and Society*, *21*(3), 307–3365. doi:10.1177/1043463109337088

Lambooij, M., Sanders, K., Koster, F., VanEmerik, H., Raub, W., Flache, A., & Wittek, R. (2003). Vearandwoording vragelijst solidarity at work ("Questionnaire solidarity at Work"). *Soldiarity at Work Paper*, *28*, 1–58. Retrieved from https://dare.uva.nl/search?identifier=eaf2f702-f012-4e64-8131-756e82e517f0

Lambooij, M., Sanders, K., Koster, F., & Zwiers, M. (2006). Human resource practices and organisational performance: Can the HRM-performance linkage be explained by the cooperative behaviours of employees? *Management Revue*, *17*(3), 223–240. Retrieved from https://www.jstor.org/stable/i40083536

Lankau, M. J., Ward, A., Amason, A., Ng, T., Sonnefeld, J. A., & Agle, B. R. (2007). Examining the impact of organizational value dissimilarity in top management. *Journal of Managerial Issues*, *19*(1), 11–34. Retrieved from https://www.pittstate.edu/

Leiner, B. M., Cerf, V. G., Clark, D. D., Kahn, R. E., Kleinrock, L., Lynch, D. C., . . . Wolff, S. (2009). A brief history of the Internet. *ACM SIGCOMM Computer Communication Review, 39*(5), 22–31. doi:10.1145/1629607.1629613

Leiner, B. M., Cerf, V. G., Clark, D. D., Kleinrock, L., Lynch, D. C., Postel, J., . . . Wolff, S. (2003). *Brief history of the Internet*. Reston, VA: Internet Society. (Original work published 1999)

Licklider, J., & Clark, W. (1962, August). *On-line man computer communication*. Boston, MA: MIT Series of Memos.

Ligon, G. S., Harris, D. J., & Hunter, S. T. (2012). Quantifying leader lives: What historiometric approaches can tell us. *The Leadership Quarterly, 23,* 1104–1133. doi:10.1016/j.leaqua.2012.10.004

Lim, B.-C., & Klein, K. J. (2006). Team mental models and team performance: A field study of the effects of team mental model similarity and accuracy. *Journal of Organizational Behavior, 27,* 403–418. doi:10.1002/job.387

Lobato, R. (2019). *Netflix nations*. New York, NY: New York University Press.

Loughry, M. L., & Amason, A. C. (2014). Why won't task conflict cooperate? Deciphering stubborn results. *International Journal of Conflict Management, 24*(4), 333–358. doi:10.1108/IJCMA-01-2014-0005

MacKenzie, S. B., Podsakoff, P. M., & Fetter, R. (1991). Organizational citizenship behavior and objective productivity as determinants of managerial evaluations of salespersons' performance. *Organizational Behavior and Human Decision Processes, 50,* 123–150. doi:10.1016/0749-5978(91)90037-T

Manz, C., & Sims, H., Jr. (1987). Leading workers to lead themselves: The external leadership of self-managing work teams. *Administrative Science Quarterly, 32*, 10–128. doi:10.2307/2392745

Marks, M. A., Zaccaro, S. J., & Mathieu, J. E. (2000). Performance implications of leader briefings and team-interaction training for team adaptation to novel environments. *Journal Of Applied Psychology, 85*(6), 971–986. doi:10.1037/0021-9010.85.6.971

MAXQDA. (2016). The software for qualitative and mized methods research [Software]. Retrieved from https://www.maxqda.com/

Maxwell, J. (1992). Understanding and validity in qualitative research. *Harvard Educational Review*, *62*(3), 279–301. doi:10.17763/haer.62.3.8323320856251826

Meyers, R. A., & Seibold, D. R. (2012). Coding group interaction. In A. B. Hollingshead, & M. S. Poole (Eds.), *Research methods for studying groups and teams: A guide to approaches, tools, and technologies* (Kindle ed., pp. 329–357). New York, NY: Routledge.

Mitchell, R., & Boyle, B. (2008). A theoretical model of transformational leadership's role in diverse teams. *Leadership and Organizational Development Journal*, *30*, 455–574. doi:10.1108/01437730910968714

Mockapetris, P. (1983a, November 1). *RFC 882 - Domain names - Concepts and facilities*. Retrieved 11 20, 2021, from Internet Engineering Task Force website: https://datatracker.ietf.org/doc/rfc882/

Mockapetris, P. (1983b, November 1). *RFC 883 - Domain names - Implementation and specification*. Retrieved 11 20, 2021, from Internet Engineering Task Force website: from https://datatracker.ietf.org/doc/rfc883/

Mockapetris, P. (1987a). *RFC 1034 - Domain names and concepts*. Retrieved November 20, 2021, from from Internet Engineering Task Force website: https://datatracker.ietf.org/doc/rfc1034/

Mockapetris, P. (1987b). *RFC 1035: Domain names - implementation and specification*. Wilmington, DE: Internet Engineering Task Force.

Moorman, R., & Blakely, G. (1995). Individual-collectivism as an individual difference predictor of organizational citizenship behavior. *Journal of Organizational Behavior*, *16*, 127–142. doi:10.1002/job.4030160204

Morgeson, F. P., & DeRue, D. S. (2006). Event criticality, urgency, and duration: Understanding how events disrupt teams and influence team leader intervention. *The Leadership Quarterly*, *17*(3), 271–287. doi:10.1016/j.leaqua.2006.02.006

Nickerson, J. V., & Muehlen, M. Z. (2006). The ecology of standards processes: Insights from the Internet standards. *MIS Quarterly*, *30*(Special Issue/August 2006), 467–488. doi:10.2307/25148769

O'Mahony, S., & Ferraro, F. (2007). The governance of open source initiatives: What does it mean to be community managed? *Journal of Management & Governance*, *11*(2), 139–150. doi:10.1007/s10997-007-9024-7

O'Neill, T. A., Allen, N. J., & Hastings, S. E. (2013). Examining the “pros” and “cons” of team conflict: A team-level meta-analysis of task, relationship, and process conflict. *Human Performance*, *26*(3), 236–260. doi:10.1080/08959285.2013.795573

Organ, D. W. (1997). Organizational citizen behavior: It's construct clean-up time. *Human Performance*, *10*, 85–97. doi:10.1207/s15327043hup1002_2

Pallant, J. (2010). *SPSS survival manual* (4th ed.). New York, NY: Open University Press McGraw Hill.

Parker, G. (2006). What makes a team effective or ineffective? In J. Gallos (Ed.), *Organizational development: A Jossey-Bass reader*. San Francisco, CA: Wiley & Sons.

Parry, K., Mumford, M. D., Bower, I., & Watts, L. L. (2014). Qualitative and historiometric methods in leadership research: A review of the first 25 years of *The Leadership Quarterly*. *The Leadership Quarterly*, *25*(1), 132–151. doi:10.1016/j.leaqua.2013.11.006

Pearson, A. W., Easley, M. D., & Amason, A. C. (2002). An assessment and refinement of Jehn's Intragroup Conflict Scale. *The International Journal of Conflict Management*, *13*(2), 110–126. doi:10.1108/eb022870/

Pethia, R., Crocker, S., & Fraiser, B. (1991). *RFC1281: Guidelines for secure operation of the Internet*. Wilmington, DE: Internet Engineering Task Force.

Podsakoff, P. M., MacKenzie, S. B., & Ahearne, M. (1997). Moderating effects of goal acceptance on the relationship between group cohesiveness and productivity. *Journal of Applied Psychology*, *82*(6), 974–983. doi:10.1037/0021-9010.82.6.374

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*(5), 879–903. doi:10.1037%2F0021-9010.88.5.879

Podsakoff, P. M., Paine, J. B., & Bachrach, D. G. (2000). Organizational citizenship behaviors: A critical review of the theoretical and empirical literature and suggestions for future research. *Journal of Management*, *26*(3), 513–563. doi:10.1016/S0149-2063(00)00047-7

Postel, J. (1981a, September 1). *RFC 791 - Internet protocol*. Retrieved November 29, 2021, from Internet Engineering Task Force website: https://datatracker.ietf.org/doc/rfc793/

Postel, J. (1981b, September 1). *RFC 793 - Transmission control protocol*. Retrieved November 20, 2021, from Internet Engineering Task Force website: https://datatracker.ietf.org/doc/rfc793/

Postel, J. (Ed.). (1991, August 1). *RFC1250 - IAB official protocol standards*. Wilmington, DE: Internet Engineering Task Force. doi:10.17487/RFC1250

Prapavessis, H., & Carron, A. (1997). Cohesion and work output. *Small Group Research*, *28*(2), 294–301. doi:10.1177/1046496497282006

Pratt, M. G., & Kim, N. (2012). Designing for drift: Planning ethnographic qualitative research on groups. In A. B. Hollingshead, & M. S. Poole (Eds.), *Research methods for studying groups and teams: A guide to approaches, tools, and technology* (pp. 6–29). New York, NY: Routledge.

Rachuri, S., Subrahmanian, I., Bouras, A., Fenves, S. J., Foufou, S., & Sriram, R. D. (2008). Information sharing and exchange in the context of product lifecycle management: Role of standards. *Computer-Aided Design*, *40*, 789–800. doi:10.1016/j.cad.2007.06.012

Randewich, N., & Krauskopf, L. (2020, February 18). *20 years after dot-com peak, tech dominance keeps investors on edge*. Retrieved from Reuters website: https://www.reuters.com/article/us-usa-stocks-dotcombust-graphic-idUSKBN20C1J7

Request for Comments Editor. (2016a, May 15). *The number of RFC published per year*. Retrieved May 16, 2016, from http://www.rfc-editor.org/rfcs-per-year/

Request for Comments Editor. (2016b, May 15). *Submissions and publications, by stream and by status*. Retrieved May 16, 2016, from http://www.rfc-editor.org/reports/subpub_stats.php

Request for Comments Editor. (2020, March 15). *RFC Index*. Retrieved from https://www.rfc-editor.org/rfc-index.txt

Rico, R., Sanchez-Manzanares, M., Gil, F., & Gibson, C. (2008). Team implicit coordination processes: A team knowledge-based approach. *The Academy Of Management Review*, *33*, 163–184. doi:10.5465/amr.2008.27751276

Rowland, P., & Parry, K. (2009). Consensual commitment: A grounded theory of the meso-level influence of organizational design on leadership and decision-making. *The Leadership Quarterly*, *20*(4), 535–553. doi:10.1016/j.leaqua.2009.04.004

Russell, A. L. (2006). 'Rough consensus and running code' and the Internet-OSI standards war. *Institute of Electrical and Electronics Engineers Annals of the History of Computing*, *3*, 48–61. doi:10.1109/MAHC.2006.42

Rysman, M., & Simcoe, T. (2008). Patents and the performance of voluntary standard-setting organizations. *Management Science*, *54*(11), 1920–1934. doi:10.1287/mnsc.1080.0919

Rysman, M., & Simcoe, T. (2010). The performance of standard setting organizations: Using patent data for evaluation. In K. Jakobs (Ed.), *New applications in IT standards: Developments and progress* (pp. 95–108). Hershey, NY: Information Science Reference.

Sanders, K., & Schyns, B. (2006a). Leadership and solidarity behavior: Consensus in perception of employees within teams. *Personnel Review*, *3*, 538–556. doi:10.1108/00483480610682280

Sanders, K., & Schyns, B. (2006b). Trust, conflict and cooperative behaviour: Considering reciprocity within organizations. *Personnel Review*, *35*, 508–518. doi:10.1108/00483480610682262

Sanders, K., & Van Emmerik, H. (2004). Does modern organization and governance threat solidarity? *Journal of Management and Governance*, *8*(4), 351–372. doi:10.1007/s10997-004-4128-9

Schyns, B., Kroon, B., & Sanders, K. (2006). Leader-member exchange and solidarity behavior: A study of reciprocity and performance. In G. Graen, & J. Graen (Eds.), *Sharing network leadership* (Vol. 4): *LMX leadership* (pp. 225–246). Greenwich, CT: Information Age Publishing.

Seers, A. (1989). Team-member exchange quality: A new construct for role-making research. *Organizational Behavior & Human Decision Process*, *43*, 118–135. doi:10.1016/0749-5978(89)90060-5

Seers, A., Petty, M., & Cashman, J. (1995). Team-member exchange under team and traditional management. *Group and Organizational Management*, *20*, 18–38. doi:10.1177/1059601195201003

Simcoe, T. S. (2007). Delays and de jure Standardization: Exploring the slowdown in Internet standards development. In S. Greenstein, & V. Stango (Eds.), *Standards and public policy* (ppp. 260–295). New York, NY: Cambridge University Press.

Simcoe, T. S. (2012). Standard setting committees: Consensus governance for shared technology platforms. *American Economic Review*, *102*(1), 305–336. doi:10.1257/aer.102.1.305

Simcoe, T. S., & Waguespack, D. M. (2011). Status, quality, and attention: What's in a (missing) name? *Management Science*, *57*(2), 274–290. doi:10.1287/msnsc.1100.1270

Smith, C., Organ, D. W., & Near, J. (1983). Organizational citizenship behavior: Its nature and antecedents. *Journal of Applied Psychology*, *68*, 653–663. doi:10.1037/0021-9010.68.4.653

Spring, M. B., Grisham, C., O'Donnell, J., Skogseid, I., Snow, A., Tarr, G., & Wang, P. (1995). Improving the standardization process: Working with bulldogs and turtles. In B. Kahin, & J. Abbate (Eds.), *Standards policy for information infrastructure* (pp. 220–254). Boston, MA: The Massuchusettes Institute of Technology Press.

Stentz, J. E., Plano-Clark, V. L., & Matkin, G. S. (2012). Applying methods to leadership research: A review of current practices. *The Leadership Quarterly*, *23*(6), 1173–1183. doi:10.1016/j.leaqua.2012.10.001

Stewart, G. L., Manz, C. C., & Sims, H. P. (1999). *Team work and group dynamics*. New York, NY: Wiley & Sons.

Story, J. L. (2010). The Jerusalem Council: A pivotal and instructive paradigm. *Journal of Biblical Perspectives in Leadership and Entrepreneurship, 1*, 33–60. Retrieved from https://www.regent.edu/journal/journal-of-biblical-perspectives-in-leadership/the-jerusalem-council-a-pivotal-and-instructive-paradigm/

Tagger, S., & Ellis, R. (2007). The role of leaders in shaping formal team norms. *The Leadership Quarterly*, *18*, 105–120. doi:10.1016/j.leaqua.2007.01.002

Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches*. Thousand Oaks, CA: Sage.

Thatcher, S. M., Jehn, K. A., & Zanutto, E. (2003). Cracks in diversity research: The effects of diversity faultlines on conflict and performance. *Group Decision and Negotiation*, *12*(3), 217–241. doi:10.1023/A:1023325406946

3GPP. (2011, November 20). *GSM spec history*. Retrieved from https://www.3gpp.org/specifications/gsm-history

Van der Vegt, G., Emans, B., & Van de Vliert, E. (1998). Motivating effects of task and outcome interdependence in work teams. *Group and Organizational Management*, *23*, 124–144. doi:10.1177/1059601198232003

Van Dyne, L., Cummings, L., & Parks, J. (1995). Extra role behaviors: In pursuit of a construct and definitional clarity (a bridge over muddied waters). In L. L. Cummings & B. Straw (Eds.), *Research in organizational behavior* (Vol. 17, pp. 215–285). Greenwich, CT: JAI.

Van Ginkel, W. P., & van Knippenberg, D. (2012). Group leadership and shared task representations in decision making groups. *The Leadership Quarterly*, *23*, 94–106. doi:10.1016/j.leaqua.2011.11.008

van Knippenberg, D. (2011). Embodying who we are: Leader group prototypicality and leadership effectiveness. *The Leadership Quarterly*, *22*, 1078–1091. doi:10.1016/j.leaqua.2011.09.004

van Knippenberg, D., & Hogg, M. A. (2003). A social identity model of leadership effectiveness in organizations. *Research in Organizational Behavior*, *25,* 243–295. doi:10.1016/S0191-3085(03)25006-1

van Woerkom, M., & Sanders, K. (2010). The romance of learning from disagreement. The effect of cohesiveness and disagreement on knowledge sharing behavior and individual performance within teams. *Journal of Business and Psychology, 25*(1), 139–149. doi:10.1007/s10869-009-9136-y

Villa, J. R., Howell, J. P., Dorfman, P. W., & Daniel, D. L. (2003). Problems with detecting moderators in leadership research using moderated multiple regression. *The Leadership Quarterly*, *14*(1), 3–23. doi:10.1016/S1048-9843(02)00184-4

Wayne, J. H., & Cordeiro, B. L. (2003). Who is a good organizational citizen? Social perception of male and female employees who use family leave. *Sex Roles*, *49*, 233–247. doi:10.1023/A:1024600323316

Werle, R., & Iverson, E. J. (2006). Promoting legitimacy in technical standardization. *Science, Technology & Innovation Studies*, *3*, 19–39. Retrieved from https://eprints.utas.edu.au/1365/

West, J. (2007). *The economic realities of open standards: Black, White, and many shades of grey*. New York, NY: Cambridge University Press.

Williams, L., & Anderson, S. (1991). Job satisfaction and organizational commitment as predictors of organizational citizenship and in-role behaviors. *Journal of Management*, *17*(3), 601–617. doi:10.1177/014920639101700305

Winston, B., Fields, D., & Cabanda, E. (2011). *Organizational leadership Ph.D.: Qualitative methods*. New York, NY: McGraw-Hill Higher Education.

Yukl, G. (2010). *Leadership in organization*. Upper Saddle River, NJ: Prentice-Hall.

Yukl, G. A., Wall, S., & Lepsinger, R. (1990). Preliminary report on the validation of the managerial practices survey. In K. Clark & M. Clark (Eds.), *Measures of leadership* (pp. 223–237). Greensboro, NC: Center for Creative Leadership.

Zaccro, S., Ritman, A., & Marks, M. (2001). Team leadership. *The Leadership Quarterly*, *12*, 451–483. doi:10.1016/S1048-9843(01)00093-5