© 2010

Corinne M. Dalelio

ALL RIGHTS RESERVED

DEVELOPMENT OF A STRUCTURAL APPROACH TO THE STUDY OF COMPUTER-MEDIATED COMMUNICATION VIA ONLINE DISCUSSION BOARDS

By

CORINNE M. DALELIO

A dissertation submitted to the

Graduate School – New Brunswick

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Communication, Information and Library Studies

Written under the direction of

Professor Hartmut B. Mokros

And approved by

_____

_____

_____

_____

_____

New Brunswick, New Jersey

January 2010

ABSTRACT OF THE DISSERTATION

Development of a Structural Approach to the Study of Computer-Mediated Communication via Online Discussion Boards

By CORINNE M. DALELIO

Dissertation Director:

Hartmut B. Mokros

To address the research question, "How can the processes of computer-mediated communication be analyzed, described and understood through a natural history, microanalytic, structural approach?" a new analytical instrument that systematically examines the structural properties of CMC via asynchronous online discussion boards is presented. In this approach, the sequence of participant activity within discussion board threads, and the references that allow inferences about who or what a specific message points back to, are mapped in varying ways.

In this method, *posts*, or messages submitted at one time, are understood as meaningful acts that serve as building blocks of the communicative structure within a discussion board thread, or *post set*. Posts are mapped according to their order of

occurrence in order to learn more about how posting activity occurs within a post set. Posts are also linked to one another by *reference cues*, or highly reliable indicators of whom or what they are addressing. Two posts linked by a reference cue are referred to as a *post pair*, and post pairs are grouped into *chains* based on posts they share in common. This allows for the creation of new maps that represent a more meaningful account of participation, highlighting features of referencing activity in a post set. Several types of visualizations are presented, allowing for the observation and comparison of non-obvious features of communication in post sets.

The developed method is demonstrated and tested through the analysis of data from discussion boards in two distinct research settings: a college course website, and a fan site for the television series *Lost*. Features of posting activity and referencing activity are compared between and within these two data sets. Posting approaches are identified by observing individual differences in participation within and across post sets.

The method was found to successfully capture participation in the observed post sets. A majority of posts were involved in referencing activity, which was found to be indicative of interaction. The features of, and constitutive possibilities for, communication in the observed research settings are described and outlined. Suggestions for future research are made based on the discoveries of the present analysis.

ACKNOWLEDGEMENTS

I wish to thank my advisor, Dr. Hartmut B. Mokros, for his mentorship and guidance throughout my scholarly experience, for being generous with his time and support, for always challenging me, and for being a great teacher as I navigated through this project.

Thank you to my committee members, Dr. Mark Aakhus, Dr. Lynn Cockett, Dr. James Katz, and Dr. Robert Kubey, for all their valuable input and suggestions.

I would like to express my gratitude to all my friends and family for all their love, support and encouragement, especially Roland, my husband-to-be, who has been at my side through it all, always having faith in me and providing exactly what I needed to keep going.

This research was supported in part by Grant # 0540417 from the National Science Foundation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

CHAPTER 1
A STRUCTURAL APPROACH TO STUDYING COMPUTER-MEDIATED COMMUNICATION

Introduction

With the rise of the Internet, communication researchers have massive amounts of data available in the form of recorded archives of human communication. It is estimated that Internet users worldwide submit billions of messages daily via email, web conferencing, online social networks, file-sharing, discussion boards, blogs, ratings or review systems, and more. The systems that allow for these activities often keep logs of all messages, along with a date-time stamp of their submission. Although submitted messages can be in the form of images, audio, and video, a large portion are primarily text based. Online discussion boards, which deliberately archive and display text-based messages for purposes of asynchronous communication, represent an especially intriguing site of study from which we can learn more about how researchers may make use of these available data. By outlining a method that adapts the structural approach to studying communication, designed specifically for use with online discussion boards, this dissertation presents such an effort.

The focus of this research is to develop and test for the study of archived computer-mediated communication (CMC) by analyzing discussion board threads. Micro-analytical starting points for analysis are determined so that researchers can make sense of the communication that takes place, and the ways in which that communication may be understood. A discussion board may be defined as any website that allows for the submission of asynchronous, text-based messages by registered members, and which are organized into threaded displays. Currently, the 1600 most active online discussion

boards collectively contain over 1.5 billion posts, suggesting an enormous investment of time (source: http://rankings.big-boards.com as of November 13, 2008). These environments are both fascinating and edifying, as participants utilize the available tools of communication to connect with others, satisfy needs, or just pass time. It is through the use of threaded text based postings that members attempt to incorporate and express human emotions such as anger, empathy, humor, joviality, and despair.

The discoveries made in this dissertation will contribute to a growing knowledge base about online/digital communication, asynchronous communication, and computer-mediated communication. Much of the research to date has focused on communication features and outcomes on a general scale. This method provides a tool to look at the underlying structure of CMC by observing its micro-level features, something that has not yet been widely addressed (Herring, 2002; Hratsinski & Keller, 2007).

The Constitutive Perspective of Communication

The approach to study in this dissertation is based in the constitutive perspective of communication. This perspective views communication as the site in which humans construct and reconstruct their senses of reality, self, situation, and society (Mokros, 2003). Messages are conceptualized, not as the transmission of information from sender to receiver, but rather as indicators of the social world, both representing and constructing it. Individuals' interpretations of situations, identities, relationships and roles are expressed and influenced through their communication with others (Goffman, 1983). People draw upon their individual expectations and assumptions about social interaction and conventions as they work together to establish shared rules and norms for interacting (Bateson, 1996). A simple case of this occurs when one person enters a room, and another demonstrates awareness of him with a greeting. This convention establishes a

relational recognition of the other's presence through acknowledgement, respect, and acceptance (Mokros, 2006). In addition, the nature of expression is shaped, hindered, enhanced, and altered in numerous ways by both the situation and the interactive processes that take place (Mokros, 1985). There are interactive influences on what is expressed, how, when and why. Thus, macro-level understandings of the social world are produced and re-produced through everyday communicative processes, and their emergence can be analyzed and understood through micro-levels of interpretation. This process is not directly stated or even part of the participants' awareness, but it becomes apparent when interaction is observed and analyzed at a micro-level in the structural approach to analysis (Mokros, 2006).

The Structural Approach to Studying Communication

The structural approach was developed by Duncan and Fiske (1977), whose work was directly influenced by studies emerging from the collaborative *Natural History of an Interview* project of the mid 1950s (McQuown, 1991). By collecting data related to various micro-level features of a single filmed interview, the researchers involved in this project discovered that analyzing the stream of observable actions in sequence highlighted the processes by which individuals establish and interpret rules for interaction (Bateson, 1996). By mapping the sequence of an interaction, the observation of participants' actions and choices begins to present an emergence of the structure, or shared norms, rules, and conventions that govern the participation; as well as individual strategies for dealing with those conventions (Duncan & Fiske, 1985). This results from, and leads to, the development of values and variables that can be understood in terms of their patterns of persistence and change throughout the course of an interaction (Mokros,

1985). In other words, researchers employing this approach systematically map the terrain of what is observable in interaction as a way of bringing non-obvious micro-level features of actions and choices into focus. From these contextualized observations, the determination of structures, then strategies, is made possible. Once determined, new patterns and observations may be layered onto the identified structure, leading to the development and testing of new hypotheses from which even more can be learned. When this type of back and forth exchange between deductive and inductive reasoning takes place, the structural approach may be understood as abductive research (Peirce, 1955).

Rule-governed games such as chess, poker, baseball, and golf serve as a useful heuristic for understanding the structural approach. The rules of a game both regulate the play and enable the constitutive possibilities, giving rise to a unique experience each time the game is played. For a researcher to gain an analytical understanding of poker through observation, she may record data on variables such as the number of cards in a given hand, number of active players, bets made, amount of chips in the pot, and the proportion of chips in each players' possession at specific moments during play. The next step may be running correlations and other statistical tests on multiple combinations of these variables; the number of bets made with the number of face cards in hand, for example. These tests may or may not yield significant results, but in the end the researcher will be no closer to understanding how the game of poker is played. The failure to gain an understanding of the rules of the game results from attending to individual differences in play, rather than the underlying structure. This type of analysis would be diffused by player bluffs, assumptions, and mistakes, as well as the lack of any pre or post move considerations. If the researcher were to instead map the natural history of a game of

poker, the rules would emerge and moves would be contextualized. It is only after the structure is identified that analysis and identification of individual player differences and strategies can be successfully achieved.

Although the rules for human interaction are not as simple or defined as those in a game, the same concept applies. There is a conventional, social, and symbolic organization to human life that can be observed when individuals interact. This is something that researchers, as human beings themselves, possess a tacit understanding of, but often have trouble operationalizing. Like the game of poker, one can learn how interaction works, one can participate in it and describe it, but it is not until the moves are recorded and understood in sequence that the grammar of the activity can be determined. It is only then that hypotheses about moves made can be generated and tested, and made into probabilistic claims.

Interpretive Microanalysis

This dissertation builds on a body of work adopting the structural approach to the study of communication in both face-to-face and computer-mediated contexts, conducted by Mokros and his students in the development and application of the interpretive microanalytic technique, a transcriptional method that has evolved over time (Mokros, 2003). Specifically, the methods used for describing and understanding the observable features in the present research build on previous approaches to mapping and sequencing the natural history of communication, as conducted by Mokros (1984), Stephenson (1998) and Cockett (2000) for recorded face-to-face communication, as well as Karetnick (2000) and Rumsey (2001) for computer-mediated communication. Each of these studies, investigating different communicative situations recorded in varying ways,

interpreted micro-level features of interaction. This allowed for the inductive determination of sequence and structure, revealing higher-order qualities of the communication and individuals participating. Mokros, investigating the transcriptions of the dyadic introductory interactions of 16 pairs of individuals, documented patterns of persistence and change in their nonverbal actions by developing and testing signals and sequences of action variables – specifically, speaker gaze and gesture, smiling, and grammaticality of utterances – in and during speaking turns. Stephenson, analyzing features of personal identity and social interaction in a videotaped 74-minute ballroom dance lesson to identify the sequences of action, found that the instructor's strategies, communicative moves and interventions revealed his theories of practice and personhood. Cockett mapped the tape-recorded interactions of 11 people in a 25-minute group decision-making meeting, revealing that individual interpretations of one's role and definition of the situation shifted throughout the meeting based on both theories of personhood and the actions of others. Karetnick analyzed the structure and strategies of individuals participating in a multi-user dungeon (MUD), finding the communicative activities of greetings, farewells, facework and conflict avoidance to be conventions in an online environment. Finally, Rumsey discovered, by mapping the messages of individuals participating in a cancer support discussion board, that individuals' perceptions of and needs for online discussion board participation could be identified through this process. Through interpretive microanalysis, all of these studies have discovered and identified both structures and strategies through mapping of the natural history of communication in their respective research settings. The present research

extends these techniques and builds especially on Rumsey's and Karetnick's studies of interaction in computer-mediated communication situations.

Computer-Mediated Communication via Online Discussion Boards

Goffman (1983) defines social interaction as "that which uniquely transpires in social situations, that is, environments in which two or more individuals are physically in one another's response presence" (p. 2). The notion of being "physically" in one another's response presence is complicated in the computer-mediated environment of online discussion boards, where "presence [is] defined by communication alone" (Maltz, 1996). Computer-mediated communication affords participants the ability to overcome geographical, temporal, material and distributional constraints. Yet this also creates constraints of its own (Rice, 1987). The lack of physical co-presence in CMC can be distancing and confusing for participants, and may result in misunderstandings and misreading of intent (Crystal, 2006). Participants lose the immediacy of feedback witnessed in face-to-face encounters, conveyed through body posturing and positioning, facial expressions, and gestures, in addition to verbal utterances such as "uh-huh" and "yeah," that serve as within-turn back-channels in communication (Duncan & Fiske, 1977). The removal of synchronicity can also be constraining in that the social interactions are inevitably slowed down as compared to live, face-to-face environments. The time between an utterance and its response, what is often referred to as "lag," can range from seconds to weeks (Crystal, 2006). This increases not only the number of times an individual wishing to continue participating must return to a computer-mediated site over a given period of time, but it can also result in increased confusion over when, if ever, communication has ended. Finally, it is not always clear to participants with whom

they are communicating, as dropping out, joining, and rejoining can occur by any participant at any given time, without the awareness of other participants.

In computer-mediated communication, the absence of verbal and non-verbal feedback of others prevents individuals from clearly determining where they stand or how they are received. This may have significant consequences for social interaction. The largely unconscious processes of social interaction must become more conscious and explicit (Baym, 1996). Rather than expressing disapproval on one's face, for example, a person would have to consciously register that disapproval, and then decide to share it in a submitted message, before others could become aware of it. In addition, individuals' sense of the interaction, or their role in it, is less secure and assured. The experience of acknowledgement and acceptance, and the joint activity that is a constant feature of face-to-face interactions, is eliminated. It is only through the observable choices and actions of users, irrespective of these constraints, that interaction can be described, if and when it takes place (Hutchby, 2001). This raises questions about whether computer-mediated communication is, in fact, interactional, and the extent to which it can be considered social interaction in the same sense as face-to-face communication.

Yet, based on the numerous studies that have self-report measures to assess perceptions of individuals' participation, participants do perceive their engagement with CMC as genuine social experiences (Baym, 1995; Baym, 1998; Chester & Gwynne, 1998; Chiu et al., 2006; Dholakia et al., 2004; Ridings et al., 2002; Sun et al., 2006; Wang & Fesenmaeir, 2004; Wasko & Faroj, 2005; Ye et al., 2006). It has been discovered that individuals communicating via computer-mediated technologies routinely feel that they are able to gain a sense of not only what is occurring, but also of with

whom they are communicating (Baym, 1995; Rumsey, 2001; Walther & Burgoon, 1992). Baym (1995) has shown how aspects of humor and performance not only exist in this environment, but also play a large role in shaping identity formation. Several studies have found evidence of support and mentoring activities (Alexander, 1999; Aviv et al., 2004; Burnett, 2000; Dustdar & Gall, 2003; Farooq et al., 2007; Looi & Ang, 2000; Meyers et al., 2002; Rumsey, 2001; Wheeler, 2006). Finally, Burnett (2000) discovered additional features of informal communication, such as pleasantries, gossip, language games, and playful interactions.

Rather than focus on the ways in which barriers to communication are circumvented, the focus of this dissertation is to learn how to study online sites of archived human communication as it occurs through writing rather than speaking. Such writing on the Internet has created a new linguistic form, which has features of both written and spoken communication, or what Marcoccia (2004) refers to as "written conversation." It is very similar to conversation because it is highly interactive, as observed by common usage of "I" and "you" pronouns, "wh" questions, informal spelling, disclosure particles, and colloquial usage (Ferrara, Brunner & Whittemore, 1991). Yet it is also like written language in that it tends to be more elaborative and explicit than spoken (Baym, 1996, Ferrara et al., 1991). For example, in her analysis of agreements and disagreements found on an online message board about soap operas, Baym (1996) describes how users adopted the norm of being more specific in their messages as a way of circumventing the constraints of the asynchronous and text-based communicative environment. At the same time, Baym observed that the features of storing and quoting previous messages afforded users the ability to bypass the need for

the restatements that often occur in face-to-face conversations. Crystal (2006) describes digital language not as a simple hybrid of previous forms of language, but as a "third medium" of language that expands the availability of linguistic options, in the same way that new trends in fashion expand our availability of clothing options.

In computer-mediated communication via online discussion boards, use of language, in the form of written text, is an action. Using language at all brings participants into the social world of the discussion board by making presence known. Yet, without the use of language, one is simply engaging in a private act of reading. How, then, can the text-based messages found on a discussion board be considered either social or private, when they may be treated as both, or neither? Therefore, as an alternative to social interaction, it may be that computer-mediated communication is better conceptualized as a mode of language-in-use that researchers can attempt to better describe and understand for its communicative nature.

Literature Review

Research of online discussion boards, also referred to as Internet forums, message boards and, less recently, bulletin boards or newsgroups, conventionally refers to these research sites as "virtual communities" (i.e. Chiu et al., 2006; Dholakia et al., 2004; Kurabacack, 2005; Looi & Ang, 2000; Ridings & Gefen, 2004; Tremblay, 2005; Ye et al., 2006). The term "virtual community" owes its popularization to Howard Rheingold's 1993 book, *The Virtual Community*. As he describes it, "people in virtual communities do just about everything people do in real life, but we leave our bodies behind" (Rheingold, 1994, p. 3). As discussed previously, however, the extent to which online message board participation represents social processes and interaction is not clear. The term

"community," therefore, is necessarily problematized (Baym, 1998). Simply connecting people with technology does not necessarily constitute community in its traditional sense (Jones, 1998). Despite all of the research conducted on "virtual communities," researchers have yet to achieve consensus on what is meant by community (Johnson, 2001; Preece, 2000).

Baym (1998) argues that it is not just a matter of determining that members feel a "sense of community," but understanding and identifying the structures that provide that sense. This includes not only pre-existing structures like context, purpose, participant characteristics, and system infrastructure, but also emergent structures of community norms, identities, expressions, relationships (Baym, 1998), and social networks (Rice, 1987). Theoretical perspectives of the relationship between technology and social processes tend to fall on a continuum from a techno-centric point of view (Blau et al., 1976; Toffler, 1981), in which the focus is on the features of the technology and its impact on communicative structures, to an entirely relativist constructivist point of view, in which technologies can only be understood through the ways they are socially constructed and used (Grint & Woolgar, 1997). In the structurational perspective, a reasonable balance between these two opposing ends is found; technology is viewed as an open system, with structures both affected by and affecting the actions of humans who use it (Orlikowski, 1992). As a result, human uses of technologies can lead to both intended and unintended communicative constructions and consequences. In the present research, I do not attempt to define whether or not the discussion boards investigated represent virtual communities, cultures, or even interaction, but instead to present a

method that allows for the identification and description of their communicative structures as a means to determining what they do and do not represent.

Alternative Research Approaches

Previous research investigating the computer-mediated communication of discussion boards has largely been focused on understanding its macro-level features, such as aspects of learning, relationships, community, identity, or technological influences. Research methods for studying the messages and communicative patterns in these research sites have typically relied on exploratory approaches, such as ethnography (e.g. Baym, 1995; Ward, 1999), social network analysis (e.g. Aviv et al., 2003; Wasko et al., 2009), and discourse analysis (e.g. Paulus, 2007; Winzelberg, 1997). In an effort to get a more complete picture of phenomena from a variety of perspectives, many studies incorporate methodological procedures and instruments from more than one of these approaches for the purpose of analytic triangulation (e.g. An et al., 2009; Paccagnella, 1997; Wheeler, 2006).

Ethnography. Ethnographic approaches involve participant observation, describing communicative artifacts, surveying individual impressions, and sampling messages to offer qualitative descriptions of the activities of a particular research site. Through participation and engagement within a particular culture, the researcher is able to gain a better understanding of how meaning is constructed, and how the world is understood (Lindlof & Taylor, 2002). According to Hine (2000), "the aim is to make explicit the taken-for-granted and often tacit ways in which people make sense of their lives" (p. 5). The intrinsic value of online ethnographies, therefore, lies in the researcher's ability to gain insight into how participation is understood by the participants

themselves. For example, through an ethnographic study of a newsgroup dedicated to the discussion of soap operas, Baym (1995) discovered that the participants' engagements with the humorous retellings of plotlines made the experience of watching the show a more pleasurable one for them.

The method developed in the present research shares in common with ethnography the principle that observation is key to inductively understanding the features of communication. The practice of ethnography, however, pre-supposes that there is a cultural frame within which perspectives and practices are embedded. According to Duncan (1969), cultural practices are identified through observation of behaviors that represent communicative expectations, rules and norms. Approaching an online discussion board from an ethnographic perspective without first determining the extent to which such structures exist can lead to improper characterization of the origins of observed behaviors. The present method seeks instead to first describe the structures of communication by starting with micro-level observations of actions, choices and behavior. Once established, ethnographic approaches may be layered on to those structures in order to interpretively analyze their origins and outcomes.

Social network analysis. In social network analysis, researchers analyze communicative relationships, or links, between people and groups of people in order to determine the strength of the ties between those individuals or groups. Studying the social networks of discussion board users typically involves monitoring the messages posted, paying particular attention to their authors, in tandem with survey or interview instruments (Garton et al., 1997). For example, by mapping the social networks of a discussion board of a professional US legal association, Wasko et al. (2009) discovered

that a small segment of participants exhibited a high degree of centrality by managing and providing a majority of the knowledge resources.

Like the present research, social network analysis aims to provide a structural account of communication by identifying and mapping communicative patterns (Garton et al., 1997). By mapping a summary of communicative links between people, researchers adopting this approach offer an elucidating window into the connectivity of these networks. The underlying assumption of social network analysis is that such connectivity represents meaningful relationships between individuals. In computer-mediated communication, however, what is and is not meaningful cannot be assumed. The structural approach applied in this dissertation, therefore, sets its priority on mapping communicative actions at a microlevel, allowing for greater description of what they do and do not represent. The relationships discovered through social networks are important levels of analysis, but they can only be properly understood in context, after the sequence of activity has been mapped and the communicative structure has been identified (Cockett, 2000).

Discourse analysis. The origins of discourse analysis are derived from the linguistic perspective that language use involves the production of speech acts (Austin, 1962). According to this perspective, speech carries out social functions and stances, such as promising, requesting, declaring, etc., in addition to conveying information. With this focus, discourse analysis looks at the content of messages and codes the specific types of speech acts evident in those messages. Like the structural approach, this approach begins with the micro-level perspective of communication, allowing the researcher to understand more about the underlying structure of activity in a particular

research site. For example, by conducting a discourse analysis of a discussion board intended for individuals recovering from eating disorders, Winzelberg (1997) discovered that many of the activities observed served functions of requesting and providing support.

However, discourse analysis restricts itself to the functional property of utterances. It does not incorporate added communicative behaviors, nor does it consider the micro-level of context as potentially shaping the types of utterances produced. In computer-mediated communication, context takes precedence, as the very act of using language serves the function of announcing presence. When speech acts as a reference cue, it serves the function of indexing communication in the context of the immediate situation, rather than representing a function external to the immediate communicative situation. Features of the communicative environment, aside from utterances, may be overlooked or misunderstood. The functional coding of speech acts must be laid out sequentially and examined in terms of sequential contingency, in order to more generally connect those speech acts with aspects of structure.

Guiding Research Question

As the "third" medium for communication becomes more prevalent, communication researchers may be growing disenchanted with it as a novel environment for study (Herring, 2004). And yet, CMC has been integrated into our daily lives in such a short amount of time, researchers may be too quick to operate from taken-for-granted assumptions about what it represents. It is reasonable, therefore, to step back and problematize that which we call computer-mediated communication, and consider first the natural history of observed CMC. Without a structural foundation within which to contextualize research findings, CMC researchers are at risk of trying to understand the

perspectives, relationships, and behaviors of "poker players" without first understanding the rules of the game they are playing.

As noted previously, the approach outlined in this chapter emphasizes the mapping and analysis of structure, as observable in the organization of CMC. This approach seeks to address the grammatical and pragmatic properties of online communication, thereby holding off semantic analysis at the level of acts based on content. In other words, the structural approach emphasizes *context* not content as the goal of its analysis. This dissertation therefore seeks to address the guiding research question: *How can the processes of computer-mediated communication be analyzed, described and understood through a natural history, microanalytic, structural approach?*

As a distinct third medium, the unique features of discussion board communication must be considered in the analysis. Participants, free to enter and exit discussions as they please, do not all share the same time frame. As a result, their messages often create what Marcoccia (2004) terms "online polylogues." The flow of communication becomes staggered, branching off into multiple strings of related messages. In addition, the real-time lag that occurs between messages tends to fall within a wide range, from minutes to months (Marcoccia, 2004). These features of communication thus put into question the very relevance of flow as a linear sequential process. This means that it is difficult to determine, sequentially, who is addressing whom, or what is addressing what. In other words, because the entire thread of messages existing at a given point is available to all readers, the message(s) that compel(s) an individual to participate may be the first, last, or any in between. A participant's comment may address specifically a single message, the entire set of existing messages,

or bring up a new topic entirely, without any clear link to previous messages. Therefore, in addition to mapping the sequence of participant activity within the discussion board threads analyzed in this study, the analysis also examines indexical markers that allow inferences about who or what a specific message points back to, or refers to in response. These indexical or deictic markers make it possible to view the structure from a participant perspective, a perspective of choice, in how any given message links to prior activities in a site.

According to adaptive structuration theory, communication systems are designed with built-in rules or resources that become structures only when they are used (Poole & DeSanctis, 1990). Rules are "techniques or procedures," and resources are the tools that can be drawn upon in order to carry out tasks (Browning and Stephens, 2004). Rules and resources, however, are not directly observable. Instead, they are inferred from choices among patterns of actions. The use of references in discussion boards is an action that reflects a choice made to indicate indexical links to prior messages. Therefore, references are highly reliable deictic markers from which to account for the structure of communication. By presenting a method that specifically considers this aspect of CMC via discussion boards, this dissertation offers an analysis that allows messages to be mapped according to the unique characteristics of this communicative context.

In this dissertation I outline an approach that systematically examines the structural properties of CMC, defining the context in which discussion board communication is observed. This dissertation introduces an analytic vocabulary for mapping observed features and processes. It shows how these mappings may be further elaborated on when analyzing added questions about the nature of communication. The

method is both demonstrated and tested through the analysis of data from two distinct research settings. Comparing these two data sets shows how the method is able to produce interpretive findings that are general to both settings, as well as findings that highlight the differences between and within them.

Contributions of the Developed Method

The method presented in this dissertation provides a distinctive methodological starting point that provides researchers with a highly reliable and verifiable account of the micro-level structure of CMC within a specific research setting, onto which it is possible to layer additional types of coding of the data. This allows for an inductively grounded account of structure and individual differences in CMC. The method also allows for added coding methods to be anchored to the natural history. The method thereby grounds efforts to study and understand sense-making, identity, culture, relationships, social networks, individual differences, behaviors, outcomes and more, so that these phenomena may be understood in context.

In sum, this dissertation advances a structural method for describing the features of, and identifying the constitutive possibilities for, computer-mediated communication. This is achieved through the systematic mapping of the natural history of communication within a particular research site, in order to determine its organizing structure and individual differences via observable actions and choices. By placing values and variables of interest into determined sequences and structures, researchers will be able to interpret their occurrence in context, revealing more about their true nature.

Additionally, this dissertation offers a new analytical instrument that, like CMC itself, adapts and builds upon that which is utilized in more traditional modes of

communication, but creates something unique to this environment alone. In the present method, the indexical cues include the use of built-in tools provided by the system as well as emergent references that depend on individual choices. The mapping of these *reference cues* link discussion board message postings so that their organization can be described. Understanding features of CMC in this way can offer insights into system design and information architecture that address challenges faced by users in organizations (Alvesson & Kärreman, 2000), online learning (Tellent-Runnels et al., 2006), and knowledge management (King et al., 2002). Additionally, in its focus on finding ways to meet the challenge of representing multi-layered contingency data, the method suggests approaches to data representation which may prove useful to researchers studying communication and information systems that involve asynchronous and mediated actions by participants engaged in presumed "common" activity.

Organization of this Dissertation

This introductory chapter has highlighted the theoretical underpinnings that underlie both the conceptualization of computer-mediated communication via online discussion boards, and the methodological procedures applied in research both previously and presently. In the remainder of this dissertation, I expand upon, elucidate, and demonstrate the method introduced here. Chapter 2 describes the two research settings investigated, along with the sampling techniques tailored to those particular settings, and presents the research procedures for the developed method in detail, alongside the introduction and definition of related constructs. In Chapter 3, I use two examples from the data to demonstrate the process of mapping the natural history of discussion board threads. Chapter 4 examines a second level of structure, based on the links between

messages determined through the mapping of indexical markers, or references. Finally, Chapter 5 summarizes and discusses the method's practical application, and discoveries made. Chapter 5 also presents the limitations of the study, potential directions for further development of the method, and empirical questions to which it may be applied.

CHAPTER 2
RESEARCH SETTINGS AND PROCEDURES

Introduction

Implementing the structural approach to analysis, the method proposed here may be understood as an emphasis on context as constituted through participation and sequential order within discussion board threads. It first seeks to offer a unified description or account of discussion board threads based on their micro-level properties. The outcome from this analysis is a map of the natural history of participation in a thread. The method then utilizes references to determine and map a second level of structure based on indexical markers, which results in the creation of a map of the organization of participation in a thread. This chapter will first provide a description of two discussion board sites from which data were obtained for analysis, along with the procedures for sampling in these sites. It will then describe data collection, coding, and mapping procedures used in the developed method.

In the analysis, *posts* are understood to be meaningful acts that serve as building blocks of the structure of the entire thread. A post is the full block of text submitted by one participant at one time. It is represented on the discussion board within a delineated text box, along with the username of the individual who submitted it, and a time stamp indicating when it was submitted to the board. Posts are organized into threads within discussion boards. A thread begins when a post is submitted as a "new topic" on a discussion board. "New topic" posts establish the thread title. Generally, their text elicits added contributions to the thread by bringing up a question and/or topic for discussion. Subsequent posts in a thread are those posts submitted as a "reply" when viewing the

thread. The parameters of a discussion board thread, therefore, include the initiating "new topic" post and all subsequent "reply" posts submitted to it.

While the procedures for submitting these two types of posts on a discussion board are generally understood as "new topics" and their "replies," I have observed that there is consistently a small percentage of posts submitted which defy these characterizations, resulting in interesting reactions from other participants. When a "reply" post is submitted that is completely unrelated in any way to any prior posts, it may include a preface such as "Not to change the subject…," or "OT..." which stands for "off-topic." Without such a preface, other participants can be observed responding with comments such as, "And your point is?" or "What does this have to do with…?" "New topics" are also occasionally submitted with a reference or link to an already existing discussion board thread, acting as a response to that thread. These occurrences will often be replied to with remarks such as, "Why didn't you just respond in that thread?" Participant reactions have indicated that the system's built-in "suggestions" about what a post *should* represent act as descriptions of participative conventions and rules. While an interesting phenomenon for future study, the analysis of such convention-breaking situations was not a goal of the present research.

Research Settings

Data for this study was collected from two different discussion boards, each very different in purpose and environment. Both settings introduce a unique set of constraints on participation. These are described in the following sections.

The College Course

The first research setting from which data were obtained was an online discussion board used in conjunction with a one-semester college course that had one in-class session per week. The discussion board was a part of the course website, which also included class-related files, announcements, and assignments. Access to the website was restricted to the 16 students registered for the course and three classroom instructors. The discussion board was intended to provide a forum for discussing class- related concepts and topics. The 16 student users were required to participate in the discussion board as one component of this for-credit college course.

As the course teaching assistant, I began all discussion board threads with a "new topic" submission, introducing questions that students were required to address. The questions posed to the students dealt with topics and concepts central to the course. During the course's first six weeks, I posted two questions each week. During the following three weeks, I posted one question each week. I also occasionally submitted additional posts in response to technical glitches that arose. After the first nine weeks, students were no longer required to participate in online discussions, although an optional discussion board thread was started during week 10 for students wishing to continue their online engagement. Eight students posted to this thread from weeks 10 to 12, contributing one post each. Post contribution stopped completely after week 12.

Sampling

The analysis that follows makes use of the 10 discussion board threads resulting from the two questions posted each week from weeks two through six. The first week was not included because students were still enrolling in the class, and learning how to

access and navigate the discussion board. After week six, the discussion board experienced increasing instances of technical glitches that complicated the data. Specifically, participants began receiving error messages upon post submission, and their posts were not submitted. Some students dealt with this by instead sending emails to the course instructors including their comments; others gave up and mentioned it in class the following week. As a result, threads beyond week six were not included in the analysis presented in Chapters 3 & 4.

Participants

There were 18 participants (i.e. 16 students, one course instructor and myself) who posted over the course of the five weeks from which the 10 threads were obtained. As the course TA, I submitted the opening post in each of the 10 threads. I also submitted one additional post in response to a technical glitch. One of the two course instructors submitted five posts during this time frame, two in weeks two and four, and one in week five. The 16 student participants included 10 undergraduate, and six graduate level students. There were seven male and nine female students.

Students were trained in class on how to access and use the discussion board. They were informed that a combined score for online and in-class participation accounted for 20% of their final grades in the course. Each student was required to post at least one response to each of the week's questions, within three days of the time it was posted, and were expected to respond to two of their classmates' postings prior to class the following week. The manner and number of posts in which they accomplished this was not specified.

Ethical Considerations

Students were made aware that their posts would be observed and archived by the course instructors and myself. The nature of the data and research procedures was also reported to the Institutional Review Board for the Protection of Human Subjects. Approval to use these data for research was granted under expedited category 5, which allows for the use of previously collected data so long as it was originally collected for non-research purposes.

The Fan Site

The second research setting from which data were obtained was a discussion board situated on a public website that can be viewed by anyone with access to the Internet. This website is meant for fans of the popular television show *Lost.* Its main focus is the discussion board, although it also includes advertising and an online store selling products related to the show. The discussion board threads on this site are organized into several sections; each suggesting what topics might be addressed. The section titled "General Discussion" was the only section from which threads were selected for analysis. Five discussion board threads were chosen for study.

Participants

This website is very popular, and unlike the college course website with a fixed number of participants, there are thousands of active members who might potentially contribute to any discussion board thread. As stated, the posts observed on this website are viewable by the public. No researcher intervention or interaction was involved. Therefore, use of these data is not restricted by human subjects regulations. No information about the participants' demographics was available on the site, nor was such

information sought or collected. Usernames and the content of each post served as the basis of information available for inferring user identity. In order to contribute to the discussion board, all users were required to register for a free membership account. Membership required agreement by users to adhere to site rules and conventions in posting to the discussion board. Violations risked termination of membership based on the judgment of the discussion board's moderators.

Sampling

Threads chosen for analysis were selected according to the following procedure and criteria: over five consecutive weeks, on each day after a new weekly episode aired, new topics were tracked on the discussion board beginning at 8:00 A.M. (UTC -05:00 Eastern Time). The first discussion board thread to reach 15 posts for that day was chosen for analysis. With this approach, a thread was established within 5 hours each week. The five threads chosen for study were revisited every day for a period of seven days to account for new posts submitted. No posts were observed being submitted beyond 3.35 days during this time frame.

Summary of Research Settings

The data derived from the college class research setting will be referred to as "the student data," and the data derived from the fan site research setting will be referred to as "the fan data," in the remainder of this dissertation. The student data were obtained from a context that was restricted to a smaller set of users, in which participants met each other face-to-face on a weekly basis. Participation in this setting was required and specific to certain task demands. The fan data was obtained from a context that was open and public, meaning that anyone could view the discussion board and/or register a free account to

contribute posts. In addition, the fan data participation was voluntary and was not likely to involve in-person meetings among participants. The main similarity between the two research settings was that observed discussion board threads were created in response to a weekly generating event. The features of the two research settings from which data were extracted are summarized in Table 2.1.

Table 2.1

Summary of Research Setting Characteristics

	Student data	**Fan data**
Type	Educational	Recreational
Membership	Closed / Enrolled in class	Open / Free
Offline encounters	Yes	Unlikely
Participation tasks	Respond to question, respond to two others	None
Weekly generating event	Class assignment	Television episode airing

The range of discussion board thread characteristics of each research setting regarding number of posts, number of participants, and the total time frame in which they took place are summarized in Table 2.2.

Table 2.2

Summary of Discussion Board Thread Characteristics, by Research Setting

	Student data	**Fan data**
Number of posts	21-36	19-95
Number of participants	14-17	13-34
Total time frame (in minutes)	6057-13606	92-4823
Total time frame (in days)	4.20-9.45	.06-3.35

As Table 2.2 shows, the threads from the student data had more stability in their number of posts and participants, comprising smaller ranges of observed frequencies. The minimum numbers of posts and participants are comparable in the two settings, but their maximums are higher in the fan data. Despite the greater potential for participants to join threads from fan data, with its thousands of potential authors as opposed to only 18, the

maximum number of participants observed in a thread was only double that of the student data, despite its having three times as many posts.

The time frame of threads in the fan data was much shorter than those in the student data. As previously noted, all threads from the fan data ceased within 3.35 days. Alternatively, the threads from the student data continued for a minimum of 4.20 days, and in some cases even beyond the week in which they were due. The thread with the fewest posts in the fan data (n = 19) also occurred within the shortest time frame (92 minutes). Yet the thread with the fewest posts in the student data (n = 21) occurred within the longest observed time frame (9.45 days).

Research Procedures

The posts in these data are grouped and linked in varying ways into structural features of communication, termed *post sets*, *post pairs,* and *chains*. Posts may also be grouped by their *post author*, representing features of the individual participants.

Post Sets

The set of posts involved in a single discussion board thread is referred to as a *post set*. In this dissertation, the 10 post sets from the student data will be referred to as S1-S10 and the five post sets from the fan data will be referred to as F1-F5. Post sets are the site in which a natural history of posts is mapped, and their sequences and organization are determined.

Sequencing Posts

Posts in a post set are sequenced and mapped according to their order of occurrence in time. Posts in a thread are stamped with a record of the time they are submitted, and typically displayed in that order. This simplifies the process of coding

and sequencing posts for analysis, as demonstrated in Chapter 3. From the time stamp identifier, an additional variable known as *time lag,* or the number of minutes that passed since the most recent post submission, can be computed for each post.

Post Authors

The user who submitted a post is referred to as its *post author*. Each post author is assigned an absolute identifier (ID) spanning participation across observed post sets, as well as a post author value (PA) respective to participation in an individual post set. The value of a post author's ID stands in for their username, and is assigned to all posts across post sets. The value of a post author's PA is based on the order in which they joined, or first posted within a post set, and is assigned only to the posts in that post set. For example, ID9 was the 11th post author to join S2, but the 8th to join S3. Therefore, ID9 was PA11 in S2 and PA8 in S3. Looking at a post author's posts by both PA and ID proves extremely useful for investigating and determining individual differences in, and approaches to, online participation, as demonstrated in Chapters 3 & 4 of this dissertation.

Linking Posts

Sequencing posts by their order of occurrence is valuable, informative, and necessary for mapping a natural history. However, these sequences are unlike face-to-face two-party conversations, in which turns at talk are typically understood as pairs in sequence. Instead, links between specific posts may not be assumed based on sequential order alone. On discussion boards, all posts in a thread are displayed as an archive, and a submitted post can be understood as a response to any, none, or all of the posts preceding it. Post authors often *refer back to* prior post(s) or post author(s) in order to clarify the

statement(s) or individual(s) that they are addressing. Such conventions of posting may act as backwards-pointing indexical markers that allow for the establishment of links between posts. In order to understand the organization of posts in this way, such *reference cues* are used to code posts with information about the prior post(s) that it refers back to.

Reference Cues

Reference cues are features of a post that an individual post author may include in order to *refer back to* a prior post or post author. The inclusion of these cues within posts represent a choice made and action taken by the *referring* post author. As it is only the perceivable choices and actions made by individuals that the researcher has available to make sense of communication (Duncan, 1969), reference cues are treated as highly reliable deictic markers that show how post authors structure their engagement with messages and participants. The two types of reference cues utilized in the present analysis, *quoting cues* and *naming cues*, are described in the following sections.

Quoting cues. Quoting cues are present when a post author includes text from a prior post within their own post. Quoting tools enabling this type of cue are a common feature of discussion boards. Use of this tool automatically displays the text of the quoted post in a text box inside the post of the *referring* post author. Although the system automatically includes the entirety of the quoted post, the *referring* post author can edit out portions of the text prior to submission. Quoting cues can also be observed in instances where a post author has simply copied and pasted text from a previous post into their own post, without use of the quoting tool. Typically, such examples will present the quoted text as separate in some way from the *referring* post author's own text, by putting

it in quotes for example. Whether through use of the quoting tool, or the copy/paste function, the researcher can use "text matching" to find the prior post from which a quote was extracted.

Naming cues. Naming cues are instances in which a post author identifies a prior post through use of a previous post author's name within their post. In addition to a username or variant thereof, on discussion boards where post authors know each other personally, "real" names or relational identifiers may be used. There is no "naming tool" equivalent to the quote button provided by the discussion board itself, so all naming cues are simply written by the *referring* post author within the text of his or her post. It is important to read each post carefully when looking for naming cues, since names and usernames are sometimes abbreviated into an acronym or nickname version. The hypothetical username "SamsonsParadiseLost," for example, may be written as "SPL," "Samson," "Sam," "Sammy," and so on. The degree of certainty in establishing a reference back to a previous post author based on a naming cue will depend on the level of confidence that an acronym or nickname in fact represents the post author *referred to*.

A drawback to the use of naming cues as indexical markers is that they do not specifically indicate a prior post to which a backward reference may be established. This is not a problem in cases where the named post author has a single prior post in the post set. When there is more than one prior post, additionally present quoting cues that *refer back to* a post written by the named post author help to disambiguate the *referred to* post. There are two ways to handle the identification of *referred to* posts in instances with no additional quoting cues, and where the *referred to* post author has more than one prior post. The *referring* post can be understood as either *referring back to* the entire set of the

named post author's prior posts, or to that author's most recently submitted prior post. For purposes of simplification, the latter of these two options is chosen in the present analysis.

Reference cue coding scheme. To summarize, the scheme for coding posts according to their reference cues is as follows: quoting cues, the more precise reference cue, are coded as *referring back to* the specific prior post to which the quoted text was matched. Naming cues, the more ambiguous reference cue, are coded as *referring back to* the named post author's most recent prior post in the post sequence. If there is more than one type of cue in a single post, and they agree, it increases confidence in coding. If they do not agree, they are coded as *referring back to* separate prior posts. Findings from the data supporting the use of these reference cues' to establish references back to prior posts are presented in Chapter 4, and discussed in Chapter 5. The procedure for mapping posts based their references back to prior posts is described in the following sections and demonstrated in Chapter 4 of this dissertation.

Post Pairs

Once posts' backward-pointing reference cues are established and coded, new mappings of the natural history of posts can be created, representing a more emic account of participation in online discussion boards. Posts are then linked by these references, forming *post pairs*. Post pairs include the two posts associated with each reference cue, the *referring post* and the *referred to* post. A single post may be involved in any number of post pairs, depending on the number of prior posts it *refers back to*, and the number of following posts in which it is *referred to*. Once a post pair is established, a measure referred to as *sequential lag*, or the number of posts that were submitted, in sequence,

between the two linked posts, can be computed. Sequential lag is determined by subtracting the *referring* post location +1 from the location of the post it *refers back to*. For example, a post pair where a post at location 8 is linked to a post at location 7 would have a sequential lag of zero, since no posts were submitted between them. A post pair in which a post at location 18 is linked to a post at location 7 would have a sequential lag of 10, since 10 posts were submitted between them.

Post States

Posts can be categorized into one of four states based on the presence or absence of their references back to prior posts and/or of references to them in following posts. See Table 2.3 for a summary of these post states.

Table 2.3

Post State Categorizations

	Referred to in a following post	**Not referred to in a following post**
Reference back to prior post	Embedded posts*	Terminating posts*
No reference back to prior post	Originating posts*	Dormant posts

* active posts

Posts that are *referred to* in one or more following posts but do not *refer back to* any prior posts are termed *originating posts.* Posts that *refer back to* one or more prior posts but are not *referred to* in any following posts are categorized as *terminating posts*. Posts that both *refer back to* one or more prior posts and are *referred to* in one or more following posts are termed *embedded posts.* Originating, embedded and terminating posts are termed *active posts*, while posts that neither *refer back to* any prior posts nor are *referred to* in any following posts are termed *dormant posts*. Only active posts are involved in post pairs and chains.

Chains

When two post pairs share a common post, they are linked into *chains.* Chains can be comprised of only two posts from a single post pair when that post pair shares no posts with any others, or of multiple posts when post pairs are linked by shared posts. Each post pair in a chain must share in common one post with at least one other post pair in that chain, and each post pair can only be in one chain. The parameters of a chain are defined by its first originating post and its last terminating post, and include all posts of linked post pairs that occur in between. The configuration of a chain will depend on the manner in which post pairs are linked by shared posts. Because post pairs in a chain have the potential for multiple branching connections between their posts, chains are best displayed in a visual representation that maps these connections.

To illustrate the various ways the post pairs of a chain may be linked together, and the variety of post state possibilities within chains, a hypothetical post set of 10 posts shown in various chain configurations is introduced in the following maps. In these maps, the post locations run along the x-axis, and each post pair is represented as a bar along a point on the y-axis, connected by the two linked active posts, represented as gray boxes. Dormant, originating, embedded and terminating post states for all 10 posts in the post set are identified in the bottom row.

The simplest chains are those containing a single post pair. As previously stated, this occurs when a post pair contains an originating and a terminating post, neither of which are included in any other post pairs. See Figure 2.1 for an example where a terminating post at post location 5 *refers back to* an originating post at post location 2.

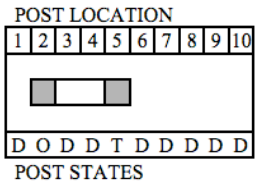

Figure 2.1. Chain example 1.

When embedded posts are included, they link post pairs, building a step-like formation in the chain. Figure 2.2 represents a chain made up of two post pairs that share a post at post location 5.

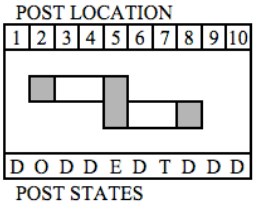

Figure 2.2. Chain example 2.

In the example presented in Figure 2.2, the originating post at post location 2 is *referred to* in an embedded post at post location 5, which is *referred to* in a terminating post at post location 8. In the map, the shared post is represented as an open box that extends its location between the two post pairs that it links.

If the originating post at post location 2 were to be *referred to* in two embedded posts, one at post location 5 which is *referred to* in a terminating post at post location 8, and the other at post location 6 which is *referred to* in another embedded post at post location 9, which is *referred to* in a terminating post at post location 10, it would look like the map shown in Figure 2.3.

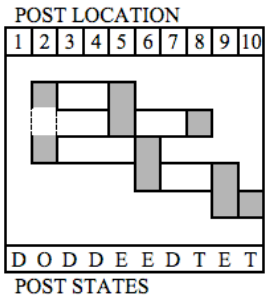

Figure 2.3. Chain example 3.

In Figure 2.3, note how the shared originating post at location 2 links two post pairs that are separated in the display by the post pair with a shared embedded post at location 5. In such instances, the axial representation becomes increasingly complex, but the most practical way to handle this is to give proximal preference to those post pairs linked by shared embedded posts. Posts that are shared by two or more post pairs but separated in the display by such prioritizations are indicated with one open box, linked by dotted lines running through the line of any of the post pairs represented between them.

There are occasions where a single post *refers back to* more than one prior post. When this occurs, post pairs that are linked by shared embedded posts, are again given precedence over proximity in the map. Figure 2.4 represents a situation similar to the chain example 3, with the exception that the terminating post at post location 8 also *refers back to* an originating post at post location 4.

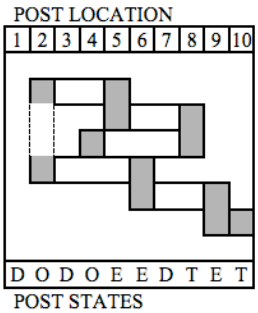

Figure 2.4. Chain example 4.

Finally, it is important to note that embedded posts are not required to build chains of more than one post pair. The only requirement is that more than one post pair share in common one post. Figure 2.5, for example, represents a series of originating-terminating post pairs where an originating post at post location 2 is *referred to* in terminating posts at post locations 5, 7 and 8.

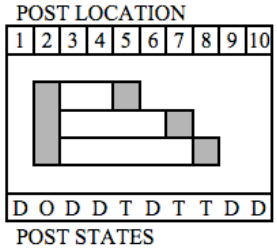

Figure 2.5. Chain example 5.

These hypothetical examples are intended to illustrate one approach to mapping such post pairs and chains as observed. The combination of linked post pairs that may be present in a single post set can be quite complex, and although guidelines are offered here, the representation of the most complex situations will have to be left up to the discretion of the individual researcher. Chapter 4 presents a version of these linked chain maps from actual post sets derived from the data, along with more detailed discussion of their use.

Summary of the Present Study

The method presented in this dissertation maps the natural history of online discussion board threads, using the unit of the post as the building block for structural analysis. Post sets include all of the posts in a discussion board thread. Posts are sequenced according to their order of occurrence, and active posts are organized into post pairs and chains. Post pairs, or pairs of posts linked by reference cues, and chains, or

groups of post pairs linked by shared posts, represent a second, more emic, level of structure. Differences in individuals' participation are analyzed by observing the posts of a post author within or across post sets. Both the sequencing and linking of posts are mapped in various ways, presented in Chapters 3 and 4. From these maps, the stream of action in the online discussion board threads can be analyzed and interpreted.

As this chapter has elucidated, reference cues are found in content, but they are indicators of context. Context in this analysis is understood not in the broader sense, but at the micro level. This may then be viewed as micro-context analysis of communication. Without directly applying semantic interpretation, researchers can identify the word(s) and designs of references that may be understood as indexical markers so they may be pragmatically assigned to that which they point back to. The reliability of these indicators will be related to the level to which they can be confidently linked to prior utterances. Once the researcher is satisfied that reference cues have been identified, tested, and mapped for a particular communicative situation, a new structure that lays out the base context for communication, onto which other interesting layers of data can be laid out, is determined.

The data collected for the present study included 15 post sets obtained from two different research settings. They total 553 posts contributed by 105 post authors. In all of the data, 367 post pairs and 79 chains were determined. See Table 2.4 for a summary of the relevant data derived from each research setting.

Table 2.4

Summary of Data Units, by Research Setting

	Student data		**Fan data**	
	Frequency	% of total	Frequency	% of total
Posts (*n* = 553)	311	56.2%	242	43.8%
Post authors (*n* = 105)	18	17.1%	87	82.9%
Post sets (*n* = 15)	10	66.7%	5	33.3%
Post pairs (*n* = 367)	220	59.9%	147	40.1%
Chains (*n* = 79)	44	55.7%	35	44.3%

A more in-depth presentation of the method's application, using examples from the data, is provided in conjunction with discussion of analysis procedures and findings in the next two chapters. Chapter 3 details the process of sequencing posts and Chapter 4 details the process of linking them. These analyses serve not only as examples for demonstrative purposes and testing grounds for the method, but also as investigative entry points into the data, from which insights can be inductively derived and tested. Considerations for the differing characteristics of each research setting are discussed and dealt with as they arise in these chapters. Chapter 5 presents an overall summary of the effectiveness of the method's application and a synthesis of the research findings, along with conclusions, implications, limitations and suggestions for future research.

CHAPTER 3
MAPPING THE NATURAL HISTORY OF ONLINE POSTINGS

Introduction

This chapter describes the approach to mapping the natural history of online posts' occurrence within a post set. As previously stated, a post set is the set of posts involved in a single discussion board thread. One post set from each research setting has been chosen from the data to provide an example of how the method is applied, and offer an investigative entry point for learning more about the data derived from these settings.

Examples from the Data: Post Sets S7 & F2

S7 is the post set from the student data used as an example in this chapter, and was chosen because it was the post set with the largest total number of posts. S7 had 36 posts, submitted by 17 authors over approximately five days. The posts were submitted during the fifth week of class, when the protocol for posting to the weekly discussion boards was well established. Its participants included all 16 students, one instructor, and me, the teaching assistant. I posted only the first post, a discussion question which opened the thread. The example post set from the fan data is F2, and was selected because it was the post set most similar to S7 in total number of posts, post authors, and total time span. F2 consisted of 27 posts by 16 post authors, and was collected during the second week of sampling. Posts were submitted over a period of approximately three and a half days.

Activity

The initial step to mapping a natural history of a post set is to code posts on the following four dimensions:

- POST LOCATION (PL): sequential location of post in relation to total posts in post set, 1 to n.

- POST AUTHOR (PA): total prior post authors +1
- POST NUMBER (PN): 1 to n, where n = total posts by that post author
- TIME LAG (TL): time, in minutes, since most recent prior post (rounded to the nearest minute)

These values are listed in order of the post location, in a post sequence map.

Table 3.1

Post Sequence Map, S7

PL	PA	PN	TL
1	1	1	
2	2	1	1196
3	3	1	99
4	4	1	68
5	5	1	3
6	5	2	5
7	6	1	11
8	4	2	21
9	7	1	203
10	8	1	48
11	4	3	873
12	9	1	80
13	5	3	43
14	3	2	23
15	6	2	48
16	10	1	100
17	11	1	316
18	4	4	18
19	8	2	1462
20	12	1	861
21	13	1	< 1
22	12	2	3
23	13	2	11
24	14	1	462
25	15	1	781
26	15	2	10
27	4	5	131
28	10	2	64
29	3	3	134
30	16	1	43
31	16	2	8
32	17	1	128
33	17	2	4
34	10	3	38
35	11	2	128
36	11	3	7

Table 3.1 displays the post sequence map for S7. Through representing post sets in this manner, researchers are able to derive information about individual posts, in the context in which they occurred. Appendix A presents the post sequence maps for all 10 student post sets and five fan post sets.

As observed in the Table 3.1 post sequence map, S7 demonstrates considerable variability in the time lag (TL) data of posts. The time lag values for this post set range from < 1 minute to 1462 minutes, with 23 of the 36 posts expressing time lags of ≤ 100 minutes, and 5 expressing time lags of > 500 minutes.

The first observable time lag in S7, located between post locations one (PL1) and two (PL2), is exceptionally long at 1196 minutes. In order to test whether this was a characteristic common to all post sets in the student data, all 10 of these post sets were compared, and time lags between PL1 and PL2 evaluated for each. These findings, along with the post sets' mean time lags, are presented in Table 3.2.

Table 3.2

Time Lag after PL1 and Mean Time Lag for S1-S10

Post set	**Time lag between PL1 and PL2**	**Mean time lag for post set**
S1	1657.00	683.30
S2	1628.00	377.53
S3	342.00	289.31
S4	2994.00	321.17
S5	24.00	466.52
S6	25.00	387.09
S7	1196.00	212.29
S8	1232.00	208.86
S9	304.00	266.65
S10	113.00*	293.97

* time lag between PL2 and PL3, because the teaching assistant initially posted twice to deal with a technical difficulty

As seen in Table 3.2, the mean time lag for the 10 student data post sets ranged from 208.86 to 683.30 minutes. Five post sets displayed a time lag of > 1100 minutes between PL1 and PL2. In each case all posts were submitted the day immediately following the first post submission, with the exception of S4, which extended two days. These 5 post sets are shaded in gray in Table 3.2. The remaining 5 post sets had a lag of < 350 minutes between PLs 1 and 2, and two of these displayed an exceptionally short time lag of < 25 minutes, which represents a considerable deviation from the mean of their respective post sets.

In this study, a unique numeric absolute identifier is attached to each individual's post submissions across post sets. This allowed for the observation that, in each post set where the time lag between PL1 and PL2 was < 350 minutes, the same individual, with the absolute ID of 14 (ID14), was the second post author (PA2).

Table 3.3

Time Lag after PL1 and Mean Time Lag for S1-S10, Revised for ID14 as PA2

Post set	Time lag between PL1 and PL2	Time lag between PL2 and PL3 in post sets where ID14 was PA2	Mean time lag
S1	1657.00		683.30
S2	1628.00		377.53
S3	342.00	1378.00	289.31
S4	2994.00		321.17
S5	24.00	1419.00	466.52
S6	25.00	1379.00	387.09
S7	1196.00		212.29
S8	1232.00		208.86
S9	304.00	998.00	266.65
S10	113.00*	136.00**	293.97

* time lag between PL2 and PL3, because the teaching assistant initially posted twice to deal with a technical difficulty

** time lag between PL3 and PL4, because the teaching assistant initially posted twice to deal with a technical difficulty

Table 3.3 serves to expand on the information contained in Table 3.2, by highlighting in gray those post sets that had a time lag extending into the following day, after either the first post or ID14's first post submission. As this table shows, nine of the 10 post sets in the student data displayed a time lag of > 900 minutes occurring between two of the first three post locations. These time lags again represent extensions into the day following that of the first post's submission. In the four data sets newly shaded in gray in Table 3.3, all PL2 submissions not extending into the following day were discovered to be posted by ID14.

Table 3.4

Post Sequence Map, F2

PL	PA	PN	TL
1	1	1	
2	2	1	6
3	3	1	2
4	4	1	1
5	5	1	1
6	1	2	1
7	2	2	3
8	5	2	< 1
9	1	3	4
10	6	1	< 1
11	3	2	2
12	7	1	2
13	1	4	1
14	5	3	1
15	8	1	3
16	1	5	1
17	3	3	2
18	9	1	235
19	10	1	10
20	9	2	1
21	11	1	25
22	3	4	8
23	12	1	70
24	13	1	17
25	14	1	17
26	15	1	4319
27	16	1	91

To look at time lag in the fan data, the post sequence map for F2 is laid out in Table 3.4. Posts were again coded according to four dimensions: post location (PL), post author (PA), post number (PN), and time lag (TL). As this table shows, the time lag in F2 generally appears to be much shorter than what was observed in S7. Nineteen of the 27 posts had a time lag of ≤ 10 minutes. Only two posts had a time lag of > 200 minutes. The time lag at PL26 is at the extreme, measuring 4319 minutes. To learn more about differences in time lag between the two data sets, the average time lag of posts were compared by data set, and are presented in Table 3.5.

Table 3.5

Descriptive Statistics for Posts' Time Lag, by Data Set

	Total # posts	**Time lag**	
		Mean	**SD**
Student Data	311	352.57	622.98
Fan Data	242	55.01	356.17

As Table 3.5 shows, posts from the fan data had much shorter average time lags than posts from the student data.

An additional observation that can be ascertained from the Table 3.4 post sequence map, is that the first 17 posts in F2 were submitted within 30 minutes of the first post submission, all with very short time lags. In order to determine if a grouping of posts with time lags ≤ 10 minutes typically followed the first post in the fan data, all five post sets from this data set were investigated. The summary of these findings is presented in Table 3.6. The third column in this table shows that time lags of ≤ 10 minutes followed the first post in all five of the fan data post sets. Additionally, four of these post sets demonstrating runs of short time lag, ranging from 2 posts in F1 to 27 posts in F4, are highlighted in gray in Table 3.6.

Table 3.6

Runs of Time Lag of ≤ 10 Minutes at the Beginning of Fan Data Post Sets

Post set	Total # posts	Total # of posts following PL1 with a TL ≤ 10 minutes	First observed TL of > 10 minutes
F1	35	2	29
F2	27	16	235
F3	95	27	17
F4	66	1	46
F5	19	17	33

Despite the longest observed run occurring in F4, only the runs in F2 and F5 were found to encompass a majority of the post set's posts. As column four in Table 3.6 demonstrates, three out of four of the short time lag post runs, in addition to the single post in F4, are followed by a time lag of < 50 minutes. F2 is the only post set where this was not the case, with a following time lag of 235 minutes.

The observation that time lags of early posts in the fan data were particularly short stands in contrast to what was observed at the beginning of post sets in student data, with early time lags lasting into the following day. These wide-ranging differences led to the empirical observation that the relative distribution of posts, in time, might be different in the post sets from the two research settings.

To determine if this was the case, the total time frame of each post set was split into ten standardized segments. The bar graphs in Figure 3.1 represent the standardized time distribution of posts in each data set. As these graphs show, the distribution of posts in the student data was more evenly dispersed throughout the post sets' time frames, with minor spikes occurring in the middle, and the highest proportion occurring in the final segment. Alternatively, the fan data demonstrated the highest proportion of posts

occurring in the beginning post set time segments, with a slight resurgence in post activity at the terminal segment.

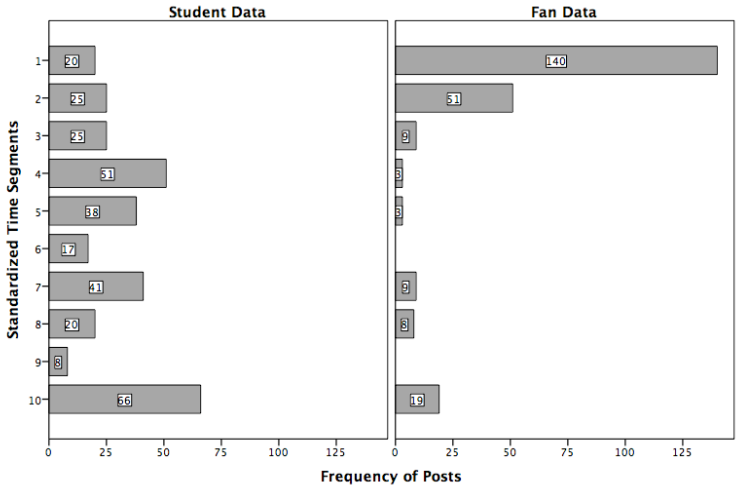

Figure 3.1. Distribution of posts, in standardized time, by data set.

To determine if these participation spikes may correspond with common times of the day in which posting activity was more or less frequent, all posts were investigated for the hour in which they were posted according to UTC -05:00 Eastern Time (US & Canada). Figure 3.2 represents the distribution of daytime hours for which posts were submitted in both the student and fan data.

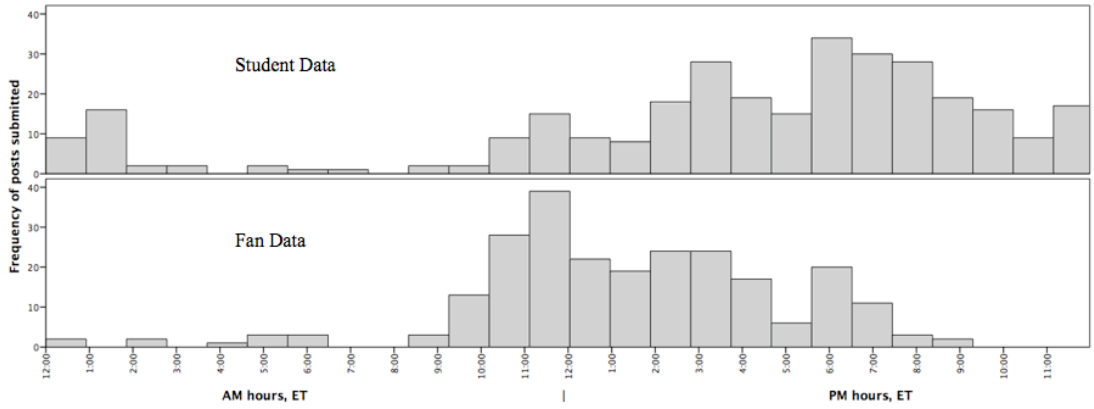

Figure 3.2. Distribution of posting activity during Eastern Time zone hours, by data set.

The bar graphs presented in Figure 3.2 show that participants in the two data sets displayed certain tendencies towards when they submitted online posts. In both data sets, posts were less likely to be submitted between the hours of 2:00 A.M. and 8:00 A.M. (UTC -05:00 Eastern Time). This observation was not surprising for the student participants, considering that these hours typically correspond to a time period during which they may be expected to be asleep in the Eastern Time zone. Posts in the student data were generally distributed across a wider range of hours, mostly occurring between 10:00 A.M. and 2:00 A.M. (UTC -05:00 Eastern Time). The greatest proportion of posts in this data set occurred during the 6:00 P.M. (UTC -05:00 Eastern Time) hour.

It is unknown which time zones the fan data participants resided in, but because *Lost* is an American TV show, and thread sampling was based on the airing dates in America, it may be assumed that the majority of these participants were residing within the United States. It should be noted that, while these discussion boards incorporate customized time features representative of the time zone from which a user is accessing the site, they do maintain a fixed forum time based on the location of the server, resulting in all message postings being viewed in real time after their submission. The lowest hours of posting activity between 2:00 A.M. and 8:00 A.M. (UTC -05:00 Eastern Time), therefore, generally correspond to a time period during which participants across the country may be expected to be sleeping. Specifically, these low-contributing hours would represent 11:00 P.M. to 5:00 A.M. (UTC -08:00 Pacific Time), 1:00 A.M. to 7:00 A.M. (UTC -07:00 Mountain Time), and 12:00 A.M. to 6:00 A.M. (UTC -06:00 Central Time). As seen in Figure 7, posts primarily occurring between the hours of 10:00 A.M. and 7:00 P.M. (UTC -05:00 Eastern Time) in the fan data were more concentrated among a smaller

span of hours than the corresponding student data posts. Considering time zones again, these hours equate to 7:00 A.M. to 4:00 P.M. (UTC -08:00 Pacific Time), 8:00 A.M. to 5:00 P.M. (UTC -06:00 Central Time), and 9:00 A.M. to 6:00 P.M. (UTC -07:00 Mountain Time). The greatest proportion of posts was submitted during the 11:00 A.M. (UTC -05:00 Eastern Time) hour.

Participants

Returning to ID14, the post author who uniquely posted on the same day of the first post's submission in four of the post sets from the student data, it can be observed that certain post authors have tendencies or preferences in their approaches to posting. ID14, for example, displayed a tendency for contributing to post sets early. In addition, ID14 was the highest contributing post author observed, contributing to nine of the 10 post sets. While the mean number of total posts across post sets was 17.28 (SD = 8.34) for post authors in the student data, ID14 contributed a total of 38. It can be stated, therefore, that ID14 displays a uniquely enthusiastic approach to posting; one in which posts are submitted early, and often.

By looking at the IDs of all post authors from the student data, it was determined that one additional individual, ID13, can also be classified as an enthusiastic post author. ID13 also contributed to nine of the 10 observed post sets, with a total of 25 posts. Neither ID13 nor ID14 contributed to S1, the post set in the student data with the fewest total posts. Both of these individuals adopted the "post early post often" approach in a majority of the post sets contributed to. Table 3.7 displays the characteristics for these two post authors, as well as their post author (PA) values for each post set, identifying the order in which they joined by contributing a post.

Table 3.7

Characteristics of Enthusiastic Post Authors in the Student Data

Absolute ID	Total # of post sets contributed to	Total # of posts, across post sets	Post sets contributed to	PA value per post set
13	9	25	S2	4
			S3	3
			S4	3
			S5	3
			S6	3
			S7	5
			S8	4
			S9	3
			S10	3
14	9	38	S2	7
			S3	2
			S4	11
			S5	2
			S6	2
			S7	4
			S8	5
			S9	2
			S10	2

As Table 3.7 shows, in all but one instance both of these individuals were amongst the first half of the set of post authors to contribute posts, as indicated by their PA values. In addition, the two enthusiastic post authors identified in the student data were two of the students receiving the highest grades for the course, even after the online participation component of grade calculation was removed.

Enthusiastic post authors were looked for in the fan data as well. The fan data was sampled from a much larger corpus of discussion board threads, and the mean number of total posts across all post sets for post authors in the fan data was only 2.79 (*SD* = 3.82). It was surprising, therefore, to discover that 13 individuals, three of whom had more than 15 total posts, participated in more than one of the sampled threads. In the fan data, these three individuals are identified as enthusiastic post authors, having post

contributions in more than one collected post set, and contributing a total of > 15 posts total across all post sets. Table 3.8 presents characteristics for these three post authors.

Table 3.8

Characteristics of Enthusiastic Post Authors in the Fan Data

Absolute ID	Total # of post sets contributed to	Total # of posts, across post sets	Post sets contributed to	PA value per post set
41	4	19	F1	23
			F2	5
			F3	3
			F5	4
50	3	23	F2	9
			F3	11
			F4	5
73	2	16	F3	19
			F4	4

The PA values presented in the fifth column of Table 3.8 identify ID41 as being among the first half of post authors in three of the four post sets contributed to, whereas IDs 50 and 73 were among the first half in only one post set each. It can also be observed from Table 3.8 that each post set in the fan data had at least one contributing enthusiastic post author, with post set F3 consisting of three, and both F2 and F4 having two each.

Based on the observations of ID14, it was hypothesized that the posts of enthusiastic post authors would, on average, have shorter time lags as compared to other posts in their respective data sets. The Figure 3.3 bar graph comparing mean time lags for the posts of enthusiastic post authors versus other post authors, by data set, shows that enthusiastic post authors' posts in both data sets tended to have a shorter time lag than other posts. This is especially prevalent in the fan data, as shown in Figure 3.3.

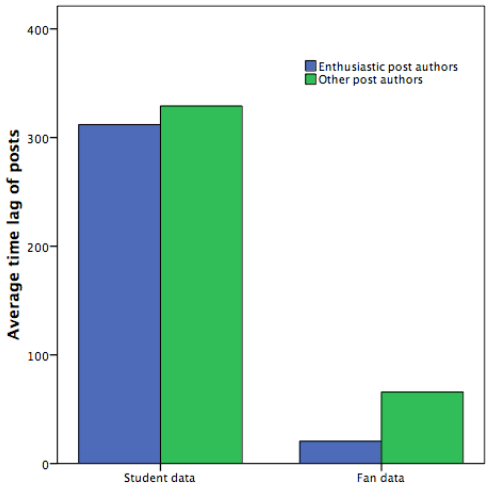

Figure 3.3. Mean time lag of posts of enthusiastic post authors versus other post authors, by research setting.

A post author sequence map is useful to discover more about individual differences in post author participation. Figure 3.4 shows a map of this type for S7, visualizing the sequence of posts in a post set by individual post author.

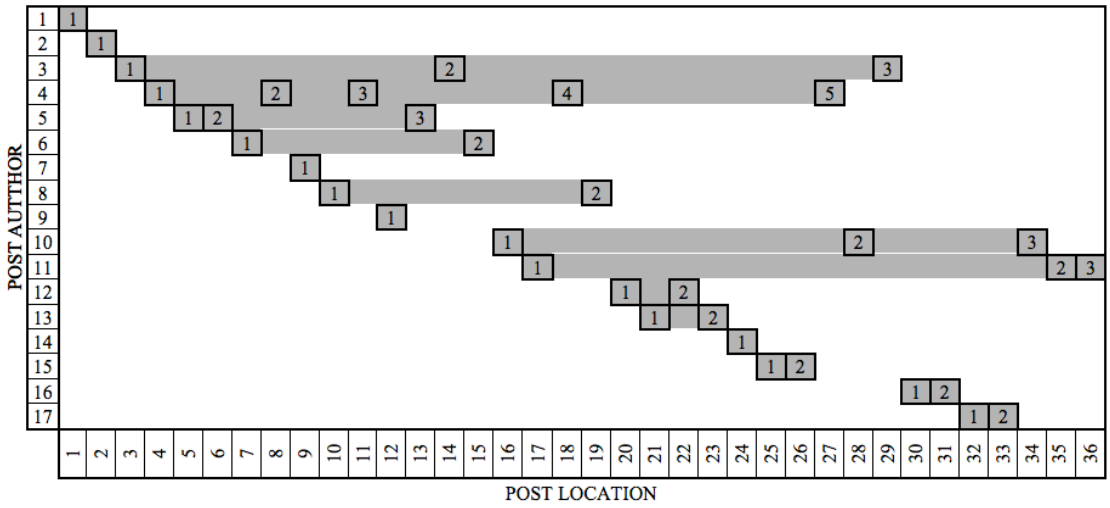

Figure 3.4. Post author sequence map, S7.

In the post author sequence map, the x-axis displays post locations, the y-axis post authors, and outlined gray boxes represent individual posts. The numbers in the gray boxes represent the post number for that individual post author, and the gray bars running

between the post boxes represent the span of post locations that occur between an individual post author's posts in the post set's sequence.

The post author sequence map in Figure 3.4 highlights the activity of the individual post authors in a post set, allowing observations of their individual differences to be made more easily. It can be observed from this map that five post authors in S7 posted once, seven posted twice, four posted three times, and one posted five times. Both of the enthusiastic post authors from the student data posted in this post set, and were among the highest contributing post authors. ID14 can be observed as PA4 in this post set and ID13 can be observed as PA5.

It can be also observed in the map in Figure 3.4 that five of the seven post authors who posted twice, (PA12, PA13, PA15, PA16 and PA17), did so in a very short time span and in close sequence, possibly indicating an additional posting approach, in which post authors submit the entirety of their posts at one time. To learn more about such an approach, it is useful to map additional posts according to their occurrence in time. The map in Figure 3.5 is termed a post time sequence map, and displays posts according to their placement within a post set's total time frame. In order to show the most detail without losing the visual quality of these maps, I split the time frame into ten standardized segments, each containing four quadrants. The post time sequence map for S7 is presented in Figure 3.5. In this map, posts are again represented as outlined gray boxes, but with numbers corresponding to the post author (PA) value for the post. The y-axis displays the post location, and the x-axis displays the ten segments of the post set's time frame. Gray bars represent a projection of time into the future, extending through the segment in which a post author next posts.

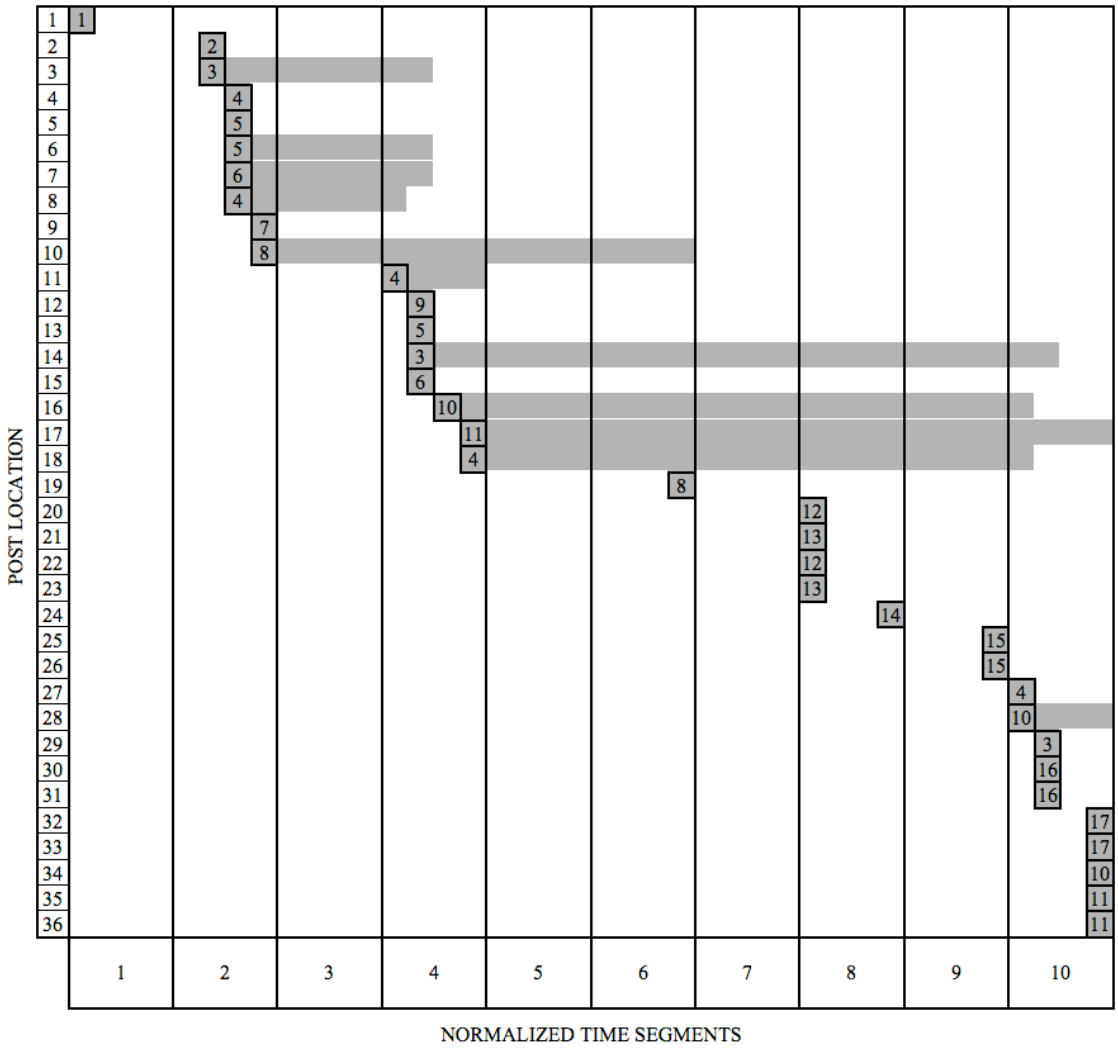

Figure 3.5. Post time sequence map, S7.

The post time sequence map in Figure 3.5 shows that in S7, the posts of all five previously identified two-post authors, (PA12, PA13, PA15, PA16 and PA17), submitted their two posts during the same standardized time segment. This supports the hypothesis that they had a single contributing visit to the discussion board thread, while also serving to solidify the identification of an additional posting approach. Post authors adopting this approach will be referred to as *transient post authors*, and defined as any post author having more than one post, and who submit the sum total of their posts in a comparatively short time span.

In looking at all 10 post sets in the student data, it was discovered that a total of nine had at least one transient post author, with S7 containing the most. Table 3.9 summarizes the characteristics of transient post author participation within the post sets of the student data.

Table 3.9

Frequency and Proportion of Transient Post Authors in the Student Data

Post set	Total number of post authors	Number of transient post authors	Proportion of post authors that were transient	Transient PA values
S1	14	1	.071	14
S2	17	4	.235	11, 14, 16, 17
S3	15	4	.267	7, 8, 12, 15
S4	14	4	.286	8, 12, 13, 14
S5	14	1	.071	14
S6	15	1	.067	15
S7	17	5	.294	12, 13, 15, 16, 17
S8	17	2	.118	12, 13
S9	15			
S10	14	2	.143	5, 12

The PA values in column five of Table 3.9 indicate that in a majority of cases in which a post author was transient, they were also among the latest post authors to join the post set. Of the 24 observed instances of transient post authors, 20 had a PA value of 11 or higher.

By looking at the transient post author IDs, it was observed that 10 individuals produced all 24 instances of transient posting. Table 3.10, which provides a summary of the characteristics of individuals' transient post author participation, indicates that while some post authors adopted a transient posting approach in only one post set, others adopted it for multiple post sets. In addition, the two post authors who contributed to the least number of post sets overall, also adopted the transient approach in a majority of those that they contributed to, with post author ID12 adopting this approach for all post

sets contributed to. Highlighted in gray in Table 3.10, these two post authors were also among the lowest contributing post authors observed in the student data.

Table 3.10

Frequency and Proportion of Transient Post Author Characteristics in the Student Data

Absolute ID	Total number of post sets contributed to	Number of post sets where transient	Proportion post sets where transient	PA value when transient
2	8	1	.125	17
3	10	2	.200	12, 12
5	10	1	.100	7
6	8	1	.125	15
9	6	5	.833	11, 8, 8, 12, 12
10	8	3	.375	13, 13, 5
11	9	3	.333	14, 13, 16
12	4	4	1.000	14, 14, 15, 17
17	7	1	.143	16
18	8	3	.375	15, 14, 12

To learn more about posting approaches observed in the fan data, F2's post author sequence map, presented in Figure 3.6, is a useful starting point.

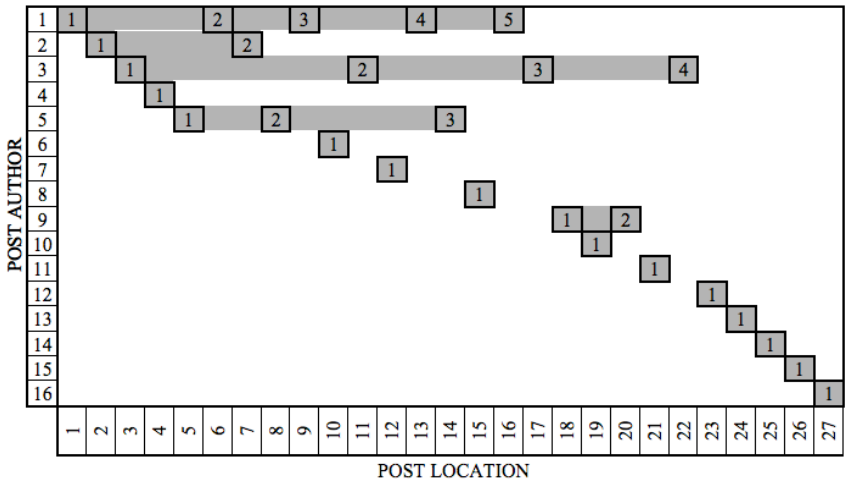

Figure 3.6. Post author sequence map, F2.

As the map in Figure 3.6 indicates, the first post author (PA1), had more posts than any other post author, and contributed to the post set until beyond its midpoint. To determine

if this was common among first post authors in the fan data, the activity of PA1 in the remaining four post sets from the fan data was analyzed, the results of which are presented in Table 3.11.

Table 3.11

PA1 Characteristics in Post Sets from the Fan Data

Post set	Total number of posts in post set	Number of posts by PA1	PL of PA1's last post in the post set
F1	35	2	6
F2	27	5	16
F3	95	13	75
F4	66	2	10
F5	19	2	7

The rightmost column in Table 3.11 shows that, in both F2 and F3, PA1 contributed posts until beyond the post set's midpoint. These post sets are highlighted in gray in Table 3.11. In the remaining three post sets from the fan data, PA1 only contributed one additional post, which in all three cases was submitted prior to the post set's midpoint. Thus, the first post authors observed in the fan data appeared to take either a vested approach to the post sets they started, by contributing a high number of posts beyond its midpoint; or a detached approach, only returning to the post set to contribute once, and well before the midpoint.

The post time sequence map in Figure 3.7 offers more insight into the time frame of participation of the post authors in F2. F2's post time sequence map shown in Figure 3.7 presents very differently from that of S7, shown in Figure 3.5. This is largely due to the relatively long time lag of 4319 minutes observed in F2, between post locations 25 and 26. Compared to other posts in F2, this represents an increase of more than 20 times

the longest observed time lag. Such relatively long breaks in standardized time were also observed in post sets F1 and F4.

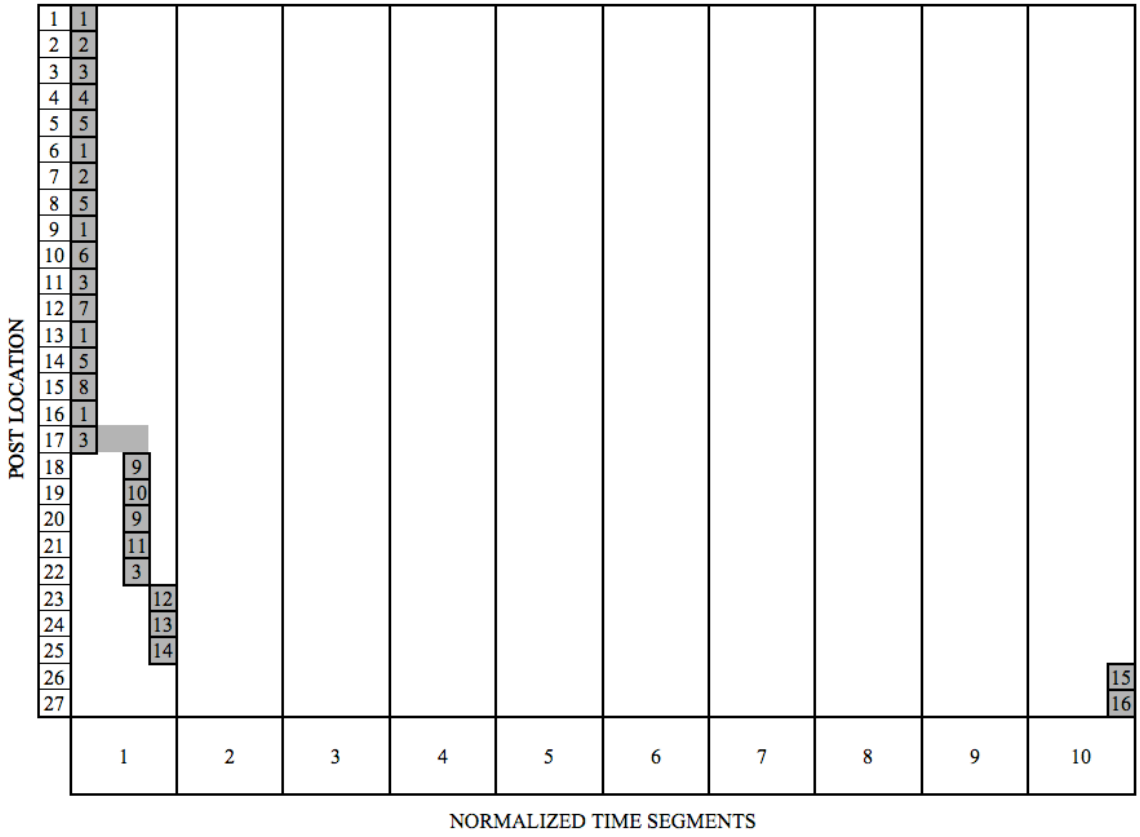

Figure 3.7. Post time sequence map, F2.

It can be observed in the post time sequence map for F2 presented in Figure 3.7 that, after the run of 17 posts submitted in the first 30 minutes, a single post author, PA3, returned to contribute to the post set, doing so just once. Of the eight additional post authors who only contributed to the post set after that point, six posted just one time. Returning to the post author sequence map for F2 in Figure 3.6, this terminal string of one-post contributing authors is clearly identified, represented by a single diagonal line of boxes at the end of the post set.

The post time sequence map in Figure 3.7 shows F2 to have three distinct clusters of posting activity occurring in shared or consecutive segments of standardized time.

These will be referred to as *posting spurts*. Posting spurts may be observed in post time sequence maps by searching for strings of successive posts, which occur within common or consecutive standardized segments of time. In F2, posting spurts first occur between post locations 1-17, next between 18-25, and third between 26-27. A review of the post time sequence maps for all fan post sets reveals a total of 17 observable posting spurts in the fan data.

In the post time sequence map presented in Figure 3.7, the posting spurts in F2 were observed to be initiated by, and primarily involved, new post authors to the post set. The exception was PA3, who was re-engaged in the second posting spurt, but not the third. To see if such "newcomers" were typically found in secondary and later posting spurts from the fan data, all 13 posting spurts that did not begin with the first post, were looked at. It was discovered that all 13 of these posting spurts were initiated by newcomer post authors. In just five of these instances, post authors from earlier in the post set were re-engaged in the posting activity. In all 13 cases, the posting spurt comprised mostly of posts by additional newcomer post authors.

The newcomers joining post sets in a new time frame and regularly restarting posting activity, are referred to as *restarting newcomers*. Newcomer post authors who follow restarting newcomers are referred to as *invited newcomers* because they are, in a sense, "invited" to contribute by a restarting newcomer's post. Post authors involved in a "restarted" posting spurt who were also observed engaging in earlier posting activity are termed *re-engaged post authors*.

In addition to the 13 posting spurts initiated by restarting newcomers, three instances were found in which new post authors joined a post set in a new time frame, but

no posting spurt was formed, as no following posts were submitted in the same or consecutive time frame. Therefore, there were a total of 16 observed cases of a restarting newcomer posting approach, 3 of which were unsuccessful at restarting posting activity. The associated values and variables for these 16 cases are presented in Table 3.12.

Table 3.12

Restarting Newcomer, Invited Newcomer, and Re-engaged Post Author Characteristics in Post Sets from the Fan Data

Post Set	Restarting newcomer PA value	Restarting post location	Restarting post time lag	Invited newcomer PA values	Re-engaged PA values
F1	17	26	1965	18, 19	9
	20	30	185	21, 22	
	23	33	680	24, 25	
F2	9	18	235	10, 11	3
	15	26	4319	16	
F3	16	44	54	17, 18, 19, 20, 21	1, 2, 3, 4, 6
	22	74	88	23, 24	1, 7
	25	79	199		
	26	81	57	27, 28	
	29	84	259	30, 31, 32, 33, 34	3, 5, 9, 14, 19
F4	17	66	2802		
F5	2	2	6	3, 4	
	5	5	4	6	
	7	8	9	8, 9	3
	10	13	10	11, 12	4
	13	19	33		

As the last two columns in Table 3.12 show, the restarting newcomers in F3 were more successful in inviting new post authors and re-engaging earlier post authors than those in the other post sets. In addition, F4 was the only post set in which no restarted posting spurts occurred.

The absolute IDs of restarting newcomer post authors in the fan data revealed that ID55 was the only participant to adopt this approach in two post sets. The first post is at PL26 in F2, and the second at PL66 in F4. Interestingly, these are the only two posts

contributed by ID55 in the observed post sets, and they are also the two posts with the longest time lag of all posts from the fan data. In one of these instances, a single invited newcomer posted one time, and no earlier post authors were re-engaged in either instance. This may indicate an extreme or unique variant of the restarting newcomer approach to posting.

Summary

The coding and mapping of sequences of posts in S7 and F2 has shown how this method may be used to highlight non-obvious features of communication in discussion board threads. It allowed for the discovery of features previously unnoticed within the two post sets, and individual differences in posting approaches. Additionally, it led to the formulation of new insights into the nature of participation within the two data sets.

Discoveries were also made regarding the nature of participation with respect to posting activity in time. For example, tendencies for posting during certain times of the day were discovered in both data sets. Students tended to submit posts in the late afternoon or early evening, while the fans tended to submit more often during morning and afternoon hours. In both research settings, relatively few posts were submitted between 2:00 A.M. and 8:00 A.M. (UTC -05:00 Eastern Time).

In addition, the distribution of posts in a post set's time frame varied between the two data sets. Post sets from the fan data tended to have more posts early on in the time frame. Those from the student data expressed a more even distribution of posts throughout the time frame, with slightly more occurring in the final time segment.

It was found that the time lags of posts from the student data were longer, on average, than those from the fan data. Specifically, differences were observed with

respect to time lag in the beginning of post sets. It was found that post sets from the student data had longer time lags after the first post or after the second post in instances where an enthusiastic post author posted second. Alternatively, post sets from the fan data tended to begin with very short time lagged posts, typically in runs of successive posts. Posting spurts, which are clusters of posting activity in shared or consecutive segments of standardized time, were discovered in the post time sequence maps. In the fan data, posting spurts following those started by the first post of a post set were always initiated by a new post author, and they generally involved a set of mostly new post authors.

Several posting approaches were discovered by observing individual differences in participation. Enthusiastic post authors posted more often, earlier, and across a wider range of post locations and post sets than others from their research setting, and were identified in the data from both research settings. Transient post authors were identified only in the student data. This approach was adopted more often by low contributing post authors, as they were observed submitting the whole of their posts in quick succession. Two first post author approaches were discovered in the fan data: a vested approach, in which the first post author contributed most often to the post set and beyond its midpoint; and a detached approach, in which a contribution of a single additional post was made well before the post set's midpoint. Finally, restarting newcomer post authors were identified in the fan data. These are post authors who join post sets in a new time frame and generally initiate new posting spurts, involving mostly new post authors (invited newcomers) but sometimes re-engaging earlier participants (re-engaged post authors).

In the next chapter, the process of linking posts and creating post pairs and chains will be demonstrated, allowing for additional observations and discoveries to be made. Subsequently, Chapter 5 presents a discussion that serves to synthesize the findings from Chapters 3 and 4 to determine final conclusions, both about the data analyzed in this study, as well as the methodological procedures developed and applied.

CHAPTER 4
MAPPING THE NATURAL HISTORY OF LINKED POSTS

Introduction

The natural history of posts sequenced according to recorded time stamps offers a map of post authors' participation in a post set, but does not describe individual post linkages. Mapping the natural history of linked posts requires further coding to include the indexical or deictic markers within posts, which point back to a prior post or post author. Only by mapping out indexically linked posts is it possible to talk of post sets as interactional.

In addition to demonstrating the process of coding indexical links and using them to create post pairs and chains, it was a goal of this analysis to test the use of reference cues as indexical markers. This chapter presents the discoveries and findings that resulted from the observation and analysis of linked post mapping. One post set from each research setting has been chosen from the data to offer an investigative entry point for learning more about the data derived from these environments.

Examples from the Data: Post Sets S1 & F5

The process for coding and mapping indexically linked posts is inherently more complex than a time-based natural history. In the interest of clarity, the two post sets used as examples in this chapter were those with the fewest total number of posts. The post set chosen from the student data was S1, consisting of 21 posts submitted by 14 post authors over approximately nine and a half days. The post set chosen from the fan data was F5, consisting of 19 posts, submitted by 13 post authors over a total of 92 minutes. It was collected during the final week of sampling for this research setting, and was the only post set lasting less than one day.

Determining and Coding Linked Posts

In order to determine the organization of participation in the observed discussion board threads, indexical links between posts are layered on to the natural history of post sets. To do this, additional values representing references to prior posts must be coded. In the linked post map, posts are listed according to the following values:

- POST LOCATION (PL): sequential location of post in relation to total posts in post set, 1 to n.
- POST AUTHOR (PA): total prior post authors +1
- POST NUMBER (PN): 1 to n, where n = total posts by that post author

In cases where a reference cue is present, posts are additionally coded for the following:

- LINKED POST LOCATION (LPL): location of referred to post
- LINKED POST AUTHOR (LPA): post author of referred to post
- LINKED POST NUMBER (LPN): post number of referred to post
- REFERENCE CUE TYPE (RCT): Q for quoting, N for naming and Q/N for combination quoting/naming cues

As with the post sequence maps presented in Chapter 3, in linked post maps, all posts are listed in sequence. Mapping linked posts involves the use of additional representative dimensions to portray the added structural level that they describe. Posts that are *referred to* in following posts are highlighted in bold. Posts that *refer back to* more than one prior post are additionally displayed in italicized text. Pairs of posts linked by reference cues, or post pairs, are highlighted in gray. The linked post map for S1 is contained in Table 4.1. The linked post maps for the 10 student data post sets and five fan data post sets are presented in Appendix B. Listing the posts in this manner allows the researcher to make observations about the features of linked posts. It is from these initial observations that an additional level of structure can be identified, and the description of interactive processes can emerge.

Table 4.1

Linked Post Map, S1

PL	PA	PN	LPL	LPA	LPN	RCT
1	1	1				
2	2	1				
3	**3**	**1**				
4	**4**	**1**				
5	**5**	**1**				
6	2	2	3	3	1	Q
6	*2*	*2*	4	4	1	Q
6	*2*	*2*	5	5	1	Q/N
7	**6**	**1**				
8	**7**	**1**	7	6	1	Q/N
9	8	1				
10	8	2	4	4	1	N
11	**9**	**1**				
12	**4**	**2**	11	9	1	Q/N
13	10	1				
14	**9**	**2**	12	4	2	N
15	**5**	**2**	8	7	1	Q/N
16	4	3	14	9	2	N
16	*4*	*3*	15	5	2	Q/N
17	**11**	**1**				
18	12	1				
19	13	1				
20	14	1				
21	14	2	17	11	1	Q

Activity

Looking at the first posts to be *referred to* (in bold), and the first posts *referring back* (outlined) in sequence indicates when referencing activity began in a particular post set. As seen in Table 4.1, the first post to be *referred to* in S1 was at post location (PL) 3. The first post to *refer back to* a prior post was at PL6.

To determine when referencing activity typically began in the student data post sets, Table 4.2 lists the first *referred to* and first *referring* post locations in all 10 post sets from this data set. As this table shows, the first *referred to* post locations in the student data ranged from PL2 - PL4. First *referring* post locations ranged from PL3 - PL9. One post set, S8, is identified as a *late-starting post set*, because of the comparatively late inception of referencing activity. S8 is highlighted in gray in Table 4.2. Its first *referred*

to post location was PL4, while that of the other nine post sets was PL2 or PL3. S8's first *referring* post location was PL9, while that of the other nine ranged from PL3 - PL6.

Table 4.2

First Referred to and First Referring Post Locations in the Student Data, by Post Set

Post set	First referred to post	First referring post
S1	3	6
S2	2	5
S3	3	5
S4	3	6
S5	2	3
S6	2	3
S7	3	6
S8	4	9
S9	2	3
S10	3	4

To learn more about when referencing activity typically began in the fan data, the linked post map for F5, presented in Table 4.3, provides a useful starting point.

Table 4.3

Linked Post Map, F5

PL	PA	PN	LPL	LPA	LPN	RCT
1	**1**	**1**				
2	2	1	1	1	1	Q
3	**3**	**1**				
4	4	1				
5	**5**	**1**	3	3	1	Q
6	**6**	**1**	1	1	1	Q
7	1	2	6	6	1	Q
8	7	1	6	6	1	Q
9	**8**	**1**				
10	**3**	**2**	5	5	1	Q
11	9	1				
12	5	2	10	3	2	Q
13	10	1	9	8	1	Q
14	**4**	**2**	9	8	1	Q
15	**10**	**2**	14	4	2	Q
16	**11**	**1**	9	8	1	Q
17	12	1				
18	4	3	15	10	2	Q
19	13	1	16	11	1	Q

The first *referred to* post in the map in Table 4.3 is PL1, and the first *referring* post is PL2. This suggests that referencing activity may have a tendency to start earlier in the fan data. To determine if this was the case, the first *referred to* and first *referring* posts in the five fan data post sets are listed in Table 4.4.

Table 4.4

First Referred to and First Referring Post Locations in the Fan Data, by Post Set

Post set	First referred to PL	First referring PL
F1	5	13
F2	2	5
F3	1	3
F4	1	3
F5	1	2

Table 4.4 indicates that the range of first *referred to* post locations in the fan data was PL1 - PL5, and the range of first *referring* post locations was PL2 - PL13. As was observed in the student data, one post set, F1, began referencing activity markedly later than the rest. This late-starting post set, F1, is highlighted in gray in Table 4.4. F1's first *referred to* post location was 5 while that of the other four was PL1 or PL2. The first *referring* post in F1 was PL13, compared to the other four, which ranged from PL2 - PL5.

Excluding the two late-starting post sets, the ranges for the start of referencing activity is similar in the two data sets. See Table 4.5 for a comparison.

Table 4.5

Excepting Late Starting Post Sets, Range of Post Locations When Referencing Activity First Began, by Data Set

Data set	Post sets	First referred to PL range	First referring PL range
Student	S1, S2, S3, S4, S5, S6, S7, S9, S10	2 - 3	3 - 6
Fan	F2, F3, F4, F5	1 - 2	2 - 5

As Table 4.5 demonstrates, ranges in the student data are one post location later than in the fan data. Because the first posts, or PL1s, in the student data were unique in that they were opening questions submitted as an assignment and not topics voluntarily submitted by the participants themselves, they may be considered exceptions with regard to referencing activity. The post location ranges, therefore, are virtually identical in both data sets for all post sets not classified as late-starting.

Reference Cues

Additional differences with respect to reference cues can also be observed through the comparison of linked post maps. Returning to the linked post map for S1 in Table 4.1, it can be observed that while 11 reference cues were identified, only 8 posts included reference cues. This is because two posts included more than one reference cue. PL16 included two reference cues, and PL6 included three. Posts with more than one reference cue account for the fact that in the entire student data set, 220 reference cues were found; yet only 179 posts included reference cues. The frequency of posts in the student data including reference cues, and the number of reference cues they include, are summarized in Table 4.6.

Table 4.6

Frequencies and Proportions of Reference Cue Use Characteristics in the Student Data

Total posts	**Total posts with reference cues**		**Posts with 1 reference cue**		**Posts with 2 reference cues**		**Posts with 3 reference cues**	
	Freq.	Prop.	Freq.	Prop.	Freq.	Prop.	Freq.	Prop.
311	179	.576	147	.821	23	.128	9	.050

Table 4.6 shows that overall, just over half of the posts in the student data included reference cues. Of those 179 posts, 147 included just one reference cue, 23 posts included two reference cues, and nine included three.

Alternatively, as the F5 linked post map in Table 4.3 shows, all 13 posts had just one reference cue each. This suggests that the posts in the fan data were less likely to have more than one reference cue. The frequencies and proportions of reference cue use in posts in the fan data are summarized in Table 4.7.

Table 4.7

Frequencies and Proportions of Reference Cue Use Characteristics in the Fan Data

Total posts	**Total posts with reference cues**		**Posts with 1 reference cue**		**Posts with 2 reference cues**	
	Freq.	Prop.	Freq.	Prop.	Freq.	Prop.
242	141	.583	135	.957	6	.043

As Table 4.7 demonstrates, only six of the 141 posts including reference cues in the fan data included more than one. None of these posts included more than two reference cues. Of note, it can also be observed that the proportion of posts in the fan data including reference cues, .583, was very similar to the proportion found in the student data, .576.

Types. There are three types of reference cues linking back to a prior post: naming cues, quoting cues and naming/quoting cues. The linked post map presented in Table 4.1 shows that S1 included three naming cues, three quoting cues, and five quoting/naming cues.

To determine if such a distribution of reference cue types was typical of the student data, the frequencies and proportions of reference cue types in the 10 post sets were compared. These are summarized in Table 4.8. This table shows that although all reference cue types were observed in each of the 10 post sets, overall quoting cues were used the most, followed by naming cues, and then quoting/naming cues in the student data.

Table 4.8

Reference Cue Type Frequencies and Proportions in the Student Data, by Post Set

Post set	**Total # reference cues used**	**Naming cues**		**Quoting cues**		**Quoting/Naming Cues**	
		Freq.	Prop.	Freq.	Prop.	Freq.	Prop.
S1	11	3	.273	3	.273	5	.455
S2	24	12	.500	8	.333	4	.167
S3	21	5	.238	7	.333	9	.429
S4	23	7	.304	12	.522	4	.174
S5	22	6	.273	9	.409	7	.318
S6	23	4	.174	12	.522	7	.304
S7	23	3	.130	14	.609	6	.261
S8	14	2	.143	6	.429	6	.429
S9	34	17	.500	14	.412	3	.088
S10	25	8	.320	10	.400	7	.280
Total	**220**	**67**	**.305**	**95**	**.432**	**58**	**.264**

In F5, on the other hand, only quoting cues were used, as demonstrated in its linked post map presented in Table 4.3. To determine if F5's demonstrated preference for quoting cues was typical of the fan data, the frequencies and proportions of reference cue types used in all five post sets were compared. These are summarized in Table 4.9.

Table 4.9

Reference Cue Type Frequencies and Proportions in the Fan Data, by Post Set

Post set	**Total # reference cues used**	**Naming cues**		**Quoting cues**		**Quoting/Naming Cues**	
		Freq.	Prop.	Freq.	Prop.	Freq.	Prop.
F1	11	0	.000	11	1.000	0	.000
F2	14	3	.214	11	.786	0	.000
F3	69	3	.043	65	.942	1	.014
F4	40	3	.075	37	.925	0	.000
F5	13	0	.000	13	1.000	0	.000
Total	**147**	**9**	**.061**	**137**	**.932**	**1**	**.007**

As Table 4.9 demonstrates, all five post sets in the fan data had a high proportion of quoting cues. Like F1, the *referring* posts in F5 used quoting cues 100% of the time. A

total of nine naming cues were discovered in the fan data, used three times each in F2, F3 and F4. Only one quoting/naming cue was used, and this was in F3.

Naming cues were further investigated across both data sets in order to test the coding scheme, which linked posts with naming cues to the most recent prior post submitted by the named post author. As discussed in Chapter 2, when a naming cue *refers back to* a post author who has only submitted one prior post, there is no ambiguity as to which post it should be linked to. Of the 76 total naming cues identified, 48 linked to a post author's first post, or a PN1. Ambiguity as to which post should be linked, therefore, was only a factor in the remaining 28 naming cues linking to a post author's second post or later, or PN > 1.

When compared to the other reference cue types, the proportion of naming cues linking to a PN > 1 was no greater than that of quoting cues or quoting/naming cues. These findings are presented in Table 4.10.

Table 4.10

Proportion of Reference Cues Pointing to PN1s and PNs > 1, by Reference Cue Type

	Total	**PN1**		**PN > 1**	
		Freq.	Prop.	Freq.	Prop.
Naming cues	76	48	.632	28	.368
Quoting cues	232	120	.517	112	.483
Naming/Quoting cues	59	34	.576	25	.424

As can be observed in Table 4.10, the proportion of naming cues linking to a PN1 was actually somewhat higher than that of quoting cues and quoting/naming cues, and the proportion linking to a PN > 1 lower.

Post States

An additional observation that can be made from the linked post maps in Tables 4.1 and 4.3 is that not all posts are involved in referencing activity, meaning that they did

not refer back to any prior posts, and were not *referred to* in any following posts. These posts that are uninvolved in referencing activity are termed *dormant posts*. Dormant posts are neither bold nor outlined in the linked post maps, and they are not included in any post pairs.

Posts that are included in at least one post pair have either *referred back to* a prior post, were *referred to* in a following post, or both. These, therefore, are coded as *active posts.* Once the active and dormant *post states* for each post have been established, they can be analyzed in sequence. For example, in S1, the sequence of active and dormant post states is as follows, where D represents dormant posts and A represents active posts:

D D A A A A A A D A A A D A A A A D D D A

From this sequence, it can be seen that 14 of the posts in S1 were active and seven were dormant. To see if these proportions were typical in the student data, the active and dormant posts in all ten post sets are summarized in Table 4.11.

Table 4.11

Active and Dormant Post Frequencies and Proportions in the Student Data, by Post Set

Post set	**Total # posts**	**Dormant posts**		**Active posts**	
		Frequency	Proportion	Frequency	Proportion
S1	21	7	.333	14	.667
S2	33	5	.152	28	.848
S3	33	8	.242	25	.758
S4	30	6	.200	24	.800
S5	30	6	.200	24	.800
S6	33	7	.212	26	.788
S7	36	8	.222	28	.778
S8	30	10	.333	20	.667
S9	35	7	.200	28	.800
S10	30	4	.133	26	.867
Total	**311**	**68**	**.219**	**243**	**.781**

As demonstrated in Table 4.11, all post sets in the student data had mostly active posts. S1 and S8 had the lowest proportions of active posts. In the case of S8, this may be

related to its late start to referencing activity, previously discussed. S1 was one of the first post sets to take place, and students were still learning how to use the discussion board, which may account for its lower proportion of active posts.

To investigate the active and dormant post proportions in the fan data, these five post sets were compared as well. As an example, the sequence of active and dormant posts in F5 is as follows:

A A A D A A A A A A D A A A A A D A A

The frequencies and proportions of active and dormant posts for each post set in the fan data are summarized in Table 4.12.

Table 4.12

Active and Dormant Post Frequencies and Proportions in the Fan Data, by Post Set

Post Set	**Total # posts**	**Dormant posts**		**Active posts**	
		Frequency	Proportion	Frequency	Proportion
F1	35	20	.571	15	.429
F2	27	9	.333	18	.667
F3	95	14	.147	81	.853
F4	66	15	.227	51	.773
F5	19	3	.158	16	.842
Total	**242**	**61**	**.252**	**181**	**.748**

Table 4.12 shows that a majority of posts in the fan data were active. However, one post set, F1, had more dormant posts than active posts.

As the only post set in all of the data to be observed with a majority of dormant posts, F1 was further investigated to determine why this may have been the case. In addition to being a previously identified late-starting post set, it was found that F1 had a remarkably high proportion of post authors contributing only one post. Comparing the proportion of post authors contributing only one post in all five fan post sets, as well as the 10 student post sets, a general trend was observed, where post sets with a higher

proportion of one-post contributing authors tended to have a lower proportion of active posts. This trend is represented in the line graphs in Figure 4.1, which plots the post sets in each data set according to their proportions of active posts and one-post contributing authors.

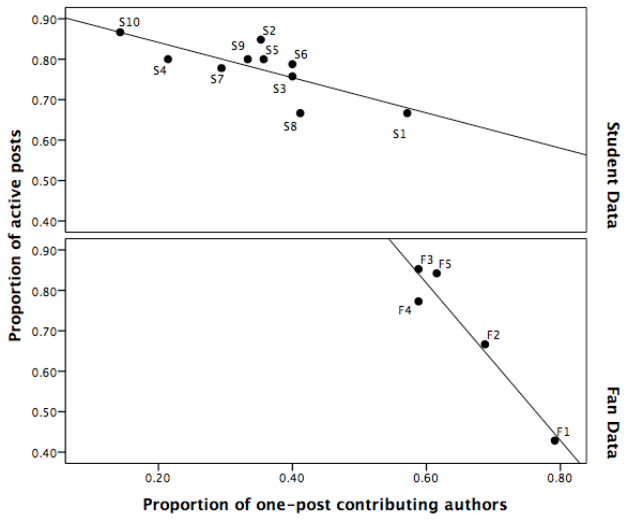

Figure 4.1. Negative relationship between post sets' proportions of active posts and one-post contributing authors, by data set.

As seen in Figure 4.1, the negative relationship between active post and one-post contributing author proportions was observed to varying degrees in both data sets, but was stronger in the fan data. F1 in the lower right corner of the graph can be seen as the most extreme case. S1 and S8 of the student data were also found to have higher proportions of one-post contributing authors, which may additionally account for the lower proportions of active posts.

Post author sequence. Now that post states have been determined, they may be layered on as an additional dimension in the post author sequence maps, presented in Chapter 3. Figure 4.2 demonstrates this layering in a post author sequence map for S1. Post author sequence maps containing the added dimension of active posts are presented

for all 10 post sets from the student data, and the five post sets from the fan data, in Appendix C.

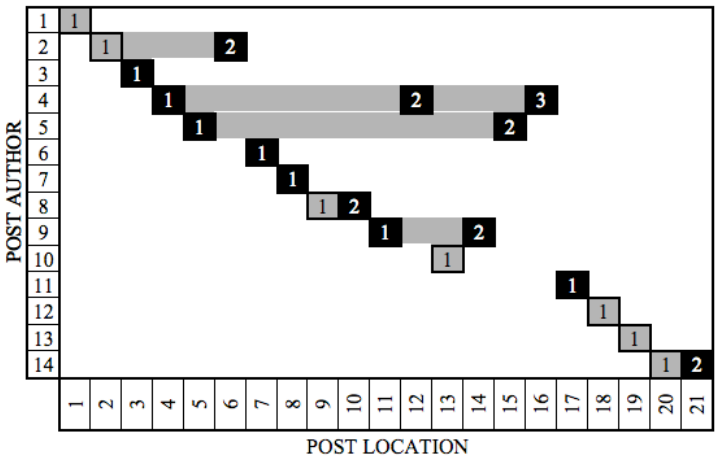

Figure 4.2. Post author sequence map with active posts, S1.

Post author sequence maps like the one presented in Figure 4.2 represent post locations along the x-axis, and post author values along the y-axis. In the revised version of the map, active posts are represented by black boxes, and dormant posts as outlined gray boxes. Regardless of post state, all post boxes show the individual post author's post number for that post. Gray bars between boxes represent the post locations occurring in sequence between an individual post authors' posts.

In the post author sequence map presented in Figure 4.2, it can be observed that every second and third post submitted by a post author is active in S1. The post author sequence maps presented in Appendix C show that PNs > 1 were typically found to be active in the student data. Supporting this observation, the post states of PN1s were compared to those of PNs > 1. Table 4.13 presents the frequency and proportion of active posts in these two categories. As this table shows, approximately 60% of post authors' PN1s were active in the student data. Posts with a PN > 1, however, were active 95% of the time.

Table 4.13

Frequencies and Proportions of Dormant and Active Posts in PN1s versus PNs > 1 in the Student Data

Post number	**Total posts**	**Dormant posts**		**Active posts**	
		Frequency	Proportion	Frequency	Proportion
PN1	152	60	.395	92	.605
PN > 1	159	8	.050	151	.950

To learn more about the post states of PN1s and PNs >1 in the fan data, the post author sequence map for F5, presented in Figure 4.3, is a useful starting point.

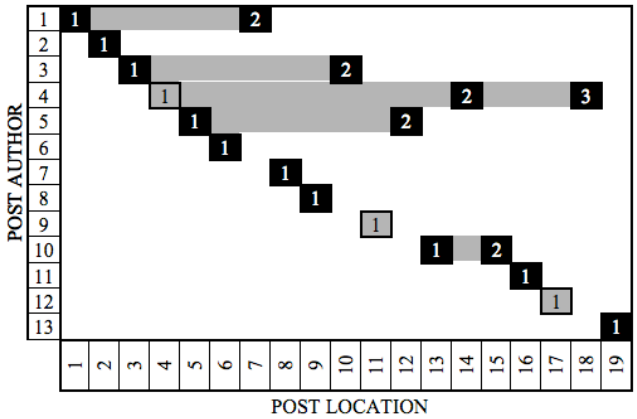

Figure 4.3. Post author sequence map with active posts, F5.

As demonstrated in Figure 4.3, all posts with a PN > 1 were active in F5 as well. To learn more about this occurrence in the fan data, Table 4.14 shows the frequencies and proportions of active posts in PN1s and PNs > 1 in the fan data.

Table 4.14

Frequencies and Proportions of Dormant and Active Posts of PN1s versus PNs > 1 in the Fan Data

Post number	**Total posts**	**Dormant posts**		**Active posts**	
		Frequency	Proportion	Frequency	Proportion
PN1	105	40	.381	65	.619
PN >1	137	21	.153	116	.847

Table 4.14 shows that, like the student data, a greater proportion of PN1s, and a much greater proportion of PNs > 1, were active in the fan data. Comparing the proportions of

dormant and active PN1s in Tables 4.14 and 4.13, it can be observed that they were similarly distributed in both data sets. The PNs > 1 in the student data, however, were active 95.0% of the time, while in the fan data they were active just 84.7% of the time.

Post time sequence. Post states may similarly be layered onto the post time sequence maps introduced in Chapter 3. In these revised maps, active posts are again represented as black boxes, and dormant posts as outlined gray boxes. The PA value for the post author who submitted the post is displayed within these boxes. The y-axis displays the post location and the x-axis displays the ten standardized segments of the post set's time frame, each of which is broken down into quadrants. The gray bars persist through the time segments, up to and including that in which the same post author next submits a post. Figure 4.4 demonstrates such a map for S1. These maps for all 10 post sets from the student data, as well as the five post sets from the fan data, are presented in Appendix D.

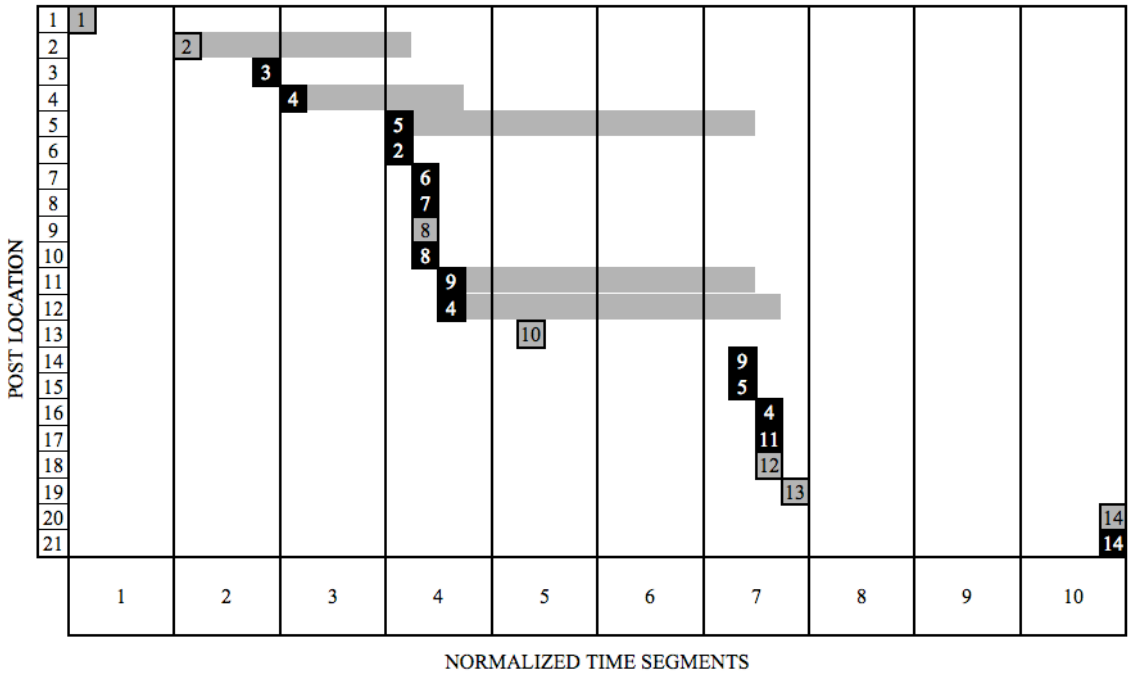

Figure 4.4. Post time sequence map with active posts, S1.

As the map in Figure 4.4 shows, all active posts tended to be submitted within posting spurts, or clusters of posting activity within shared or consecutive segments of time, previously discussed in Chapter 3. Three of the seven dormant posts, however, were submitted outside of posting spurts. Thus, it was hypothesized that the time lag surrounding active post submission would be shorter than that of dormant posts in the student data. The mean time lags of active and dormant posts in the student data are presented in Table 4.15.

Table 4.15

Mean Time Lag of Active and Dormant Posts in the Student Data

	Total posts	**Time lag**	
		M	SD
Active posts	243	290.51	463.55
Dormant posts	68	450.88	999.58

As this table shows, the time lag of active posts, on average, was shorter than that of dormant posts.

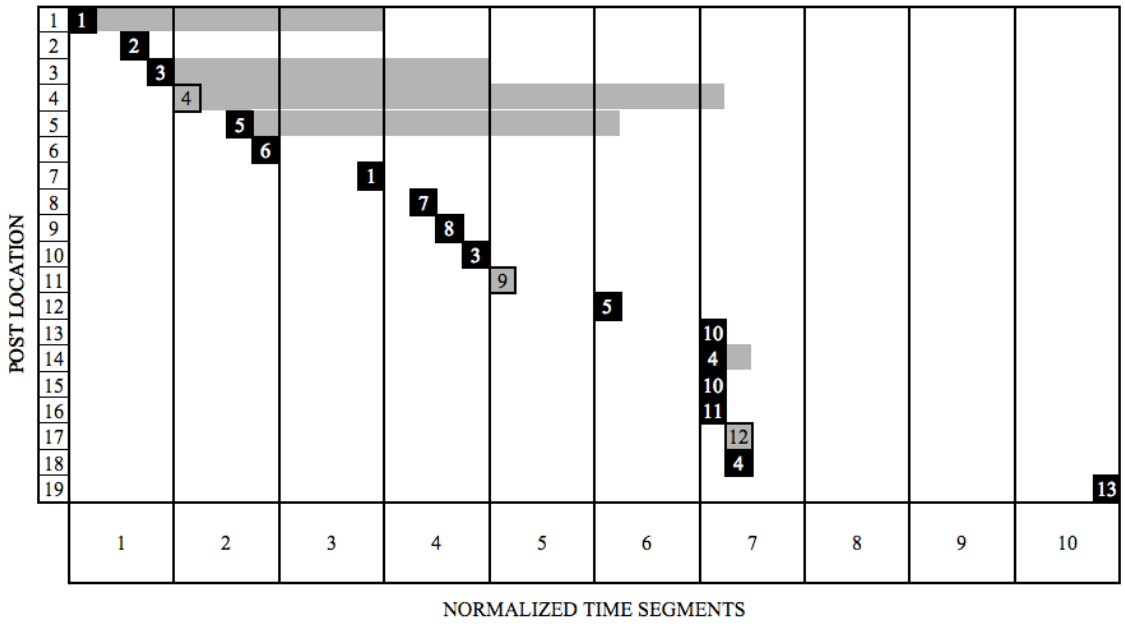

Figure 4.5. Post time sequence map with active posts, F5.

To investigate this phenomenon in the fan data, the post time sequence map for F5, presented in Figure 4.5, can be observed. This figure shows that, in F5, all but one active post was submitted within a posting spurt, and all three dormant posts were submitted within posting spurts. This suggests that there may be no difference between the time lags of active and dormant posts in the fan data. Their average time lags are presented in Table 4.16.

Table 4.16

Mean Time Lag of Active and Dormant Posts in the Fan Data

	Total posts	**Time lag**	
		M	SD
Active posts	181	40.46	256.26
Dormant posts	61	98.18	556.77

Table 4.16 shows that the mean time lags of active posts were shorter than dormant posts in the fan data as well. F5, then, appears to be somewhat of an anomaly in the fan data, likely resulting from its comparatively short 92-minute time frame.

Active post states. As previously discussed in Chapter 2, active post states can be further specified into three additional categories. These *active post states* are based on the presence or absence of post reference cues, and/or of references to them in following posts. These can be determined by looking at post position(s) in post pairs. Posts that are found in the first position only are deemed to be in an *originating state*, because they were *referred to* in one or more following posts, but did not *refer back to* prior posts. Posts that are found in both the first and second position of post pairs are deemed to be in an *embedded state*, because they were both *referred to* in one or more following posts, and *referred back to* one or more prior posts. Finally, posts located only in the second position of post pairs are deemed to be in a *terminating state*, because they *refer back to*

one or more prior posts, but are not *referred to* in any following posts. Note that originating and terminating posts may be present in one or more post pairs, but embedded posts must be present in at least two.

Once identified, all post states can also be observed in sequence. For example, by observing posts' positions in post pairs in the S1 linked post map presented in Table 4.1, the post states in sequence can be determined as follows, where D represents dormant posts, O represents originating posts, E represents embedded posts, and T represents terminating posts:

D D O O O T O E D T O E D E E T O D D D T

In S1, it appears that while all three types appear throughout the post set, more originating posts occur at the beginning of the sequence, while embedded and terminating posts are more prevalent in the middle and end. This was observed in the sequence of post states for F5 as well:

O T O D E E T T O E D T T E E E D T T

It seems reasonable to expect originating posts to be found more often at the beginning of the post set, since posts submitted earlier have fewer prior posts to *refer back to*, but there is more opportunity for following posts to *refer back to* them. The later a post is submitted in the sequence, the more posts it has available to *refer back to*, thereby increasing the probability of it being in an embedded or terminating state. *Referring* posts that are submitted towards the end of a sequence have fewer following posts to *refer back to* them, increasing their likelihood of existing in a terminating state.

To learn more about the placement of the four post states, all 15 post sets are visualized in Figure 4.6.

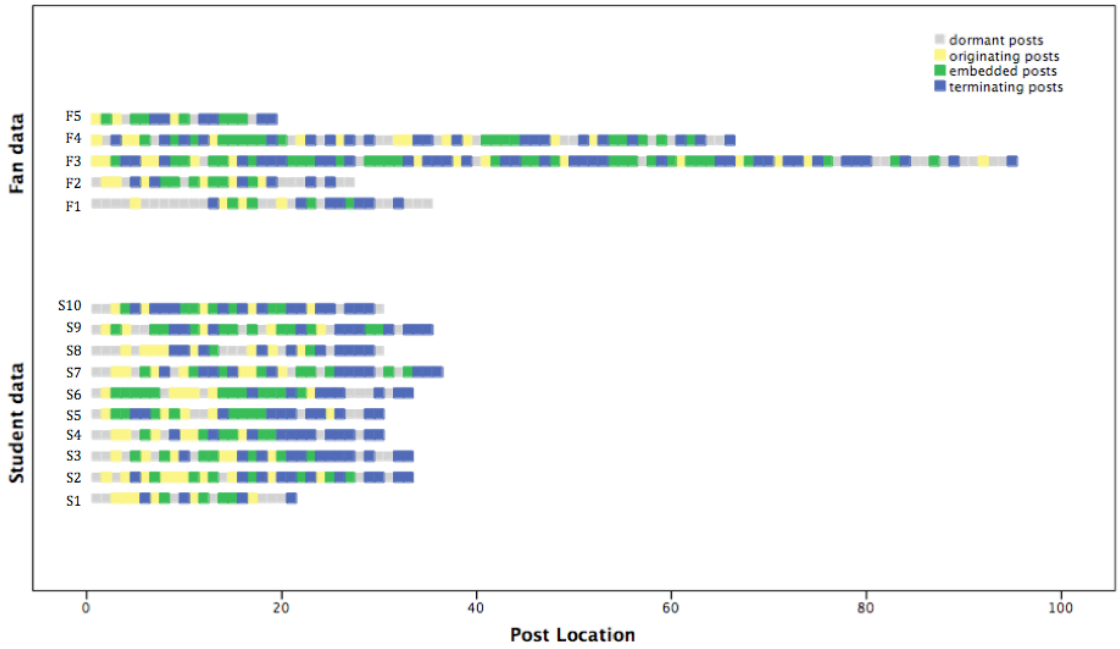

Figure 4.6. Post state sequences, all post sets.

Figure 4.6 shows each post set as a line of posts, in sequence, displayed as colored boxes representing their post state. Dormant posts are colored in gray, originating posts in yellow, embedded posts in green, and terminating posts in blue. From this figure, it can be observed that in most post sets, originating posts tend to occur more in the first half of post sets, and terminating posts tend to occur more in the second half. Embedded and dormant posts are shown to occur throughout post sets.

Supporting this observation, the mean post location for the posts in each post state category of both data sets is presented in Figure 4.7. As Figure 4.7 shows, the mean post locations of post states are comparatively similar in the two data sets. In both the student data and the fan data, the mean post location of originating posts is earlier, the mean post location of terminating posts later, and the mean post location of dormant and embedded posts positioned in between.

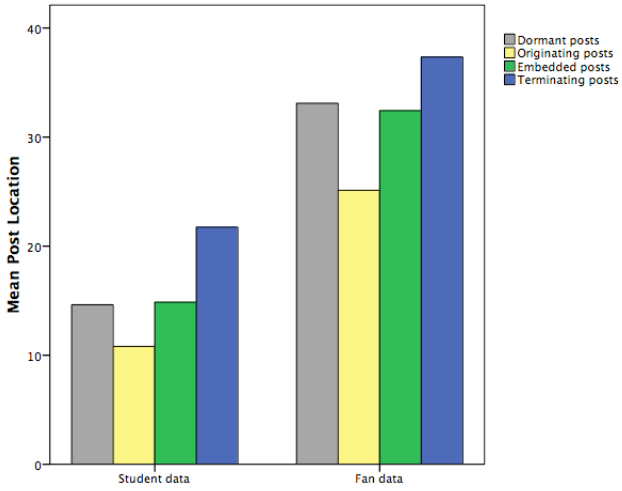

Figure 4.7. Mean post location of posts in dormant, originating, embedded, and terminating states, by data set.

Participants

To determine if there were individual differences between post authors with respect to referencing activity, the post pair involvement of all 18 post authors in the student data was investigated. All but one of these post authors was found to be involved in at least four post pairs. The one exception was myself, ID1, who submitted a total of 11 posts. Ten of these posts were the "opening question" posts for the weekly online class discussions. The 11th was a follow-up post to test a technical glitch occurring on the discussion board. The first post in a post set cannot *refer back to* prior posts, since there are none. The other post authors in the student data may not have felt the need to *refer back to* my posts, perhaps because they represented an assignment they were assumed to be fulfilling with their responses. Therefore, ID1 as an explainable exception was removed from the analysis.

Figure 4.8 displays a bar graph of the remaining 17 post authors in the student data, in ascending order based on the total number of post pairs involved in. The bars represent the post author's total number of post pairs, with the green portion representing

the proportion in which they were the *referring* post author, and the blue in which they were the *referred to* post author.

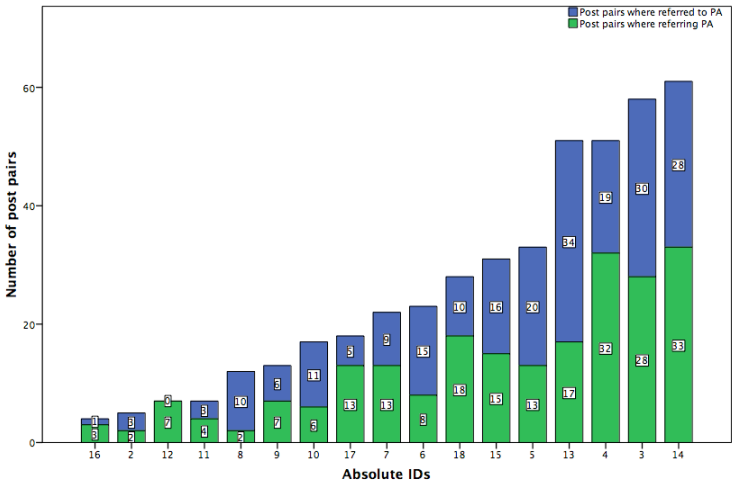

Figure 4.8. Post pair involvement of post authors in the student data.

As Figure 4.8 shows, only one post author from the student data was never *referred to* in any following posts. This was ID12, who was identified in the previous chapter as the post author most often adopting a transient approach. While never *referred to* in any following posts, ID12 *referred back to* a total of seven prior posts. This unique situation is likely due to ID12's tendency to post later in the sequence of a post set, a tendency often observed in transient post authors. Posts that occur later in a post set's sequence have less opportunities for following posts to *refer back to* them. This likely accounts for the fact that ID12's consistently late posts were never referred back to. The remaining 16 student post authors were all *referred to* in a following post at least one time, and *referred back to* a prior post at least twice.

From the graph presented in Figure 4.8, it can be observed that some post authors display inclinations towards either *referring*, or being *referred to*. ID12, along with IDs 4, 16, 17, and 18, all *referred back to* prior posts at least 1.5 times more often than they were *referred to* in following posts. These are *greater-referring post authors*. IDs 5, 6, 8,

10, and 13 were all *referred to* in following posts at least 1.5 times more than they *referred back to* prior posts. These are called *greater-referred-to post authors*.

Based on the analysis of ID12's unique reference proportions, it was hypothesized that greater-referring post authors may have a tendency to join the post set later on, providing less opportunity for following posts to *refer to* them, and greater-referred-to post authors a tendency to post earlier in the sequence, allowing for more such opportunities. By splitting the set of post authors for each post set into halves, these post authors were classified as either early or later joining post authors for each post set they participated in. The results are presented in Table 4.17. As this table demonstrates, all five of the greater-referred-to post authors were found to be early joining post authors in three to 10 post sets. Greater-referred-to post authors, therefore, may tend to be *referred to* in following posts more often as a result of their common early entry into post sets. Four of the five greater-referring post authors were found to be later joining in four to seven of the post sets they participated in. This suggests that greater-referring post authors have a later-joining tendency in the student data.

Table 4.17

Early and Later Joining Tendencies of Greater-Referred-To and Greater-Referring Post Authors in the Student Data

Greater-referred-to post authors				**Greater-referring post authors**			
Abs. ID	# post sets	# early joining	# later joining	Abs. ID	# post sets	# early joining	# later joining
5	10	10	0	4	10	10	0
6	8	5	3	12	4	0	4
8	10	3	7	16	5	0	5
10	8	4	4	17	7	0	7
13	9	9	0	18	8	1	7

ID4 as the early-joining exception of greater-referring post authors suggests that there may be additional factors that influence the greater-referring status of post authors.

It was hypothesized that the greater-referring characterization may additionally be related to a tendency to include reference cues in posts. The proportion of greater-referring post authors' posts with reference cues was compared to that of other post authors. The results of this comparison are presented in Table 4.18.

Table 4.18

Frequency and Proportion of Greater-Referring and Other Post Authors' Posts Including Reference Cues in the Student Data

	Total posts	**Posts with reference cues**		**Posts without reference cues**	
		Freq.	Prop.	Freq.	Prop.
Greater-referring post authors	98	61	.622	37	.378
Other post authors	213	118	.554	95	.446

As Table 4.18 demonstrates, greater-referring post authors were found to have higher proportions of posts including reference cues than other post authors.

For purposes of comparison, the 10 post authors with ≥ five posts in the fan data were looked at as well, as the lowest number of posts observed for post authors in the student data was five. These 10 post authors were found to be involved in at least eight post pairs. Figure 4.9 displays a bar graph of these 10 post authors, in ascending order based on number of post pairs involved in. Again, the bars represent the post author's total number of post pairs, with green representing the proportion in which they were the *referring* post author, and blue representing the proportion in which they were the *referred to* post author. As the graph in Figure 4.9 shows, all 10 of these post authors were *referred to* in a following post, and *referred back to* a prior post, at least three times. Only two individuals, IDs 44 and 50, may be characterized as greater-referred-to post authors. IDs 57, 59, 64 and 91 may be characterized as greater-referring post authors.

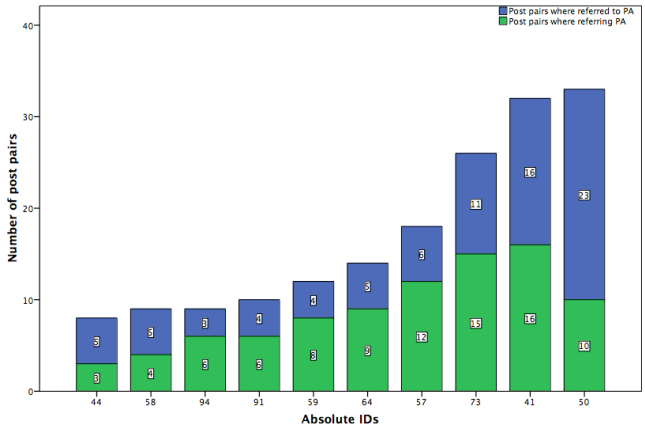

Figure 4.9. Post pairs for post authors contributing ≥ five posts in the fan data.

To see if early and later-joining tendencies influenced these characterizations in the fan data, the number of post sets in which they joined early and late were looked at. These are summarized in Table 4.19.

Table 4.19

Early and Later Joining Tendencies of Greater-Referred-To and Greater-Referring Post Authors in the Fan Data

Greater-referred-to post authors				**Greater-referring post authors**			
Abs. ID	# post sets	# early joining	# later joining	Abs. ID	# post sets	# early joining	# later joining
44	1	1	0	57	1	1	0
50	3	2	1	59	1	1	0
				64	1	1	0
				91	1	1	0
				94	1	0	1

Table 4.19 shows that both greater-referred to post authors were early joining in a majority of the post sets contributed to. Only one of the greater-referring post authors was found to be later-joining in the post set contributed to. Therefore, while it may be said that greater-referred to post authors in the fan data had an early-joining tendency, greater-referring post authors did not have a later-joining tendency. To see if the greater-referring post authors in the fan data may reflect a tendency to include reference cues in a

higher proportion of their posts, their reference cue inclusion in posts was compared to that of other post authors in the fan data. These results are summarized in Table 4.20.

Table 4.20

Frequency and Proportion of Greater-Referring and Other Post Authors' Posts Including Reference Cues in the Fan Data

	Total posts	**Posts with reference cues**		**Posts without reference cues**	
		Freq.	Prop.	Freq.	Prop.
Greater-referring post authors	49	43	.878	6	.122
Other post authors	193	99	.513	94	.487

As Table 4.20 demonstrates, the proportion of posts including reference cues is much higher for greater-referring post authors in the fan data. Therefore, while in the student data the greater-referring characterization is related to a late-joining tendency combined with a tendency to included reference cues, in the fan data it appears to be related only to a strong tendency to include reference cues.

Linking and Mapping Chains

The next step in linking the posts of a post set is to map chains by linking post pairs. Post pairs are linked when they share a common post. As stated in Chapter 2, chains are displayed best in a visual map that shows the post pairs' linking connections. The ordering of post pairs displayed in chains should ensure to the greatest extent possible that shared posts be placed one after another, giving proximal priority to shared embedded posts over shared originating and shared terminating posts. The linked chain map for S1 is presented in Figure 4.10. Linked chain maps for all 10 post sets in the student data, as well as the five post sets in the fan data, are presented in Appendix E.

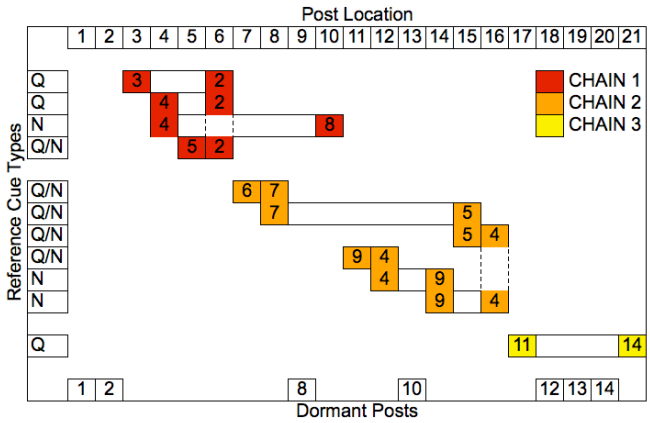

Figure 4.10. Linked chain map, S1.

In linked chain maps like the one presented for S1 in Figure 4.10, the post locations are presented along the x-axis, with the posts in each post pair represented by colored boxes. The white boxes at the bottom of the graph represent dormant posts. The numbers inside each box display the author of that post. The first colored box in each post pair represents the *referred to* post, and the second represents the post that *referred back to* it. Posts that are shared by and therefore link post pairs, are presented either as a single open box or, when separated by other post pairs in the display, two or more open boxes joined by a dotted border. The white bars represent the number of posts that occur between two posts in a post pair, or its *sequential lag*. Each individual chain is assigned a number and color, which is displayed in the key in the upper right corner of the map. The y-axis shows the reference cues used to create the post pairs in the same row. This visualization allows for comparison of the post pairs and chains within a given post set. In addition, this can be used to provide an overall portrayal of the linked structure of a post set, which can then be used for comparison to others.

Activity

Linked chain maps provide a view of the combinations of four possible active post states in each post pair. The four possible *post pair states* are originating-embedded

(O-E), originating-terminating (O-T), embedded-embedded (E-E), and embedded-terminating (E-T). As can be observed in the linked chain map presented in Figure 4.10, the post pair states in the chains of S1, where O represents originating posts, E represents embedded posts, and T represents terminating posts, are as follows:

CH1: O-T, O-T, O-T, O-T

CH2: O-E, E-E, E-T, O-E, E-E, E-T

CH3: O-T

Both CH1 and CH3 in S1 are comprised of only O-T post pairs. CH2, on the other hand, has no O-T post pairs, but two O-E post pairs, two E-E post pairs, and two E-T post pairs. For this reason, CH2 is the only chain in Figure 4.10 with the step-like formations created by embedded posts.

To learn more about the frequencies of post pair states in the student data, all 220 post pairs were compared, by post set and overall. These are presented in Table 4.21.

Table 4.21

Frequencies and Proportions of Post Pair States in the Student Data, by Post Set and Overall

Post Set	**O-T post pairs**		**O-E post pairs**		**E-E post pairs**		**E-T post pairs**		**Total post pairs**
	Freq.	Prop.	Freq.	Prop.	Freq.	Prop.	Freq.	Prop.	
S1	5	.455	2	.182	2	.182	2	.182	11
S2	8	.333	6	.250	2	.083	8	.333	24
S3	8	.381	4	.190	3	.143	6	.286	21
S4	6	.261	6	.261	2	.087	9	.391	23
S5	4	.182	4	.182	5	.227	9	.409	22
S6	6	.261	4	.174	10	.435	3	.130	23
S7	5	.217	5	.217	4	.174	9	.391	23
S8	10	.714	2	.143	0	.000	2	.143	14
S9	5	.147	5	.147	14	.412	10	.294	34
S10	7	.280	4	.160	4	.160	10	.400	25
Totals	**64**	**.291**	**42**	**.191**	**46**	**.209**	**68**	**.309**	**220**

As this Table 4.21 shows, E-T post pairs were most commonly found in the student data, and O-E post pairs were least commonly found. There was no "typical" proportion of post pair states in the post sets. The late-starting post set S8, however, had an unexpectedly high proportion of O-T post pairs, and was the only post set to have no E-E post pairs.

To begin investigation into the post pair states in the fan data, the linked chain map for F5 is presented in Figure 4.11.

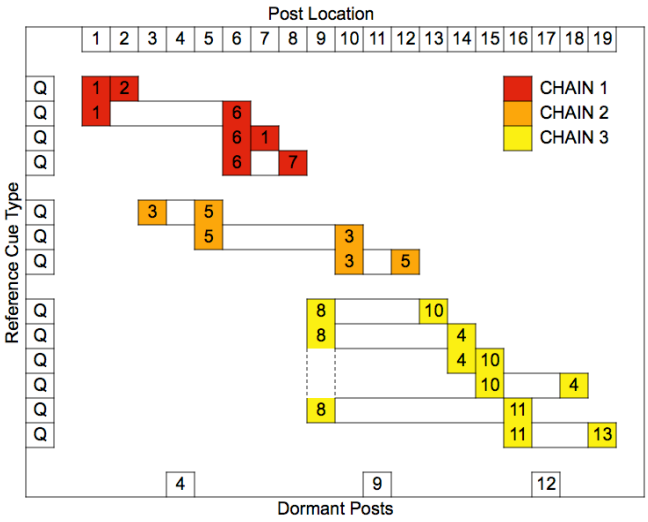

Figure 4.11. Linked chain map, F5.

From the linked chain map presented in Figure 4.11, it can be determined that the post pair states in the chains of F5 are as follows:

CH1: O-T, O-E, E-T, E-T

CH2: O-E, E-E, E-T

CH3: O-T, O-E, E-E, E-T, O-E, E-T

In F5, CH1 and CH3 have at least one of each type of post pair state. CH2 has no O-T post pairs, one O-E post pair, one E-E post pair, and one E-T post pair.

To learn more about the frequencies of post pair states in the fan data, the 147 post pairs in this data set were investigated. The frequency and proportion of post pairs in each post pair state, by post set and overall, is presented in Table 4.22.

Table 4.22

Frequencies and Proportions of Post Pair States in the Fan Data, by Post Set and Overall

Post Set	**O-T post pairs**		**O-E post pairs**		**E-E post pairs**		**E-T post pairs**		**Total post pairs**
	Freq.	Prop.	Freq.	Prop.	Freq.	Prop.	Freq.	Prop.	
F1	4	.364	3	.273	1	.091	3	.273	11
F2	5	.357	3	.214	4	.286	2	.143	14
F3	11	.159	11	.159	20	.290	27	.391	69
F4	12	.300	7	.175	12	.300	9	.225	40
F5	1	.077	4	.308	3	.231	5	.385	13
Totals	**33**	**.224**	**28**	**.190**	**40**	**.272**	**46**	**.313**	**147**

As Table 4.22 shows, O-E post pairs were the least common in the fan data, and E-T post pairs were the most common. This is similar to what was observed in Table 4.21 for the student data. The fan data, however, had a somewhat lower proportion of O-T pairs than the student data, and a somewhat higher proportion of E-E pairs.

Like S8 in the student data, F1, the late-starting post set in the fan data, also had the lowest proportion of E-E post pairs in its data set. Comparing the structure of the linked chain maps of both S8 and F1 to the others in Appendix E, their lack of deep, step-like formations, created by runs of linked E-E post pairs, gives these maps a comparatively stagnant appearance.

Participants

The linked chain maps show that, in both S1 and F5, each chain represents a distinct group of post authors in the post set. As shown in Figure 4.10, in S1, PAs 2, 3, 4, 5, and 8 participate in the first chain (CH1), PA4 and PA5 again participate in CH2 along

with PAs 6, 7 and 9, while CH3 consists of a single post pair between PA11 and PA14. Although PA4 and 5 are both involved in CH1 and CH2, it is notable that they were only directly linked in a post pair in CH2. Thus, they did not directly refer to one another in CH1. As shown in Figure 4.11, F5 demonstrates no crossover post authors between chains at all, where PAs 1, 2, 6 and 7 participate in CH1 only, PA3 and PA5 in CH2 only, and PAs 8, 10, 4, 11 and 13 in CH3 only. By looking at the chains of all post sets in the linked chain maps listed in Appendix E, it can be concluded that chain post author sub-groupings were a phenomenon common to both data sets. The number of post authors in these sub-groupings ranged from 1-11. The characteristics of all chains, and the post author sub-groupings they create, are summarized in Appendix F.

As can be observed in Figure 4.10, PA4 and PA9 engaged in a back-and-forth referencing situation in CH2 of S1. Figure 4.11 shows a similar situation between PA3 and PA5 in CH2 of F5. Such back and forth activity may be reflective of an ongoing dialogue between two post authors. Thus, it was hypothesized that, when two post authors were involved in more than one post pair in the same post set, they would be likely to be grouped into the same chain. A total of 47 instances where sets of post pairs involving the same two post authors in a post set were discovered. Nine of the ten post sets from the student data, and four of the five from the fan data, had at least one such set of post pairs. The two late-starting post sets did not share this characteristic, further underscoring their lower levels of referencing activity. The number of post pairs between the same two authors ranged from 2-13. The characteristics of these 47 instances are presented in Table 4.23.

Table 4.23

Instances of Two Post Authors Involved in Two or More Post Pairs in the Same Post Set

Post set	Post author Abs. IDs	# of post pairs between post authors	# of chains linked into	Value(s) of chain(s) linked into
S1	4 & 7	3	1	2
S2	3 & 13	2	2	2, 5
S3	13 & 14	2	2	1
	3 & 14	2	1	1, 4
	3 & 15	2	1	3
S4	3 & 4	2	1	1
	3 & 6	2	1	1
	13 & 14	3	1	1
	14 & 15	2	1	1
S5	3 & 4	2	1	1
	3 & 14	2	1	1
	3 & 15	2	1	1
	4 & 15	3	1	1
	13 & 14	2	1	1
S6	3 & 14	8	1	1
	13 & 15	3	2	1, 3
	14 & 15	2	1	1
S7	3 & 13	2	1	1
	6 & 14	2	1	2
	5 & 14	2	1	3
	14 & 17	2	1	5
S9	3 & 5	2	1	1
	3 & 6	2	1	1
	3 & 14	2	1	1
	4 & 13	2	1	1
	7 & 15	3	1	1
	10 & 14	2	1	1
	4 & 6	2	2	1, 3
	4 & 5	2	1	4
S10	4 & 6	2	1	1
	4 & 14	3	1	1
	13 & 14	2	1	1
	17 & 18	2	1	1
F2	41 & 44	4	1	2
	44 & 46	3	1	2
F3	41 & 64	4	1	1
	59 & 64	5	1	1
	41 & 59	5	2	1, 2
	57 & 58	4	3	1, 6, 7
	70 & 71	3	1	2
	58 & 71	2	2	2, 8
F4	50 & 73	13	5	3, 4, 6, 10, 11
	50 & 91	9	4	4, 5, 6, 10
	73 & 94	2	2	6, 10
	50 & 94	7	4	6, 10, 11, 12
F5	100 & 101	3	1	2
	41 & 65	2	1	3

By looking at the fourth column in this Table 4.23, it can be observed that, in 76.6%, or 36 of the 47 instances, the set of post pairs between the same two post authors were contained in the same chain. In these 36 instances, the number of post pairs between the same two authors in one chain ranged from 2-8.

The remaining 11 sets of post pairs between the same two post authors were further investigated by studying the content of the posts, to determine the reasons they were not linked into the same chain. In three cases, those occurring in S2, S3, and between ID57 and ID58 in F3, it was clear that different topics were being addressed in each chain. In S3, for example, the two post authors were joking about an unrelated issue in CH1, and engaging with class topics in CH4. Therefore, the separate chains represent separate conversations between the two post authors.

In the eight remaining cases, it was observed that one of the two post authors submitted an originating post, thereby starting one of the chains they were involved in. In three of these cases, those occurring in S6, S9, and between ID58 and ID71 in F3, the text of the originating posts indicated a generalized response to no one in particular, and introduced new questions and/or concepts to be discussed. For example, the originating post in which this occurred in S6 began with the words "To respond in general…" The chains in these three instances also occurred in separate time frames, and there was no overlap between the posts' locations in the post set's sequence. These three cases, then, also appear to represent separate segments of conversation between the two post authors. Therefore, in 89.3%, or 42 of the 47 cases observed where more than one post pair involved the same two post authors, all post pairs appeared to be placed into the appropriate chains.

The final five cases were all derived from the fan data, and four of these were from F4. These are highlighted in gray in Table 4.23. Investigating these instances further, the originating posts starting new chains were found to begin with words such as, "but," and "so" and addressed "you" and "your," clearly indicating that the post author was addressing a person or statement. One post, for example, stated simply, "Your explanation kind of made no sense to me just now." Reading the posts prior to these originating posts, it became apparent that these posts were intended as responses to recently submitted posts, yet no reference cue to those posts were included. Therefore, such *non-referring* posts should have been shared between post pairs of chains that were in fact left unlinked. ID50 did this a total of seven times, at PLs 5, 13, 22, 26, 39, 48 and 52 in F4. Two other post authors did this once each, ID41 at PL6 in F3, and ID91 at PL33 in F4. Had these post authors used reference cues to *refer back to* the posts to which they were responding, CH1 and CH2 in F3 would have combined into one chain, as would CHs 3, 4, 5, 6, 8, 9, 10, 11, and 12 in F4. Although this analysis has demonstrated that the method used is generally very good at capturing distinct post author sub-groupings in chains, these three post authors may represent a *non-referring* posting approach that proves problematic for the method.

To learn more about this approach, by whom it is adopted, and in what situations, a more detailed investigation into the nature of participation in these instances was conducted. Relevant data for these three post authors for the post sets in which they were non-referring, are summarized in Table 4.24. As the fourth and fifth columns in this table show, all three non-referring post authors contributed several posts across the majority of F3 and F4.

Table 4.24

Non-Referring Post Author Characteristics

Abs. ID	Post set where non-referring	PA value in post set	Total # of posts in post set	Range of post locations in post set	Non-referring post numbers (PNs)
41	F3	3	12	3-93	2
50	F4	5	20	5-59	1, 3, 7, 9, 13, 16, 18
91	F4	9	10	14-49	8

ID50 and ID41 were previously identified as enthusiastic post authors in Chapter 3. This suggests that higher contributing authors may be more likely to submit non-referring posts. The third non-referring post author, ID91, submitted a non-referring post that was clearly intended to respond to one of ID50's posts. This led to the speculation that perhaps ID91 adopted this approach after observing ID50's non-referring behavior in four prior posts.

Both ID50 and ID41 adopted the non-referring approach in the post set they contributed to the most, and neither ID41 nor ID50 adopted this approach in the other post sets they were observed contributing to. ID91 was observed contributing to F4 only. Additionally, all three non-referring post authors submitted other posts within the same post set that did include reference cues. This suggests that the non-referring posting approach may only be applied in certain posts, and in certain situations. Since these two post sets have the most posts of all observed in this study, it is possible that the situations result from, or in, post sets with more posts.

In all but one of the nine posts identified as non-referring, the time lag ranged from 0-8 minutes. Thus, it was hypothesized that, in general, these non-referring posts may have tended to take place during posting spurts. Because all of the non-referring posts observed took place within the first 400 minutes of the post sets, these first minutes

of both post sets can be displayed together without having to standardize time into segments. Figure 4.12 shows the occurrence of non-referring posts in time, in order to look at the surrounding activity of these posts.

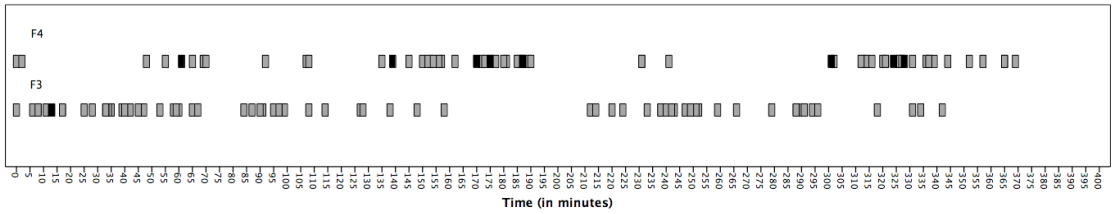

Figure 4.12. Non-referring posts, in time, in F3 and F4.

Posts in Figure 4.12 are represented by gray rectangles, with the exception of non-referring posts, which are displayed in black. As can be observed, all of the non-referring posts were found within a distinct cluster of posts, submitted within a very short period of time, suggesting that the posts were involved in posting spurts. In fact, comparing post time sequence maps in Appendix C shows that all 9 non-referring posts took place during the two largest posting spurts observed in the data, incorporating 43 posts in F3, and 65 posts in F4. It may then be argued that the non-referring posting approach to posting may be used, particularly by high-contributing post authors, during extended posting spurts. These extended posting spurts may also be responsible for the greater number of posts in these two post sets.

Capturing Participation

The maps presented in this dissertation have successively built upon one another. Each has highlighted different features of the natural history of post sets, adding new dimensions from which to understand the organization of participation. In this section, a final type of map, referred to as a *sequential lag map*, is presented. The sequential lag map provides a somewhat different perspective of post sets by placing posts not only in

sequence of occurrence but also according to the placement of the posts that they *refer back to*. This view provides a snapshot of referencing activity, and the relationships between linked posts in a given post set, ideal for making comparisons between post sets. The sequential lag maps for S1 and F5 are presented together in the following figures, Figure 4.13 and Figure 4.14. The sequential lag maps for all 10 student post sets, and the five fan post sets are presented in Appendix G.

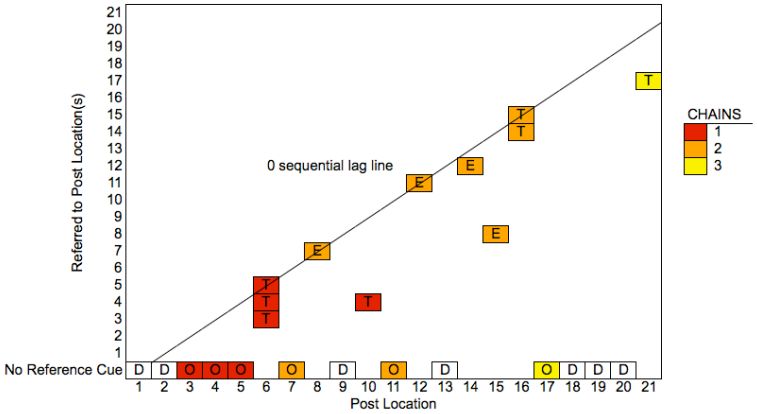

Figure 4.13. Sequential lag map, S1.

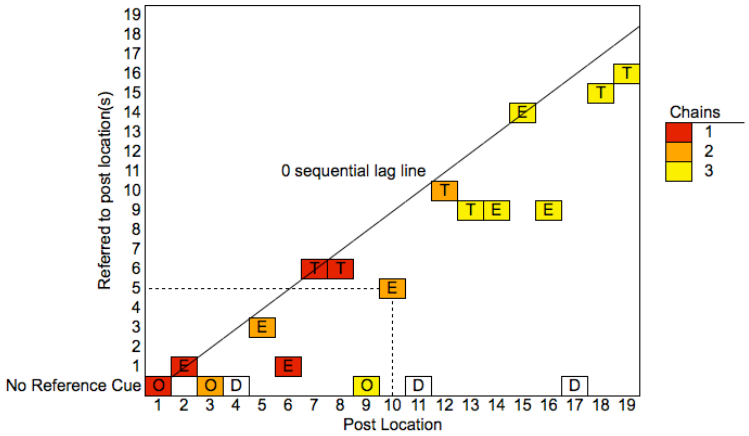

Figure 4.14. Sequential lag map, F5.

In the sequential lag maps presented in Figures 4.13 and 4.14, the x-axis represents post locations in a post set and the y-axis represents the location of the posts they *refer back to*. The dotted lines in Figure 4.14 are included to demonstrate how these

maps are read, with respect to *referring* and *referred to* post locations. In the highlighted example, the post at PL10 is *referring back to* the post at PL5. The zero-sequential lag line, running diagonally across the map, represents the potential placement of posts that *refer back to* those in the location immediately preceding it in the post sequence, or those with a *sequential lag* of zero. Deviation from the zero-sequential lag line represents the degree of sequential lag in post pairs, or the number of posts occurring between the two posts. The sequential lag is the number of posts between two posts in a post pair. This is determined by subtracting the PL of the *referring* post +1 from the PL of the *referred to* post. In the example highlighted by the dotted lines n Figure 4.14, the sequential lag between PL10 and PL5 is four. Thus, the corresponding box is placed four spaces to the right of the sequential lag line. Posts that do not include any reference cues are displayed along the zero value of the y-axis, as they do not *refer back to* any prior posts. Again, boxes are color-coded according to their chain, and displayed as white when dormant, or not involved in any chains. The letter in each box represents the post's state, where D is dormant, O is originating, E is embedded and T is terminating.

In a single representation, sequential lag maps visualize multiple levels of information related to the posts in a post set. Post states in sequence, for example, are made apparent not only by their representing letters, but also by the placement of their boxes. Dormant and originating posts are always placed along the bottom row, and embedded and terminating posts always above it. Each box located above this row can additionally be understood as representations of the post set's post pairs. The locations of both posts in a post pair, and the degree of sequential lag between them can be determined by observing the placement of their representative boxes within the map. The

number of posts *referred back to* in a single *referring* post is represented by the number of boxes in a single column along that post's location on the x-axis. The number of times a post is *referred to* in following posts is represented by the number of boxes in a single row along that post's location on the y-axis. The number of posts and post pairs in each chain is highlighted by their color-coding, which serves to also give an impression of the proportion of total posts it encompasses, and how small or large a segment of the post set's sequence it spans.

The spatial representation of posts in a sequential lag map can be understood as degrees of deviation from testable models. A sequential staircase structure in which every post *refers back to* the one immediately preceding it, for example, would plot all boxes along the sequential lag line. This model provides the testable assumptions that: 1) all posts would have equal probability of being active, 2) all posts but the first and last would have equal probability of being embedded (both *referred to* and *referring*), and 3) all post pairs would have a sequential lag of zero. Both S1 and F5 deviate to varying degrees from such a structure.

As an additional means to representing the features displayed in sequential lag maps, the values along the x and y axes for each box in a sequential lag map may be set to music. I have experimented with a software program called GarageBand to create music out of the sequential lag graphs. This can be a relatively simple process, where each of the colored boxes is treated as a single note on a scale, or more complex, where dormant posts are represented as drum beats, chains as various musical instruments, and the notes of *referred to* posts held through their sequential lag until their *referring* post occurs. Time may additionally be layered on by representing time lags as notes or breaks

of varying lengths, allowing researchers to hear musically, the relative length of time between post submissions. Each time the same note is heard it represents a post being repeatedly *referred to*, and each time a dyad or chord is heard, it represents a post that *referred back to* more than one post. When represented musically, a site following a sequential staircase model would sound like a scale. S1 and F5, on the other hand, exhibit more of a melody. "Listening to the data" in this way provides alternative perspectives on the communicative activity of post sets.

As Figures 4.13 and 4.14 show, the range of sequential lag observed in both S1 and F5 is 0-6. To see if this range was typical of the two data sets, the frequency and proportion of sequential lag values for all post pairs, by data set and overall, are compared in Table 4.25.

Table 4.25

Post Pair Sequential Lag Value Frequencies and Proportions, by Data Set and Overall

Sequential	**Student data**		**Fan data**		**Overall**	
lag	Freq.	Prop.	Freq.	Prop.	Freq.	Prop.
0	35	.159	37	.252	72	.196
1	31	.141	34	.231	65	.177
2	25	.114	27	.184	52	.142
3	20	.091	9	.061	29	.079
4	18	.082	9	.061	27	.074
5	15	.068	4	.027	19	.052
6	14	.064	2	.014	16	.044
7	8	.036	5	.034	13	.035
8	12	.055	1	.007	13	.035
9	5	.023	3	.020	8	.022
10	9	.041	0	.000	9	.025
> 10	28	.127	16	.109	44	.120
Total post pairs	**220**		**147**		**367**	

As Table 4.25 shows, the post pairs in the fan data tended to have shorter sequential lag values, with a majority (66.7%) ranging from 0-2. Those in the student data were more

distributed, with a majority (71.8%) ranging from 0-6. This can be observed in detail by comparing the sequential lag maps for the post sets from these two data sets, presented in Appendix G. The sequential lag map for S9 presented in Figure G9 is particularly demonstrative of a post set with a more distributed range of sequential lag values. The sequential lag map for F4 presented in Figure G14 is more indicative of a post set with a relatively narrow sequential lag range. Of note, in the entirety of the data, F4 is the closest to having a sequential staircase structure.

Summary

The process of linking and mapping posts allows a new level of structure to emerge, providing a great deal of additional information about posts, post sets, post authors, and the newly determined post pairs and chains. New insights into the organization of participation as it occurred in these sites can be compared within and between the data sets, offering new views into their commonalities and differences.

Findings that were common to both data sets were all related to general features of referencing activity. The range of post locations in which referencing activity began was similar in the two data sets. Just under 60% of posts included reference cues in both data sets. A majority of posts were found to be active, and tended to have shorter time lags than dormant posts. Originating posts tended to occur earlier in the post sequence, while terminating posts tended to occur later.

General features of referencing activity of post authors were similar in the two data sets as well. It was discovered that posts submitted by post authors subsequent to their first post had a strong tendency to be active. It was also found that post authors who posted earlier in the post sequence were likely to be *referred to* in following posts more often. All post authors with five or more posts were found to be involved in at least four

post pairs. Finally, smaller sub-groups of post authors communicating within chains were discovered in all post sets.

Each data set had a particularly late-starting post set, in which referencing activity began at a later point in the sequence than other post sets in their respective research settings. These two post sets, S8 and F1, were also found to have comparatively lower proportions of active posts than the other post sets. Finally, they were determined to have lower proportions of E-E post pairs.

Observed differences between the two data sets allowed for the comparison of referencing features and activity. In the fan data, post authors showed a tendency to use mostly quoting cues when *referring back to* prior posts. In addition, *referring posts* tended to *refer back to* posts that were closer in the post sequence more often than those in the student data. In the student data, they tended to include naming cues more often, both alone and in concert with quoting cues. In addition, posts including more than 1 reference cue were more commonly found in the student data than in the fan data.

Greater-referring post authors, who *refer back to* prior posts more often than they are *referred to* in following posts, were discovered in both data sets. In the student data, greater-referring post authors tended to join post sets later and include reference cues in a higher proportion of their posts. In the fan data, they showed no tendency to join post sets later, but displayed a stronger tendency to include reference cues in a higher proportion of their posts.

An additional posting approach, the non-referring approach, was identified in the fan data alone. This occurs when a post author submits a post clearly intended as a reply to a recently submitted post but does not include a reference cue to that post. It was

found that this approach tended to be utilized, particularly by enthusiastic post authors, in extended posting spurts. These extended posting spurts also appear to result in a greater number of total posts being submitted to the post set.

In the following and final chapter, a synthesis of the discoveries made in this and the previous chapter is presented. A final assessment of what was discovered about online communication via discussion boards, both in general and in the different research settings investigated, is discussed. In addition, aspects of the method, its strengths and weaknesses, application and usability, and potential for interpreting the structure of communication via discussion boards, are discussed.

CHAPTER 5
CONCLUSIONS

This dissertation has presented the development and application of a method, based in the structural approach, for mapping the natural history of micro-level features of computer-mediated communication via online discussion boards. This approach has allowed for the description of the organization of participation in online discussion board threads, taking into account the unique features of their communicative environment. The focus of the present research aimed to test the process of applying the method with data for the purpose of identifying structural features of communication. It was found that the method presented in this dissertation successfully captured participation in the observed discussion board threads, allowing for the description of the features of, and constitutive possibilities for, communication in these environments. In this chapter, discoveries pertaining to the discussion board threads observed and their communicative features and structures are described and compared. Finally, limitations of the present study are identified, with suggestions for future research and development.

Testing the Method with Data

One of the main goals of this dissertation was to develop and test the method with actual data. By applying the method to data derived from two distinct research settings, the concepts introduced, as well as their operational definitions and use in interpretation, could be assessed. In particular, these include the terms *reference cues*, *active posts*, *post pairs* and *chains*. Findings related to referencing activity indicated features representing interaction between participants, supporting the conceptual definitions introduced in the dissertation.

Reference Cues

Reference cues have been defined in this dissertation as "backwards-pointing indexical markers that allow for the establishment of links between posts." By looking at two specific types of reference cues, *naming cues*, or those in which a prior post author has been named within the text of a post; and *quoting cues*, or those in which text from a prior post has been included in a post; it was established that a particular post was referring to, and likely responding to, a specific prior post. Reference cues were tested by looking at the regularity of their use as well as the different types used, in the observed discussion board threads, or *post sets*.

It was observed that reference cues were included in approximately 58% of posts in both data sets. Of those posts that did not include reference cues, 43.3% were *referred to* in at least one following post, indicating that 76.7% of posts in total were found to be *active*, or in some way involved in referencing activity. Additionally, 14 of the 15 discussion board threads observed were found to have a majority of active posts. The likelihood of active posts encompassing a majority of the observed discussion board threads shows that referencing activity is commonplace, supporting the justification of its use for mapping and interpreting discussion board communication.

Posts including reference cues were found to include more than one cue 17.9% of the time in the student data, but only 4.3% of the time in the fan data. This may be a result of some students' interpretations of the assignment, specifically to respond to at least two of their classmate's posts by the following week. In other words, some students may have decided to fulfill this requirement in a single post. Post authors in the fan data, however, had no such requirement, indicating that there may exist a natural tendency

towards submitting a single post for each reference to a prior post. If referring posts are to be understood as "responses" to that which they are referring, then it would seem reasonable that in the interest of clarity and continuity, each response be submitted separately.

Both naming cues and quoting cues were found to be used individually more often than in combination. The likelihood of using quoting/naming cues was .264 in the student data and only .007 in the fan data. Had these proportions been higher, it may have suggested that the use of quoting/naming cues reflects an inherent weakness in quoting and/or naming cues when used alone. The low probability of the use of such combination cues, however, indicates that both naming and quoting cues were used sufficiently on their own.

As discussed in Chapter 2, in cases where the *referred to* post author has more than one prior post, there is some ambiguity regarding naming cues' indexical links. In the analysis, this was addressed by linking posts containing naming cues to the most recent prior post submitted by the named post author. It was discovered in the data that a greater proportion of posts with naming cues linked to post authors' first posts, than of posts with quoting cues or quoting/naming cues. As such, the probability of a naming cue pointing back to a post author's second or later post was .368, as opposed to .483 for quoting cues, and .423 for quoting/naming cues, emphasizing the fact that it was no more likely for naming cues to be linked to subsequent posts of post authors. This provides evidence that the coding scheme applied to naming cues had little influence, if any, over the frequency of such instances. In addition, the observed tendency to use quoting or quoting/naming cues more often when referring to posts of post authors with more than

one prior post suggests that there may be some awareness on the part of *referring* post authors that naming cues are more ambiguous in these instances.

Active Posts

As stated previously, *active posts* have been operationally defined as posts that *refer back to* a prior post, are *referred to* in a following post, or both. The allocation of the term "active" was chosen because it suggested that such posts represented more interactive situations between participants in a discussion board thread. This assumption was tested by looking at features of active posts and their authors in the data from both research settings.

It was found that in both data sets active posts were submitted after shorter periods of time lag, on average, than were dormant posts. Shorter time lags can be suggestive of a number of factors, for example a popular time frame for posting tendencies by the population using the discussion board. For example, several students were observed posting at the end of a post set's time frame, typically the night before class. The high concentration of posts within this short time frame may have resulted in shorter time lags. Shorter time lags may also be representative of the influence of a feature of the discussion board system that "highlights" threads. This was observed in the fan site, in which hundreds or thousands of threads are displayed in a paginated list, sorted by their most recent post. When posts are submitted to a thread in this setting, therefore, it is moved back up to the top of the list, bringing it into focus, and into the awareness of current users. Finally, periods of shorter time lag may result from characteristics of the posts being submitted that serve to motivate contribution. For example, in both F3 and F4 of the fan data, post authors were discussing and debating the

complicated topic of time travel in the TV show *Lost*. The inherent complexity of the issue being discussed was reflected in post authors' continued persistence in contributing posts. The comparatively high frequency and close submission of posts in these two post sets, therefore, may reflect a situation in which knowledge-sharing was occurring. Knowledge sharing is typically characterized by activities such as providing and receiving explanations, engaging in conflict and controversy, and building on others' ideas (Sawyer and Berson, 2004). Regardless of the factors influencing time lag lengths, shorter time lags in consecutive posts represent situations that reflect, or at least establish a foundation for, increased interaction between users. The fact that active posts tend to be submitted after shorter time lags provide supporting evidence that they are more likely to represent interactive situations.

The post set with the lowest proportion of active posts also had the highest proportion of one-post contributing authors. An inverse relationship between these two variables was found in both data sets. When individuals submit only one post to a discussion board thread, there is little opportunity for interaction with others. Therefore, it would be expected that post sets with a high proportion of one-post contributing authors would be generally less interactive. The fact that posts had a lower probability of being active in these less interactive post sets provides evidence that the frequency of active posts reflects the overall level of interaction associated with a post set.

When individuals submitted more than one post, their additionally submitted posts were found to be active 90.2% of the time across both data sets. Alternatively, their first posts were found to be active only slightly more than 60% of the time. This finding supports the concept that individuals returning to the discussion board thread are more

likely to be involved in interaction with others. The slightly higher proportion of active PNs > 1 in the student data may be related to the fact that students were encouraged to respond first to the opening question, and then to their classmates' posts. These findings again suggest that active posts may be indicators of interaction between participants in discussion board threads.

Post Pairs and Chains

Reference cue-based indexical links were used to group active posts into two additional units. The first was *post pairs*, or pairs of posts linked by reference cues, and the second was *chains*, or groups of post pairs linked by shared posts. Despite the fact that both post pairs and chains were created based on features of posts unrelated to their post authors or related values, post author activity was found to be captured within them. This provides the strongest supporting evidence that this method may point to features of communication as it occurs between participants in discussion board threads.

For example, post authors with five or more posts were consistently observed to be involved in at least four post pairs across both data sets. All but one of these post authors had a post that was *referred to* in at least one following post, and *referred back to* a prior post at least twice. The high probability of the most active post authors being involved in at least four post pairs shows that the post authors communicating the most were appropriately incorporated into the established segments of interaction.

It was also discovered that chains consistently captured distinct sub-groups of post authors in each post set. If chains are to be understood as interactive segments of a post set, then it is reasonable to accept that related posts are submitted by smaller groups of interacting post authors. Sets of post pairs involving the same two post authors in the

same post set were discovered in 13 of the 15 post sets. When two post authors in the same post set are observed referencing each other, it is likely indicative of a continuing dialogue between them. The fact that these sets of post pairs were found to be incorporated into the same chain 76.6% of the time, suggests that chains have a high probability of capturing this dialogue.

Comparative Analyses

In the process of applying and testing the method with data, a great deal was learned about the nature of participation in the observed post sets, both generally and with respect to the research settings from which they were derived, thereby providing insights that may be expanded upon with future research. In this section, discoveries made with respect to posting activity in post sets are discussed. By comparing data from two research settings, similarities and differences between and within the data sets provided insights into how the features of discussion board participation can vary, as well as identifying potential influencing factors. While notable differences between the features of communication in the two data set highlighted preferences of, and influences on, the two posting populations, observed similarities offer a description of the constitutive possibilities for communication in these environments.

General Features of Participation

In comparison to the fan data, post sets in the student data showed greater consistency with respect to who participated and how often. This is a reflection in part of the contextual differences between the two sites. The longest observed post set in the student data, S7, had 36 posts, while in the fan data two post sets had a much greater number of posts, F4 with 66 posts and F3 with 95. This may be a reflection of the more

established nature of the student groups, having a shared history, and the more ad-hoc nature of the fan groups, with zero-history. In a study of group information sharing in CMC, Mennecke and Valacich (1998) found that established groups tended to share less information than ad-hoc groups.

The unique time-related features of asynchronous communication were attended to in the present research as well. Overall, it was observed that posts in both research settings were submitted more often during specific times of the day. The hours during which posts were most commonly submitted varied by research setting, emphasizing the respective posting habits of these two groups. In the student data, they reflect what can be considered a typical schedule for college students, attending class during the day, and completing assignments in the evening. For the fans, posts were most often submitted during hours that tend to correspond to a the standard work day of most Americans, raising the question of just how "leisurely" the time is in which the discussion board was accessed. The overall time frames of post sets, or length of time for which posts continued to be submitted, varied by research setting as well. In the student data, this range was 4.21 - 9.44 days, with an average of approximately one week. The time frame of the post sets in the fan data was much shorter, ranging from 92 minutes to 3.35 days. In the absence of participation demands, the average sustained continuation of post submission was approximately two days. The participation requirements placed on students likely contributed to their continued participation, as compared to the fan data.

One main difference was found with respect to features of referencing activity between the two data sets, in the distribution of reference cue types used. In the fan data, quoting cues were used alone 93.2% of the time, whereas the students expressed a more

even distribution of all three types, using quoting cues 43.2% of the time, naming cues 30.5% of the time, and quoting/naming cues 26.4% of the time. The demonstrated preference among student participants for using naming cues alone and in combination with quoting cues, may be a reflection of the face-to-face component of relationships between participants in this research setting. The students knew each other personally, meeting each week during class, thus increasing the likelihood that they would refer to one another by name. Again, the participants in the fan site were ad-hoc, zero history groups, and not likely to have known each other in this way. This indicates that the tendency to state the name of the previous post author to whom one is referring is more likely when groups have a shared history, and offline relationships exist. Previous research finding behavioral norms representing group cohesion to be present in more established groups supports this discovery (Dorfman & Stephan, 1984; Menneck & Valacich, 1998).

Individual Differences in Participation

Individual differences in orientation to posting and referencing activity additionally emerged through the mapping of the natural history of post sets. By observing post author activity both within and across post sets, characterizations of post authors and their adoption of identified posting approaches can be described. While some post authors were observed consistently applying particular posting approaches, others were observed adopting such approaches only some of the time. This implies that while individual differences may reflect certain posting preferences or tendencies, situational factors may also play a role. Considerations as to what those factors may be are discussed

in this section, in conjunction with a summary of the most salient posting approaches identified.

Enthusiastic Post Authors

It was observed that certain post authors tended to display an enthusiastic approach to posting, joining post sets often, early, and posting throughout. In the student data, two post authors clearly stood out as enthusiastic post authors, displaying this approach in the majority of observed post sets. As these two post authors received the two highest grades in the course, it is likely that their consistent enthusiastic approach was a reflection of a motivation to perform well. In the fan data, enthusiastic post authors were not as consistently found across post sets, because only five discussion board threads out of hundreds were sampled. Three enthusiastic post authors in the fan data were identified by their participation in multiple post sets and adoption of an enthusiastic approach in at least one. Their contribution to more than one of the sampled post sets suggests that these post authors may generally tend to contribute to threads in the "General Discussion" section of the discussion board more often than other post authors in the fan data. The fact enthusiastic post authors were only observed adopting this approach in some of the post sets they were observed contributing to indicates that, in the more natural setting of the fan data, the enthusiastic posting approach may be more related to a genuine interest in the topic(s) being addressed.

Transient Post Authors

A transient approach to posting, in which post authors submit the entirety of their posts in close sequence, and within the same segment of standardized time, was identified only in the student data. Transient post authors were regularly observed joining post sets

later, and their participation suggested a single contributing visit to the discussion board thread. Two of the 10 transient post authors were observed adopting this approach in a majority of the post sets they contributed to. ID9 was observed doing so 83.3% of the time, and ID12 100% of the time. Overall, both of these students were also very low contributing post authors relative to their data set. This indicates that the transient strategy may be adopted more often by post authors who have generally less motivation to contribute. The remaining eight post authors that were observed adopting the transient approach did so in much lower proportions, ranging from .100 to .375 of the post sets contributed to. The fact that eight of the 10 transient post authors only occasionally adopted this approach suggests that its adoption may be more related to mitigating factors or circumstances. For students, this may reflect the occasional week in which completing assignments on time suffered as a result of coursework overload or personal challenges.

ID12 was also uniquely characterized as the only post author with more than five posts that was never *referred to* in following posts. This was likely due to ID12's consistently late entry into the post set, observable in all four of the post sets contributed to. Because very few students were observed using the message board after it was no longer a required component of the course, a majority may have considered it of little intrinsic value, and were unlikely to respond to such late-submitted posts. Thus, late submissions of transient post authors were not likely to be *referred to*. For this reason, transient post authors were responsible for the "flattened" appearance of the one-post contributing author tails, which were observed in S1, S2, S5 and S6. These tails in the student data may be described more appropriately as fewer-post contributing author tails, since they all include the late two to three post submissions of transient post authors.

Restarting Newcomer Post Authors

An additional posting approach observed only in the fan data was that of restarting newcomer post authors. Restarting newcomer post authors were characterized by joining post sets in a new time frame, generally at the start of a posting spurt involving mostly new post authors. This is likely a result of a feature in the discussion board software, which highlights recently posted-to threads by relocating them back to the top of the threaded queue. As previously discussed, this feature was especially important in the research setting of the fan data, because the post sets observed were situated amongst hundreds of other discussion board threads, collectively categorized under the "General Discussion" section of the site.

Because restarting newcomers submitted posts at a later time than the previous set of post authors, it may be that an entirely different group of participants were using the website in this new time frame. This likely accounts for the set of mostly new post authors joining the post set at the time the thread was "restarted". These new post authors were referred to as invited newcomers, since the thread highlighting likely brought it into their awareness, in essence "inviting" them to participate. Restarting newcomers were successful at restarting new posting activity 82.3% of the time. They were successful at re-engaging post authors who participated in earlier time frames 29.4% of the time. These earlier post authors were thusly characterized as re-engaged post authors. It is probable that in these cases, re-engaged post authors were still using the site, or perhaps just happened to have accessed it within the same time frame that the post set was "re-started."

One uniquely extreme case of a restarting newcomer was observed in the fan data, in which a post author joined the post set by posting after an exceptionally long time lag, lasting days rather than minutes or hours. This post author, ID55, did this in two of the observed post sets, and represented the only restarting newcomer post author to adopt this approach more than once. Further, these were the only two observed posts of ID55, and they were responsible for the relatively long gaps in time in F2 and F4, as seen in their post time sequence maps, presented in Appendix D. ID55's two restarting attempts were relatively unsuccessful, as only one invited newcomer submitted a following post. As these threads were most likely deeply "buried" within the list of discussion board threads at the time of ID55's post submission, this may represent a unique approach to seeking out threads on the part of ID55, perhaps scanning pages of thread titles to locate those of particular interest, or using a keyword search tool to do so.

The restarting newcomer approach highlights the increased influence of system features and processes in highly active discussion boards. The relocation of threads to the top of the thread queue begins to play a role in who participates and when participation occurs. Additionally, the identification of threads via a keyword search function can result in users directly influencing which threads are "bumped" back into focus. Bumping and searching both require an action on the part of both the user, and a system response to that action. Such activities, therefore, support adaptive structuration theory (Browning & Stephens, 2004; Poole & Desanctis, 1990), in that they represent communicative processes as shaping, and being shaped by, use of technology-enabled resources.

Non-Referring Post Authors

Finally, a non-referring posting approach was observed in just two post sets from the fan data, F3 and F4. These represented a rare posting approach, observed in scenarios where a post author clearly intended a submission to be in direct response to a recently submitted post, but failed to include a reference cue indicating as much. Nine such non-referring post submissions were observed, all of which were submitted during extended posting spurts. All nine, in fact, were submitted during the two posting spurts encompassing the greatest number of posts in the data. It was concluded, therefore, that the non-referring posting approach was related both to participation in extended periods of shared time frames, and the frequency of posts submitted within those time frames. It is likely that in these extended near-synchronous posting situations, an understanding developed between the post authors, that they were addressing one another's posts, which had been submitted only minutes earlier, eliminating the need for quoting or naming cues within these posts.

Despite the fact that several post authors consistently used reference cues during these two extended posting spurts, three adopted the non-referring approach. Two of the non-referring post authors, IDs 41 and 91, each submitted a single non-referring post, while the third, ID50, submitted seven. All three non-referring post authors contributed additional posts in which they did include reference cues. It should be noted that ID50 had the highest total number of posts of all post authors from the fan data, and was classified, along with ID41, as an enthusiastic post author. While both ID41 and ID50 were the first to submit non-referring posts in their respective post sets, ID91 was observed doing so in direct response to one of ID50's posts. ID91, therefore, may have

adopted the non-referring posting approach only after observing some of the non-referring posts of ID50. This observation may have captured intrinsic "learning" of a new posting approach by post author ID91. A tendency to adopt the non-referring posting approach, then, may be more likely to be observed among high contributing post authors, and may increase as discussion board experience is gained.

Although in the present research only quoting cues and naming cues were sought out and coded as reference cues, it was discovered in the non-referring posts that there may be other, less obvious, indicators of reference cues, such as addressing "you" and "your" or using responsive words like "but" or "so." As the non-referring approach has demonstrated, perhaps there are other types of reference cues that can be discovered and accounted for in an additional round of "second-order" coding and analysis. This could highlight additional referencing activity that may have occurred in the observed post sets, and provide even further insights into posts' linking relationships.

Constitutive Possibilities for Communication

Revisiting the constitutive perspective, which understands everyday communication as the space in which macro-level meanings are constructed and re-constructed, it can be argued that this method has brought to light features of the framework within which such processes take place (Mokros, 2003). In the present analysis, a constitutive understanding of the observed communication was outlined by observing commonalities that emerged in discussion board use, regardless of the unique participation constraints or affordances present in the two research settings.

For example, the first 10 posts opening each post set were found to be submitted by a relatively consistent number of post authors. By looking at the post sequence maps

in Appendix A, it can be observed that, in the student post sets, these first posts were submitted by a total of 5-9 individuals (excluding myself), and in the fan post sets, a total of 6-8 individuals. This is similar to what Cockett (2000) found in her study of a face-to-face meeting of 11 participants, where just six individuals claimed the first ten turns at talk. These findings suggest that there may be a natural tendency to limit the initial "steps" taken in both types of communicative environments to a smaller number of individuals, negotiating opening moves and setting the stage for the remainder of the interaction.

A similar finding in the present research found the range of referencing activity to begin at approximately the same post locations in the observed post sets. Removing the two identified late-starting post sets, and my posts in the student data, the first post to be referenced was always the first or second, and the first post to reference another was always between PL2 and PL5. Again, this suggests that, like turn-taking, there may be particular referencing tendencies among humans in the opening moves of an interaction. Similarly, particularly in the fan data, it was observed that the final moves of an interaction tended to peter out with fewer and fewer returning post authors. These findings highlight the natural processes for which it is established who contributes, how, and when.

Other previously mentioned features of referencing activity were additionally found to be consistent across the two data sets. In particular, the finding that post authors' returning contributions were found to be active approximately 30% more often that their first posts suggest an increasing propensity for engaging with others as one returns to an interaction. In addition, the nearly equal proportions of active posts and

post including reference cues in both data sets was rather remarkable, and indicate a specific level of necessity for clarification and coherence when interacting with others in the online environment. Finally, the observation that post authors were typically observed referencing one another in distinct chain-based sub-groups mirrors processes discovered in group research where, in naturally-occurring casual public settings, people are typically observed interacting within smaller groups (Bakeman, 1974). This suggests that, in both face-to-face and computer-mediated situations, humans display a natural tendency for interacting with a smaller number of individuals.

These findings support the notion that humans tend to subconsciously negotiate and establish shared rules and norms for interacting, even in the absence of the co-presence that occurs with shared space and time frames. Once established through the method presented here, the observation of these rules and norms in practice can provide invaluable insight into how individuals are able to make sense of what is occurring in their interaction, despite the mediated nature of the communication.

Mapping Data

The use of mapping in this method, acting as an interface between qualitative and quantitative interpretation, allowed for an analysis of the structure of communication in the observed discussion board threads. By observing these maps and the associated visualizations they create, the data begin to tell a story, providing information about the nature of participation in a post set. The structure of interaction was observed through the features of the various maps presented, which allowed for new insights into interaction to be generated and tested. The extent to which a discussion board thread represents interaction, and the nature of that interaction, was reflected in the various configurations

constructed, and their placement within these maps. An example of this is provided by the one-post contributing author tails observed in some post author sequence maps, which highlight the processes of "thread death."

Post author sequence maps (Appendix C) can also be used to define the nature of post author participation in a thread, by analyzing the placement of posts throughout the four quadrants of the map. For example, Figures C8 and C11 show that the post author sequence maps for S8 and F1, the two post sets whose late starts to referencing activity were found to indicate low levels of interaction, are noticeably empty in their upper right quadrants, a result of relatively few persistently contributing post authors. F1 is additionally characterized as having a distinct one-post contributing author tail, indicating that it died out with no re-engagement of the earlier post authors. The overall shape of these maps offers additional information about the number of participants relative to the number of posts in a post set. For example, that of F4, presented in Figure C4, was particularly rectangular in shape as a result of fewer authors contributing more posts, while that of F1 in Figure C11 was more square-like due to its uniquely high post to post author ratio.

Post time sequence maps (Appendix D) distinguish posting spurts, displaying the ebb and flow of posting activity in time. In general, these maps demonstrate the distinct nature of posting within each of the two data sets. Student post sets typically contained a higher number of posting spurts distributed throughout their time frames, corresponding to differences in posting activity relative to time of day and throughout the course of a week. The post time sequence map for S3 presented in Figure D3 shows a particularly distinct example of such a structure. The fan post sets demonstrated tendencies for a

single, large posting spurt at the beginning of the post set, followed by periods of relative inactivity, with occasional small posting spurts throughout the remainder of the post set, reflecting, perhaps, decreasing interest as the specifics of the newest TV episode were exhausted. A notable example of this structure is provided in the post time sequence map for F3 presented in Figure D13.

Linked chain maps (Appendix E) represent the subsets of communicating individuals and their related messages. By showing the formations constructed by the post pair states, the level of interaction is portrayed. For example, the linked chain maps for the late-starting S8 and F1, presented in Figures E8 and E11, respectively, represent their lower proportions of embedded-embedded post pairs. The resulting structures for these two post sets, therefore, present as a series of particularly flat chains, with relatively few step-like formations. In contrast, the linked chains maps for post sets such as S9, shown in Figure E9, and S10, shown in Figure E10, contain many multi-post pair involved posts and step-like formations. Additionally, the first chains (CH1s) in both S9 and S10 encompass higher proportions of the post sets' posts and post authors. The maze-like appearance of these chains in the linked chain maps shows that this process may be driven by a smaller number of highly linked posts, which were actively involved in referencing activity. Such highly linked posts in the analysis may be representative of what is referred to as a "dramatizing message" in symbolic convergence theory (Benoit et al., 2001). Dramatizing messages use imaginative language to call to mind fantasy themes, or narratives that a particular group of people use in order to create a shared reality. When such messages occur, members of the social group often become animated or excited, participating more often and at a faster pace, with the conversation becoming

more lively (Bormann, 1972). The chains that include these highly linked posts, therefore, may represent what is referred to as a fantasy chain, or the conversation resulting from dramatizing messages.

Finally, sequential lag maps best represent the relative sequentiality of posts in a post set (Appendix G). By observing the placement of posts to the right of the sequential lag line, the distribution of post pairs and chains in the sequence of a post set can be easily identified. The sequential lag map for S9 presented in Figure G9, for example, demonstrates that the post pairs of CH1 represent posts that are farther apart from one another in the sequence. This indicates that several posts from earlier in the sequence continued to be linked to throughout the post set's natural history, thereby supporting the notion that certain posts in this post set may be particularly "activating" in nature, eliciting responses even as the discussion progressed. Alternatively, as shown in Figure G14, the same map for F4 presents a structure most like the sequential staircase model, in which posts are linked to more recently submitted prior posts. From the information learned regarding the nature of posting activity in F4, it may be that such a structure is additionally indicative of a higher level of near-synchronicity in discussion board thread posting.

All four types of maps provide useful insights into the nature and level of interaction in a post set. In addition, each map may be altered in various ways as a way to highlight different dimensions, in order to learn much more about these data than what is presented in this dissertation. The color-coding of chains, for example, could be layered onto a post author sequence map, in order to better characterize the nature of activity between the sub-grouped post authors they represent. A comparable layering of chains

onto the post time sequence maps may be used to distinguish how, specifically, they occurred within posting spurts.

Deciding what to display within "post boxes" of a particular map may be adjusted in order to obtain specialized information of particular interest to the researcher. By replacing post number values with post states in the post author sequence maps, more information can be learned regarding the evolution of post states for authors as they continue to contribute. Replacing post author values with post states in the post time sequence maps will highlight any patterns that may exist for post states' occurrence in time. Post boxes in linked post chain maps may be adjusted to include post states, instead of post authors, to provide additional insights into post pair states. To extend such an analysis, the post boxes above the bottom row in sequential lag maps may be analyzed as post pairs, showing their post pair states. Alternatively, placing post author values in the boxes of sequential lag maps can highlight additional posting approaches by identifying post authors with tendencies for *referring back to* posts nearer, or further, in the sequence. Finally, additional values and variables not considered in the present analysis, whether derived inductively or through the application of theory, can be similarly layered on to these maps so that more may be learned about their occurrence in the context of participation in discussion board threads.

Such observations as those presented in the previous paragraphs allow for the description of post sets as "more" or "less" interactional. They also provide for new insights into similarities and differences in the nature of observed interactions, generating new hypotheses for testing. As more research is conducted with the use of the method

presented, more can be learned about how post sets may be categorized and understood, and what is typical or characteristic of certain types of interactional situations.

Limitations

While the goal of this analysis was focused on the development and testing of the method, and not on determining generalizable features of discussion board communication, procedural limitations are acknowledged, to be considered with respect to final conclusions. Procedural challenges faced in data collection and sampling from the highly active fan discussion board are addressed, as well as potential consequences of the criteria used. Additionally, the non-referring approach as a potential weakness of the method is discussed in further detail.

While the student data were derived from an enclosed setting, and the time frame of posting activity had ended at the time of this analysis, the decided-upon procedure for sampling in the fan data was to discontinue observation and data collection after one week from the initial post submission. Although post contributions in all threads from this research setting ceased after 3.35 days within the week of observation, these threads remain in open and public status for an indefinite period of time. This indicates that additional posting activity may have occurred beyond the week of observation, and such activity would not have been accounted for in the present analysis. The archived nature of discussion board threads lends itself to the possibility of continued contribution beyond a chosen time frame of study, however an analytical decision about that time frame must be made in order to ensure consistency.

In my experience with discussion boards, I have observed threads where participation continues for months, or even years. These threads, however, are typically

"stickied" by discussion board moderators, which means they remain in a select group of threads consistently displayed above the list of non-stickied threads. This is normally done when topics are considered particularly relevant, important, or pervasive to the overall focus of the discussion board. The activities which surround moderator decisions to "sticky" or "unsticky" certain threads may prove an interesting layer of analysis to observe in the natural history of a post set, considering moderators often make these choices in response to requests of other discussion board members. However, because of their elevated status, participation in such threads should not be considered as typical discussion board use.

An alternative to determining set time frames for observation may be to study threads that have been "locked" by discussion board moderators, meaning they are closed to further post contributions. It should be noted that locked threads are likely to have significantly different characteristics than other threads. This is because moderators typically make the decision to lock a thread when disagreements between participants become insulting in nature, or other extreme or atypical situations. While such instances would likely provide interesting insights into the nature of discussion board communication breakdowns, the locked status of a thread should be considered as a factor.

In this study, an additional criterion for sampling threads from the fan site was that the first observed thread to reach 15 posts would be observed. This was done for the sake of consistency, and to ensure that threads would have enough posts to warrant investigation. Nonetheless, this particular research decision may have resulted in analyzing threads with unique features in the fan data, particularly with respect to early

activity. An alternative approach may be to sample threads from discussion board sites at random, regardless of their placement within the site archive. However, this approach eliminates the benefit of observing the submission of posts in real time, and may not appropriately account for the potential for moderator interventions in a discussion board thread.

Finally, although the method applied was generally found to be successful at appropriately capturing the organization of participation in discussion board threads, one weakness in non-referring posts was discovered. As previously discussed, non-referring posts are rare post submissions that do not include any reference cues, but which were discovered to be responding to recently submitted posts. Consequently, these posts and their post authors were not properly linked in a single chain. The circumstances consistently observed surrounding non-referring posts suggest that the strength of this method may decrease for high-contributing post authors, during extended periods of near-synchronous posting activity. Thus, such situations should be treated with particular attention in future similarly conducted analyses. It may be that every communicative situation has its own version of reference cues that can be used to re-arrange participation in a more meaningful way.

Directions for Future Research

When used as a basis for understanding communication in the online environment, the possibilities for which the developed method can be used are plentiful. Once the structure of communication in post sets has been identified, other important aspects of online communication may be layered on to it for comparative analysis or to learn more about how values and variables of interest occur in context. Comparative

studies using this approach could be conducted to learn more about how particular aspects of research sites and participants are related to the nature of online participation. Universally observed characteristics of computer-mediated communication in discussion boards could additionally be discovered through such analyses. Aspects of posts, such as speech acts, pronoun use or topical keywords, can be further investigated to learn more about how they occur in context, and individual differences between their authors. Finally, social networks could be identified by mapping post authors' links to one another in post pairs and chains.

Specific discoveries made in the present analysis may be expanded upon with constructs of applied theory. Aspects of knowledge-sharing in online communication can be further investigated to learn more about how such activities occur, and are identified, in discussion board participation. The content of highly linked posts and their chains can be further investigated to determine their status as dramatizing messages and fantasy chains in symbolic convergence theory. The distinct post author sub-groupings discovered may provide an avenue for learning more about the structure and processes of group communication in the online environment. Further research into zero history and history groups in online discussion boards may find additional features of computer-mediated communication typically characteristic of these groups. Social network analysis may also prove fruitful for learning more about greater-referring and greater-referred-to post authors, and to learn more about how this status may reflect the roles of such individuals in the networks of a particular discussion board.

While the present analysis looked at posts on a message board, this has implications for all sorts of communicative environments, written or spoken, computer-

mediated or face-to-face, constrained by the design of a technology or communication system. The challenge for researchers is in determining and testing the reference cues observed in each unique situation. Through the testing and screening of data, further types of cues may be determined and, again, tested.

Overall, the method presented in this dissertation proved to be extremely effective in identifying the features of computer-mediated communication, as it occurs through online discussion boards. It should be noted, however, that modification or development of software applications specific to coding and representing data in the mapping of the natural history of computer-mediated communication, will vastly increase the applicability of this method. The maps in the present analysis were created manually, with the use of Microsoft Word and Microsoft Excel. These tools are neither intended for, nor robust enough to efficiently handle the requirements of data visualization and analysis in such research. With the application of software capable of allowing for the input of thoughtfully coded and linked data points and visualizing them in useful mapped displays such as those presented in this dissertation, this approach may prove to be extremely effective as a methodological starting point for all types of research in the field of communication.

APPENDIX A: POST SEQUENCE MAPS, ALL POST SETS

Table A1

Post Sequence Map, S1

PL	PA	PN	TL
1	1	1	
2	2	1	1657
3	3	1	1054
4	4	1	270
5	5	1	1165
6	2	2	130
7	6	1	239
8	7	1	102
9	8	1	3
10	8	2	2
11	9	1	211
12	4	2	68
13	10	1	1218
14	9	2	2620
15	5	2	32
16	4	3	194
17	11	1	15
18	12	1	43
19	13	1	230
20	14	1	4351
21	14	2	2

Table A2

Post Sequence Map, S2

PL	PA	PN	TL
1	1	1	
2	2	1	1628
3	3	1	12
4	4	1	999
5	5	1	56
6	6	1	304
7	7	1	256
8	8	1	803
9	9	1	61
10	10	1	29
11	3	2	140
12	11	1	62
13	11	2	3
14	2	2	133
15	12	1	135
16	13	1	26
17	3	3	855
18	7	2	27
19	14	1	1579
20	14	2	12
21	8	2	99
22	9	2	9
23	4	2	95
24	4	3	5
25	2	3	0
26	8	3	113
27	6	2	5
28	15	1	24
29	16	1	271
30	16	2	20
31	17	1	2757
32	17	2	4
33	17	3	1559

Table A3

Post Sequence Map, S3

PL	PA	PN	TL
1	1	1	
2	2	1	342
3	3	1	1378
4	4	1	101
5	2	2	149
6	5	1	862
7	6	1	220
8	3	2	39
9	7	1	80
10	7	2	43
11	8	1	1108
12	8	2	2
13	9	1	53
14	10	1	93
15	11	1	263
16	5	2	1043
17	2	3	184
18	12	1	1153
19	12	2	4
20	12	3	10
21	12	4	7
22	3	3	251
23	4	2	12
24	12	5	82
25	2	4	1312
26	2	5	28
27	2	6	2
28	13	1	8
29	6	2	95
30	14	1	99
31	15	1	198
32	15	2	8
33	15	3	29

Table A4

Post Sequence Map, S4

PL	PA	PN	TL
1	1	1	
2	2	1	2994
3	3	1	91
4	4	1	144
5	5	1	99
6	6	1	1040
7	7	1	144
8	8	1	136
9	8	2	5
10	9	1	43
11	10	1	79
12	5	2	705
13	6	2	288
14	11	1	160
15	3	2	141
16	12	1	947
17	12	2	38
18	12	3	23
19	6	3	162
20	4	2	104
21	11	2	1461
22	11	3	4
23	11	4	2
24	13	1	63
25	13	2	8
26	2	2	5
27	7	2	241
28	14	1	165
29	14	2	10
30	14	3	12

Table A5

Post Sequence Map, S5

PL	PA	PN	TL
1	1	1	
2	2	1	24
3	3	1	1419
4	4	1	137
5	5	1	380
6	2	2	1440
7	6	1	703
8	7	1	169
9	3	2	176
10	8	1	39
11	3	3	12
12	9	1	87
13	10	1	1084
14	2	3	109
15	3	4	177
16	5	2	1038
17	4	2	121
18	5	3	65
19	6	2	397
20	2	4	784
21	7	2	264
22	9	2	99
23	8	2	19
24	8	3	103
25	11	1	147
26	12	1	172
27	13	1	348
28	14	1	4008
29	14	2	3
30	14	3	5

Table A6

Post Sequence Map, S6

PL	PA	PN	TL
1	1	1	
2	2	1	25
3	3	1	1379
4	4	1	83
5	5	1	28
6	2	2	743
7	5	2	699
8	6	1	11
9	7	1	108
10	3	2	230
11	8	1	43
12	9	1	70
13	10	1	1108
14	2	3	93
15	2	4	4
16	4	2	100
17	3	3	85
18	5	3	1166
19	5	4	8
20	2	5	1237
21	2	6	4
22	7	2	234
23	9	2	131
24	6	2	38
25	8	2	88
26	8	3	14
27	11	1	64
28	12	1	141
29	13	1	35
30	14	1	9
31	15	1	4403
32	15	2	3
33	15	3	3

Table A7

Post Sequence Map, S7

PL	PA	PN	TL
1	1	1	
2	2	1	1196
3	3	1	99
4	4	1	68
5	5	1	3
6	5	2	5
7	6	1	11
8	4	2	21
9	7	1	203
10	8	1	48
11	4	3	873
12	9	1	80
13	5	3	43
14	3	2	23
15	6	2	48
16	10	1	100
17	11	1	316
18	4	4	18
19	8	2	1462
20	12	1	861
21	13	1	< 1
22	12	2	3
23	13	2	11
24	14	1	462
25	15	1	781
26	15	2	10
27	4	5	131
28	10	2	64
29	3	3	134
30	16	1	43
31	16	2	8
32	17	1	128
33	17	2	4
34	10	3	38
35	11	2	128
36	11	3	7

Table A8

Post Sequence Map, S8

PL	PA	PN	TL
1	1	1	
2	2	1	1232
3	3	1	107
4	4	1	52
5	5	1	3
6	6	1	22
7	7	1	218
8	8	1	39
9	5	2	844
10	5	3	3
11	9	1	97
12	4	2	34
13	6	2	72
14	3	2	12
15	10	1	149
16	11	1	229
17	12	1	935
18	12	2	17
19	13	1	71
20	13	2	1
21	8	2	62
22	14	1	325
23	15	1	757
24	9	2	225
25	16	1	351
26	10	2	< 1
27	10	3	13
28	11	2	88
29	11	3	10
30	17	1	89

Table A9

Post Sequence Map, S9

PL	PA	PN	TL
1	1	1	
2	2	1	304
3	3	1	998
4	4	1	122
5	5	1	216
6	4	2	32
7	6	1	114
8	7	1	1058
9	7	2	7
10	2	2	348
11	8	1	507
12	9	1	289
13	2	3	239
14	4	3	34
15	5	2	86
16	10	1	12
17	11	1	120
18	11	2	1
19	12	1	8
20	3	2	55
21	7	3	2858
22	8	2	207
23	11	3	196
24	5	3	598
25	13	1	357
26	8	3	24
27	2	4	10
28	14	1	2
29	2	5	< 1
30	10	2	71
31	10	3	5
32	15	1	67
33	9	2	85
34	12	2	25
35	12	3	11

Table A10

Post Sequence Map, S10

PL	PA	PN	TL
1	1	1	
2	1	2	13
3	2	1	113
4	3	1	136
5	4	1	400
6	5	1	69
7	5	2	8
8	6	1	1090
9	6	2	25
10	2	2	306
11	7	1	505
12	8	1	304
13	4	2	65
14	2	3	160
15	9	1	113
16	10	1	159
17	11	1	24
18	3	2	15
19	3	3	3
20	12	1	1280
21	6	3	1578
22	7	2	202
23	10	2	199
24	12	2	1064
25	12	3	3
26	13	1	49
27	8	2	110
28	11	2	63
29	11	3	3
30	14	1	466

Table A11

Post Sequence Map, F1

PL	PA	PN	TL
1	1	1	
2	2	1	3
3	3	1	1
4	4	1	29
5	5	1	2
6	1	2	8
7	4	2	6
8	6	1	71
9	6	2	2
10	7	1	0
11	6	3	3
12	6	4	3
13	8	1	99
14	9	1	60
15	10	1	13
16	11	1	34
17	9	2	2
18	8	2	3
19	12	1	34
20	13	1	8
21	14	1	9
22	15	1	4
23	16	1	7
24	14	2	1
25	14	3	2
26	17	1	1965
27	18	1	35
28	9	3	65
29	19	1	23
30	20	1	185
31	21	1	5
32	22	1	49
33	23	1	680
34	24	1	79
35	25	1	14

Table A12

Post Sequence Map, F12

PL	PA	PN	TL
1	1	1	
2	2	1	6
3	3	1	2
4	4	1	1
5	5	1	1
6	1	2	1
7	2	2	3
8	5	2	< 1
9	1	3	4
10	6	1	< 1
11	3	2	2
12	7	1	2
13	1	4	1
14	5	3	1
15	8	1	3
16	1	5	1
17	3	3	2
18	9	1	235
19	10	1	10
20	9	2	1
21	11	1	25
22	3	4	8
23	12	1	70
24	13	1	17
25	14	1	17
26	15	1	4319
27	16	1	91

Table A13

Post Sequence Map, F3

PL	PA	PN	TL
1	1	1	
2	2	1	6
3	3	1	2
4	1	2	< 1
5	4	1	3
6	3	2	2
7	5	1	< 1
8	4	2	4
9	6	1	< 1
10	3	3	8
11	7	1	3
12	8	1	5
13	1	3	< 1
14	9	1	< 1
15	3	4	1
16	4	3	< 1
17	3	5	1
18	1	4	4
19	9	2	1
20	1	5	2
21	9	3	3
22	4	4	2
23	3	6	6
24	9	4	5
25	10	1	1
26	4	5	1
27	3	7	5
28	11	1	2
29	9	5	17
30	4	6	3
31	12	1	3
32	9	6	1
33	5	2	4
34	13	1	2
35	1	6	1
36	3	8	1
37	14	1	9
38	15	1	6
39	1	7	13
40	14	2	1
41	14	3	10
42	2	2	10
43	1	8	10
44	16	1	54
45	17	1	2
46	16	2	6
47	3	9	4
48	17	2	9
49	18	1	5
50	6	2	< 1
51	2	3	2

52	16	3	< 1
53	19	1	< 1
54	19	2	2
55	6	3	1
56	1	9	4
57	3	10	2
58	19	3	< 1
59	16	4	2
60	1	10	1
61	17	3	7
62	20	1	7
63	2	4	13
64	17	4	9
65	2	5	< 1
66	1	11	1
67	20	2	2
68	1	12	3
69	20	3	2
70	21	1	22
71	2	6	13
72	3	11	3
73	4	7	8
74	22	1	88
75	1	13	4
76	23	1	35
77	24	1	12
78	7	2	3
79	25	1	199
80	9	7	435
81	26	1	57
82	27	1	41
83	28	1	3
84	29	1	259
85	30	1	5
86	31	1	8
87	19	4	27
88	32	1	10
89	9	8	11
90	33	1	4
91	14	4	35
92	28	2	14
93	3	12	14
94	34	1	10
95	5	3	3

Table A14

Post Sequence Map, F4

PL	PA	PN	TL
1	1	1	
2	2	1	2
3	3	1	46
4	4	1	7
5	5	1	6
6	4	2	4
7	6	1	4
8	7	1	1
9	5	2	22
10	1	2	15
11	8	1	1
12	4	3	27
13	5	3	4
14	9	1	6
15	5	4	5
16	5	5	2
17	4	4	1
18	9	2	1
19	9	3	2
20	5	6	1
21	9	4	5
22	5	7	8
23	4	5	2
24	5	8	1
25	9	5	2
26	5	9	< 1
27	9	6	2
28	5	10	3
29	4	6	1
30	9	7	4
31	5	11	1
32	10	1	1
33	9	8	< 1
34	4	7	2
35	5	12	1
36	11	1	41
37	12	1	< 1
38	4	8	10
39	5	13	60
40	13	1	1
41	12	2	10
42	4	9	1
43	5	14	1
44	5	15	2
45	9	9	4

46	12	3	1
47	4	10	< 1
48	5	16	3
49	9	10	2
50	5	17	< 1
51	12	4	1
52	5	18	1
53	12	5	3
54	4	11	5
55	5	19	1
56	12	6	< 1
57	4	12	2
58	14	1	5
59	5	20	8
60	15	1	5
61	14	2	8
62	12	7	< 1
63	14	3	4
64	16	1	48
65	13	2	55
66	17	1	2802

Table A15

Post Sequence Map, F5

PL	PA	PN	TL
1	1	1	
2	2	1	6
3	3	1	1
4	4	1	4
5	5	1	4
6	6	1	3
7	1	2	9
8	7	1	3
9	8	1	4
10	3	2	1
11	9	1	2
12	5	2	9
13	10	1	10
14	4	2	< 1
15	10	2	1
16	11	1	< 1
17	12	1	2
18	4	3	< 1
19	13	1	33

APPENDIX B: LINKED POST MAPS, ALL POST SETS

Table B1

Linked Post Map, S1

PL	PA	PN	LPL	LPA	LPN	RCT
1	1	1				
2	2	1				
3	**3**	**1**				
4	**4**	**1**				
5	**5**	**1**				
6	2	2	3	3	1	Q
6	*2*	*2*	4	4	1	Q
6	*2*	*2*	5	5	1	Q/N
7	**6**	**1**				
8	**7**	**1**	7	6	1	Q/N
9	8	1				
10	8	2	4	4	1	N
11	**9**	**1**				
12	**4**	**2**	11	9	1	Q/N
13	10	1				
14	**9**	**2**	12	4	2	N
15	**5**	**2**	8	7	1	Q/N
16	4	3	14	9	2	N
16	*4*	*3*	15	5	2	Q/N
17	**11**	**1**				
18	12	1				
19	13	1				
20	14	1				
21	14	2	17	11	1	Q

Table B2

Linked Post Map, S2

PL	PA	PN	LPL	LPA	LPN	RCT
1	1	1				
2	**2**	**1**				
3	3	1				
4	**4**	**1**				
5	5	1	2	2	1	N
6	**6**	**1**				
7	**7**	**1**	4	4	1	N
8	**8**	**1**				
9	**9**	**1**				
10	**10**	**1**				
11	**3**	**2**	4	4	1	Q
12	**11**	**1**				
13	**11**	**2**	10	10	1	Q
14	2	2				
15	**12**	**1**				
16	13	1	6	6	1	N
16	*13*	*1*	7	7	1	N
16	*13*	*1*	11	3	2	N
17	**3**	**3**	13	11	2	Q
18	7	2	15	12	1	Q/N
19	**14**	**1**				
20	14	2	12	11	1	Q
21	8	2	17	3	3	N
22	**9**	**2**	10	10	1	Q/N
23	4	2	17	3	3	N
24	**4**	**3**				
25	**2**	**3**	19	14	1	Q
26	8	3	22	9	2	N
26	*8*	*3*	24	4	2	N
26	*8*	*3*	25	2	3	N
27	**6**	**2**	24	4	2	Q/N
27	***6***	*2*	25	2	3	N
28	15	1				
29	16	1	19	14	1	Q/N
29	*16*	*1*	25	2	3	N
30	16	2	27	6	2	Q
31	17	1				
32	17	2	8	8	1	Q
33	17	3	9	9	1	Q

Table B3

Linked Post Map, S3

PL	PA	PN	LPL	LPA	LPN	RCT
1	1	1				
2	2	1				
3	**3**	**1**				
4	4	1				
5	**2**	**2**	3	3	1	Q/N
6	**5**	**1**				
7	6	1				
8	**3**	**2**	6	5	1	N
9	**7**	**1**				
10	7	2	5	2	2	N
10	*7*	*2*	6	5	1	N
11	8	1				
12	**8**	**2**	6	5	1	N
13	**9**	**1**	12	8	2	N
14	**10**	**1**				
15	**11**	**1**				
16	5	2	14	10	1	Q
16	*5*	*2*	15	11	1	Q/N
17	**2**	**3**	8	3	2	Q
18	12	1	17	2	3	Q
19	**12**	**2**				
20	**12**	**3**	14	10	1	Q
21	12	4	9	7	1	Q/N
22	3	3	13	9	1	Q/N
23	**4**	**2**	20	12	3	Q
24	12	5	23	4	2	Q
25	2	4	6	5	1	Q/N
26	2	5	19	12	2	Q
27	2	6	23	4	2	Q/N
28	13	1				
29	6	2	20	12	3	Q/N
30	14	1				
31	15	1				
32	15	2	3	3	1	Q/N
33	15	3	19	12	2	Q/N

Table B4

Linked Post Map, S4

PL	PA	PN	LPL	LPA	LPN	RCT
1	1	1				
2	2	1				
3	**3**	**1**				
4	**4**	**1**				
5	5	1				
6	**6**	**1**	4	4	1	N
7	**7**	**1**				
8	8	1				
9	8	2	3	3	1	N
9	*8*	*2*	6	6	1	N
10	**9**	**1**				
11	**10**	**1**				
12	**5**	**2**	11	10	1	Q
13	6	2	12	5	2	Q
14	**11**	**1**	3	3	1	N
14	***11***	***1***	4	4	1	N
15	**3**	**2**	14	11	1	Q
16	**12**	**1**				
17	12	2	10	9	1	Q
18	**12**	**3**	7	7	1	Q
19	**6**	**3**	16	12	1	Q
19	***6***	***3***	18	12	3	N
20	4	2	14	11	1	Q
20	*4*	*2*	15	3	2	N
20	*4*	*2*	16	12	1	Q
21	11	2	16	12	1	Q
22	11	3	11	10	1	Q
23	11	4	15	3	2	Q/N
24	13	1				
25	13	2	14	11	1	Q
26	2	2	19	6	3	Q/N
27	7	2	18	12	3	Q
28	14	1				
29	14	2	11	10	1	Q/N
30	14	3	14	11	1	Q/N

Table B5

Linked Post Map, S5

PL	PA	PN	LPL	LPA	LPN	RCT
1	1	1				
2	**2**	**1**				
3	**3**	**1**	2	2	1	N
4	**4**	**1**	2	2	1	Q
5	5	1	4	4	1	N
6	2	2	3	3	1	Q/N
7	**6**	**1**	4	4	1	N
8	**7**	**1**				
9	**3**	**2**	8	7	1	Q
10	**8**	**1**				
11	3	3				
12	9	1				
13	**10**	**1**				
14	2	3	13	10	1	Q
15	**3**	**4**	13	10	1	N
16	**5**	**2**	7	6	1	Q
16	***5***	***2***	9	3	2	Q
17	**4**	**2**	15	3	4	Q
18	5	3	17	4	2	N
19	6	2	16	5	2	Q/N
19	*6*	*2*	17	4	2	N
19	*6*	*2*	18	5	3	Q/N
20	2	4	17	4	2	Q/N
21	7	2	10	8	1	Q/N
22	9	2				
23	8	2	4	4	1	Q/N
24	8	3	15	3	4	Q/N
25	**11**	**1**				
26	12	1	4	4	1	Q
27	13	1				
28	14	1				
29	14	2	25	11	1	Q
30	14	3	10	8	1	Q

Table B6

Linked Post Map, S6

PL	PA	PN	LPL	LPA	LPN	RCT
1	1	1				
2	**2**	**1**				
3	**3**	**1**	2	2	1	N
4	**4**	**1**	2	2	1	N
4	***4***	***1***	3	3	1	N
5	**5**	**1**	2	2	1	Q
5	***5***	***1***	4	4	1	N
6	**2**	**2**	5	5	1	Q
7	**5**	**2**	6	2	2	Q
8	6	1				
9	**7**	**1**				
10	**3**	**2**				
11	**8**	**1**				
12	9	1				
13	**10**	**1**				
14	**2**	**3**	7	5	2	Q
15	**2**	**4**	4	4	1	Q/N
16	**4**	**2**	10	3	2	Q
17	3	3	16	4	2	Q/N
18	**5**	**3**	15	2	4	Q
19	**5**	**4**	14	2	3	Q
20	**2**	**5**	18	5	3	Q
21	2	6	19	5	4	Q
22	**7**	**2**	20	2	5	Q
23	**9**	**2**				
24	6	2	22	7	2	Q/N
24	*6*	*2*	23	9	2	Q/N
25	8	2	13	10	1	Q/N
26	8	3	9	7	1	Q/N
27	11	1				
28	12	1				
29	13	1				
30	14	1	11	8	1	Q/N
31	15	1				
32	15	2	10	3	2	Q
33	15	3	11	8	1	Q

Table B7

Linked Post Map, S7

PL	PA	PN	LPL	LPA	LPN	RCT
1	1	1				
2	2	1				
3	**3**	**1**				
4	**4**	**1**				
5	5	1				
6	**5**	**2**	3	3	1	Q
7	**6**	**1**				
8	4	2	3	3	1	N
8	*4*	*2*	6	5	2	Q/N
9	7	1				
10	**8**	**1**				
11	**4**	**3**	7	6	1	Q/N
12	9	1	7	6	1	Q
13	5	3	10	8	1	Q
14	**3**	**2**	6	5	2	Q
15	6	2	11	4	3	Q/N
16	**10**	**1**				
17	**11**	**1**				
18	**4**	**4**	17	11	1	Q
19	8	2	16	10	1	Q/N
19	*8*	*2*	18	4	4	Q
20	**12**	**1**				
21	13	1				
22	**12**	**2**	17	11	1	N
23	**13**	**2**	18	4	4	N
24	14	1				
25	**15**	**1**	4	4	1	Q
26	15	2	22	12	2	Q
27	4	5	25	15	1	Q
28	10	2	25	15	1	Q/N
29	3	3	23	13	2	Q/N
30	16	1				
31	**16**	**2**	14	3	2	Q
32	17	1				
33	**17**	**2**	31	16	2	Q
34	10	3	20	12	1	Q
35	11	2	18	4	4	Q
36	11	3	33	17	2	Q

Table B8

Linked Post Map, S8

PL	PA	PN	LPL	LPA	LPN	RCT
1	1	1				
2	2	1				
3	3	1				
4	**4**	**1**				
5	5	1				
6	**6**	**1**				
7	**7**	**1**				
8	**8**	**1**				
9	5	2	4	4	1	Q
10	5	3	8	8	1	Q/N
11	**9**	**1**				
12	4	2	11	9	1	Q
13	**6**	**2**	11	9	1	Q
14	3	2				
15	10	1				
16	11	1				
17	**12**	**1**				
18	12	2	7	7	1	N
19	**13**	**1**				
20	13	2				
21	8	2	11	9	1	N
21	*8*	*2*	13	6	2	Q/N
21	*8*	*2*	17	12	1	Q
22	**14**	**1**				
23	**15**	**1**	19	13	1	Q
24	9	2	19	13	1	Q
25	16	1				
26	10	2	23	15	1	Q/N
27	10	3	22	14	1	Q/N
28	11	2	19	13	1	Q/N
29	11	3	6	6	1	Q/N
30	17	1				

Table B9

Linked Post Map, S9

PL	PA	PN	LPL	LPA	LPN	RCT
1	1	1				
2	**2**	**1**				
3	**3**	**1**	2	2	1	N
4	**4**	**1**				
5	5	1				
6	4	2				
7	**6**	**1**	2	2	1	N
7	***6***	***1***	3	3	1	N
8	**7**	**1**	2	2	1	N
8	***7***	***1***	3	3	1	N
9	7	2	4	4	1	Q
10	2	2	4	4	1	Q
11	**8**	**1**	4	4	1	N
12	**9**	**1**				
13	2	3	7	6	1	Q/N
14	**4**	**3**	3	3	1	Q
15	5	2	8	7	1	Q
15	***5***	***2***	11	8	1	N
16	10	1				
17	**11**	**1**	11	8	1	N
17	***11***	***1***	12	9	1	N
17	***11***	***1***	15	5	2	N
18	11	2				
19	**12**	**1**				
20	**3**	**2**	11	8	1	Q
21	**7**	**3**	14	4	3	N
21	***7***	***3***	15	5	2	Q
22	8	2	15	5	2	N
22	*8*	*2*	20	3	2	Q
22	*8*	*2*	21	7	3	Q/N
23	**11**	**3**	7	6	1	N
23	***11***	***3***	21	7	3	Q
24	**5**	**3**				
25	13	1				
26	8	3	24	4	3	N
27	2	4	23	11	3	Q
28	14	1	19	12	1	N
29	**2**	**5**	21	7	3	Q/N
30	**10**	**2**	3	3	1	Q
31	10	3	12	9	1	N
31	*10*	*3*	29	2	5	Q
32	15	1				
33	9	2	17	11	1	Q
34	12	2	23	11	3	Q
35	12	3	3	3	1	N
35	*12*	*3*	30	10	2	Q

Table B10

Linked Post Map, S10

PL	PA	PN	LPL	LPA	LPN	RCT
1	1	1				
2	1	2				
3	**2**	**1**				
4	**3**	**1**	3	2	1	N
5	4	1	3	2	1	N
5	*4*	*1*	4	3	1	N
6	**5**	**1**				
7	5	2	4	3	1	N
8	6	1	4	3	1	Q/N
9	6	2	6	5	1	Q
10	**2**	**2**	4	3	1	Q/N
11	**7**	**1**	4	3	1	N
11	*7*	***1***	10	2	2	N
12	**8**	**1**				
13	**4**	**2**	12	8	1	Q/N
14	2	3	12	8	1	Q/N
15	**9**	**1**	11	7	1	Q/N
16	10	1	6	5	1	Q
17	**11**	**1**				
18	3	2	11	7	1	Q
19	**3**	**3**	17	11	1	Q
20	**12**	**1**	12	8	1	N
21	6	3	13	4	2	Q
22	7	2	13	4	2	Q
22	*7*	*2*	15	9	1	N
22	*7*	*2*	19	3	3	Q/N
23	**10**	**2**				
24	12	2	19	3	3	Q/N
25	12	3	17	11	1	Q
26	13	1				
27	8	2	23	10	2	Q
28	11	2	3	2	1	Q
29	11	3	20	12	1	Q
30	14	1				

Table B11

Linked Post Map, F1

PL	PA	PN	LPL	LPA	LPN	RCT
1	1	1				
2	2	1				
3	3	1				
4	4	1				
5	**5**	**1**				
6	1	2				
7	4	2				
8	6	1				
9	6	2				
10	7	1				
11	6	3				
12	6	4				
13	8	1	5	5	1	Q
14	**9**	**1**				
15	**10**	**1**	14	9	1	Q
16	**11**	**1**				
17	**9**	**2**	16	11	1	Q
18	8	2				
19	12	1				
20	**13**	**1**				
21	14	1				
22	15	1	16	11	1	Q
23	**16**	**1**	20	13	1	Q
24	14	2				
25	14	3	23	16	1	Q
26	17	1	5	5	1	Q
27	**18**	**1**	17	9	2	Q
28	9	3	27	18	1	Q
29	19	1	15	10	1	Q
30	20	1				
31	21	1				
32	22	1	5	5	1	Q
33	23	1				
34	24	1				
35	25	1				

Table B12

Linked Post Map, F2

PL	PA	PN	LPL	LPA	LPN	RCT
1	1	1				
2	**2**	**1**				
3	**3**	**1**				
4	4	1				
5	5	1	2	2	1	N
5	*5*	*1*	3	3	1	N
6	**1**	**2**				
7	2	2	6	1	2	Q
8	**5**	**2**	6	1	2	Q
9	**1**	**3**	8	5	2	Q
10	6	1				
11	**3**	**2**	6	1	2	Q
12	**7**	**1**				
13	**1**	**4**	11	3	2	Q
14	**5**	**3**	9	1	3	Q
14	***5***	***3***	12	7	1	N
15	**8**	**1**				
16	1	5	14	5	3	Q
17	**3**	**3**	13	1	4	Q
18	**9**	**1**				
19	10	1	18	9	1	Q
20	9	2				
21	11	1				
22	3	4				
23	12	1	15	8	1	Q
24	13	1				
25	14	1	17	3	3	Q
26	15	1				
27	16	1				

Table B13

Linked Post Map, F3

PL	PA	PN	LPL	LPA	LPN	RCT
1	**1**	**1**				
2	**2**	**1**				
3	**3**	**1**	1	1	1	Q
4	1	2	2	2	1	Q
5	4	1	3	3	1	Q
6	**3**	**2**				
7	**5**	**1**				
8	4	2	6	3	2	Q
8	*4*	*2*	7	5	1	Q
9	**6**	**1**	6	3	2	Q
10	**3**	**3**	9	6	1	Q
11	**7**	**1**				
12	8	1				
13	**1**	**3**	11	7	1	Q
14	**9**	**1**	2	2	1	Q
14	***9***	***1***	11	7	1	Q
15	**3**	**4**				
16	4	3	14	9	1	Q
17	3	5	14	9	1	Q
18	1	4	15	3	4	Q
19	9	2	17	3	5	Q
20	1	5	13	1	3	Q
21	9	3	13	1	3	Q
22	4	4	21	9	3	Q
23	3	6	22	4	4	N
24	9	4	23	3	6	Q
25	10	1	3	3	1	Q
26	4	5	23	3	6	Q
27	3	7	26	4	5	Q
28	11	1				
29	**9**	**5**	26	4	5	Q
30	**4**	**6**	29	9	5	Q
31	**12**	**1**	3	3	1	Q
32	**9**	**6**	30	4	6	Q
33	5	2	31	12	1	Q
34	**13**	**1**				
35	1	6	34	13	1	Q
36	3	8	31	12	1	Q
36	*3*	*8*	32	9	6	N
37	14	1	3	3	1	Q
38	**15**	**1**				
39	1	7	38	15	1	Q
40	14	2				
41	**14**	**3**				
42	**2**	**2**	41	14	3	Q
43	1	8	42	2	2	Q
44	16	1	10	3	3	Q
45	**17**	**1**	6	3	2	Q/N
46	**16**	**2**	45	17	1	Q
47	3	9	45	17	1	Q
48	**17**	**2**	46	16	2	Q

49	**18**	**1**				
50	6	2	48	17	2	Q
51	2	3	48	17	2	Q
52	16	3	48	17	2	Q
53	19	1	48	17	2	Q
54	**19**	**2**	49	18	1	Q
55	**6**	**3**	49	18	1	Q
56	**1**	**9**	55	6	3	Q
57	3	10				
58	**19**	**3**	56	1	9	Q
59	16	4	49	18	1	Q
60	**1**	**10**	58	19	3	Q
61	**17**	**3**				
62	**20**	**1**	61	17	3	Q
63	**2**	**4**	60	1	10	Q
64	**17**	**4**	62	20	1	Q
65	2	5	61	17	3	Q
66	1	11	63	2	4	Q
67	**20**	**2**				
68	**1**	**12**	67	20	2	Q
69	20	3	68	1	12	Q
70	21	1	3	3	1	Q
71	**2**	**6**				
72	3	11	71	2	6	Q
73	4	7	64	17	4	Q
74	**22**	**1**				
75	1	13	74	22	1	Q
76	**23**	**1**	64	17	4	Q
77	24	1				
78	7	2	76	23	1	Q
79	25	1	54	19	2	Q
80	9	7	76	23	1	Q
81	26	1				
82	27	1				
83	**28**	**1**	58	19	3	Q
84	29	1	76	23	1	Q
85	30	1				
86	31	1				
87	**19**	**4**	83	28	1	Q
88	32	1				
89	9	8	87	19	4	N
90	33	1				
91	14	4				
92	**28**	**2**				
93	3	12				
94	34	1				
95	5	3	92	28	2	Q

Table B14

Linked Post Map, F4

PL	PA	PN	LPL	LPA	LPN	RCT
1	**1**	**1**				
2	2	1				
3	3	1	1	1	1	Q
4	**4**	**1**				
5	**5**	**1**				
6	**4**	**2**	5	5	1	Q
7	6	1				
8	7	1	5	5	1	Q
9	**5**	**2**	6	4	2	Q
10	1	2	4	4	1	Q
11	**8**	**1**	1	1	1	Q
12	4	3	9	5	2	Q
13	**5**	**3**				
14	**9**	**1**	13	5	3	N
15	**5**	**4**	14	9	1	Q
16	**5**	**5**	14	9	1	Q
17	**4**	**4**	13	5	3	Q
18	**9**	**2**	15	5	4	Q
19	9	3	16	5	5	Q
20	**5**	**6**	17	4	4	Q
20	***5***	***6***	18	9	2	Q
21	9	4				
22	**5**	**7**				
23	4	5	20	5	6	Q
24	5	8				
25	9	5	22	5	7	Q
26	**5**	**9**				
27	9	6	26	5	9	N
28	5	10				
29	4	6	26	5	9	Q
30	9	7				
31	5	11				
32	**10**	**1**				
33	**9**	**8**				
34	4	7	33	9	8	Q
35	5	12	32	10	1	Q
36	11	1				
37	**12**	**1**				
38	4	8	37	12	1	Q
39	**5**	**13**				
40	13	1				
41	**12**	**2**	39	5	13	Q
42	**4**	**9**	39	5	13	Q
43	**5**	**14**	41	12	2	Q
44	**5**	**15**	42	4	9	Q
45	9	9	44	5	15	N
46	12	3	43	5	14	Q
47	4	10	44	5	15	Q
48	**5**	**16**				
49	9	10				
50	5	17				

51	12	4	26	5	9	Q
52	**5**	**18**				
53	12	5	48	5	16	Q
54	**4**	**11**	48	5	16	Q
55	**5**	**19**	54	4	11	Q
56	12	6	52	5	18	Q
57	**4**	**12**	55	5	19	Q
58	14	1				
59	**5**	**20**	57	4	11	Q
60	15	1				
61	14	2	11	8	1	Q
62	**12**	**7**	59	5	20	Q
63	14	3	62	12	7	Q
64	16	1				
65	13	2				
66	17	1	6	4	2	Q

Table B5

Linked Post Map, F5

PL	PA	PN	LPL	LPA	LPN	RCT
1	**1**	**1**				
2	2	1	1	1	1	Q
3	**3**	**1**				
4	4	1				
5	**5**	**1**	3	3	1	Q
6	**6**	**1**	1	1	1	Q
7	1	2	6	6	1	Q
8	7	1	6	6	1	Q
9	**8**	**1**				
10	**3**	**2**	5	5	1	Q
11	9	1				
12	5	2	10	3	2	Q
13	10	1	9	8	1	Q
14	**4**	**2**	9	8	1	Q
15	**10**	**2**	14	4	2	Q
16	**11**	**1**	9	8	1	Q
17	12	1				
18	4	3	15	10	2	Q
19	13	1	16	11	1	Q

APPENDIX C: POST AUTHOR SEQUENCE MAPS, ALL POST SETS

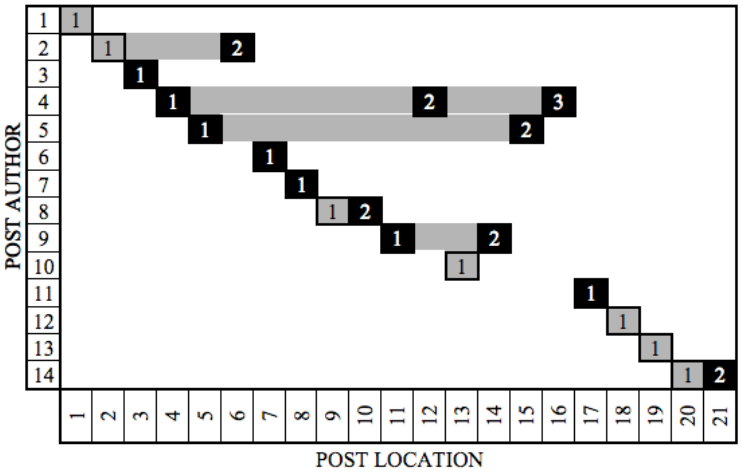

Figure C1. Post author sequence map, S1.

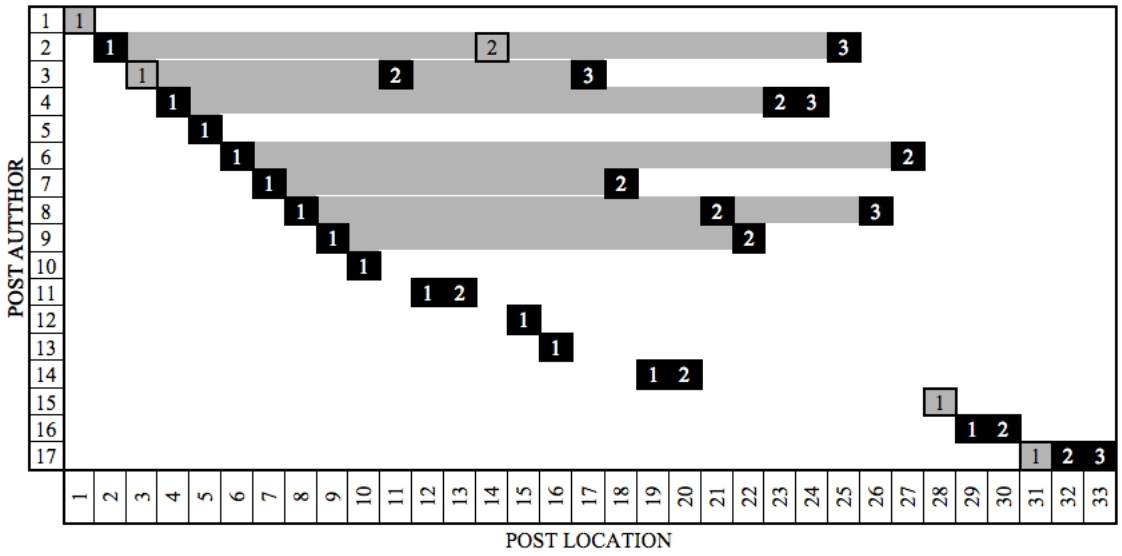

Figure C2. Post author sequence map, S2.

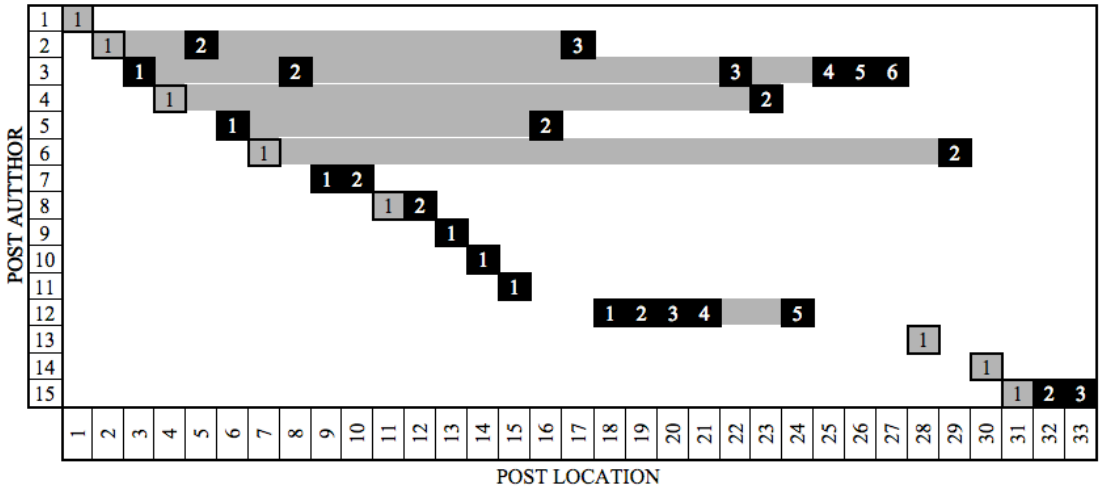

Figure C3. Post author sequence map, S3.

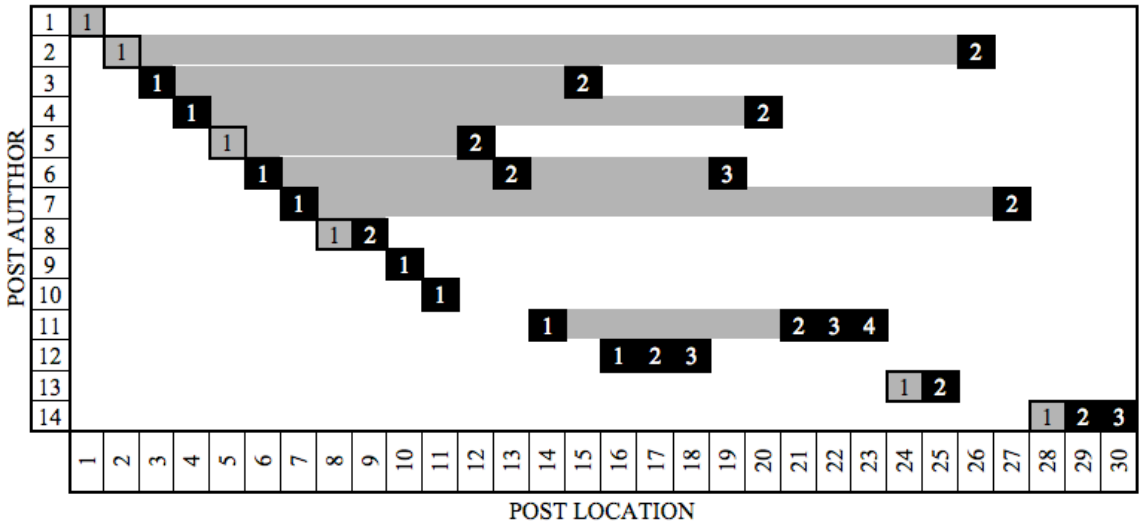

Figure C4. Post author sequence map, S4.

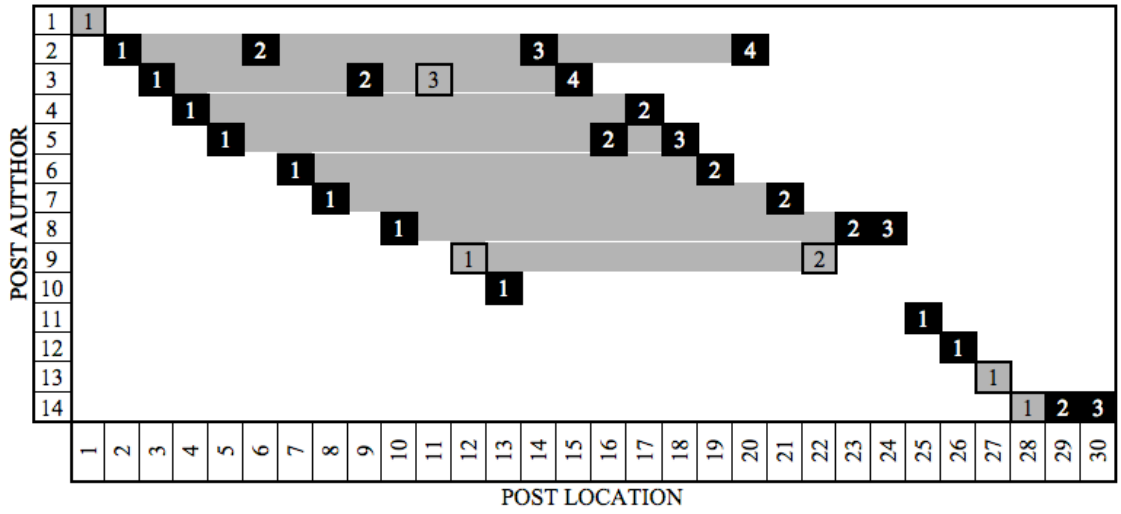

Figure C5. Post author sequence map, S5.

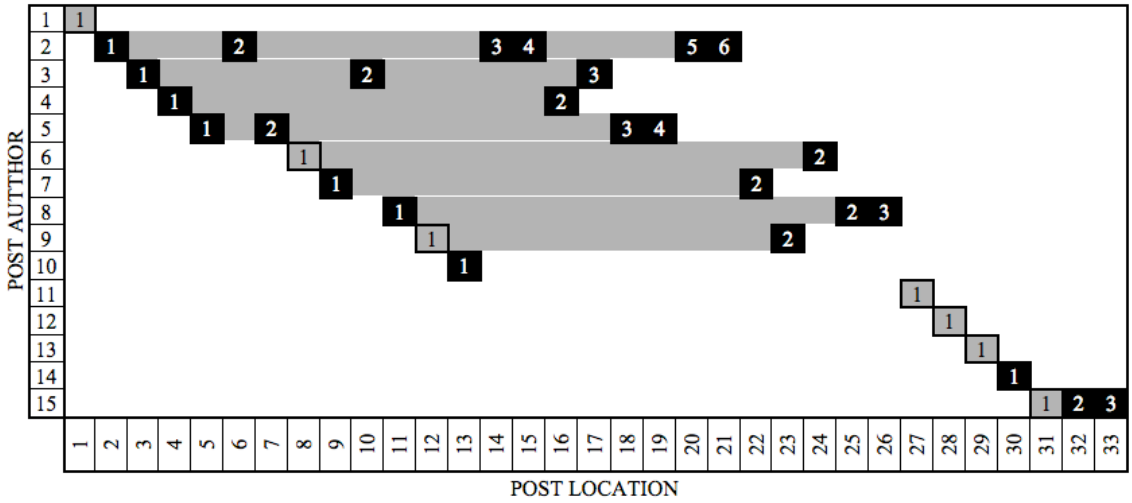

Figure C6. Post author sequence map, S6.

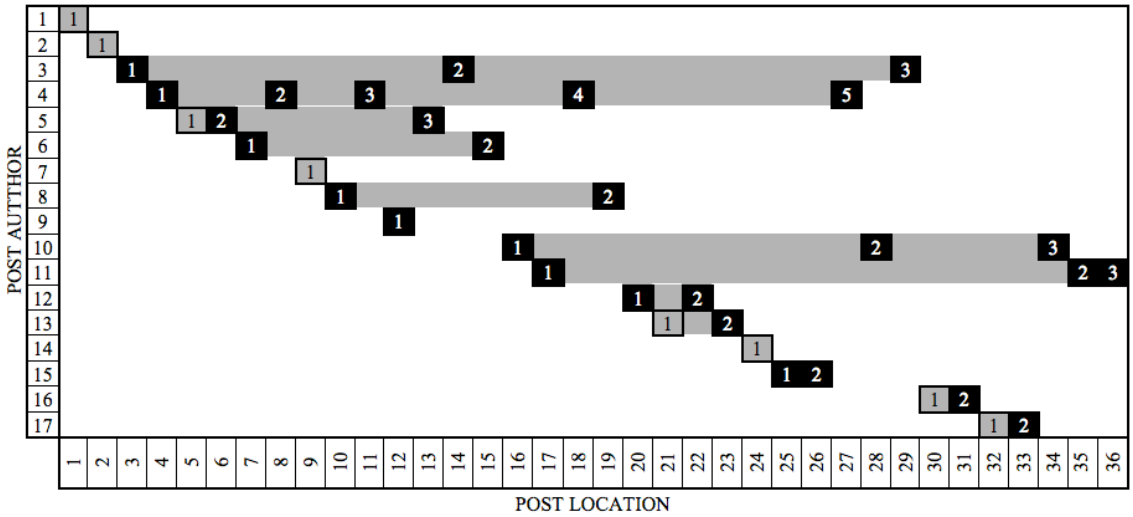

Figure C7. Post author sequence map, S7.

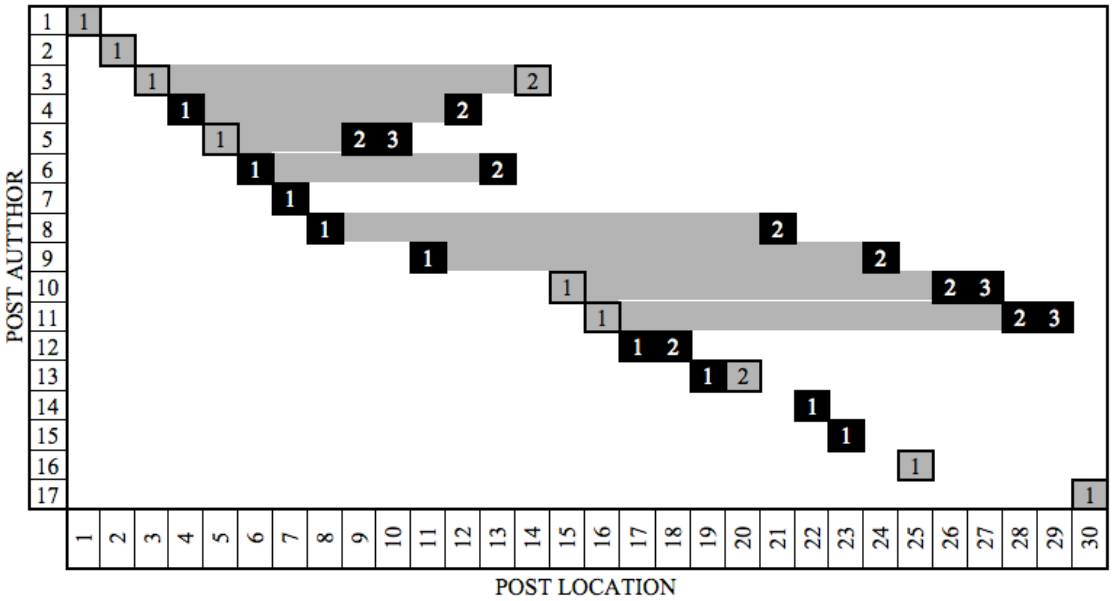

Figure C8. Post author sequence map, S8.

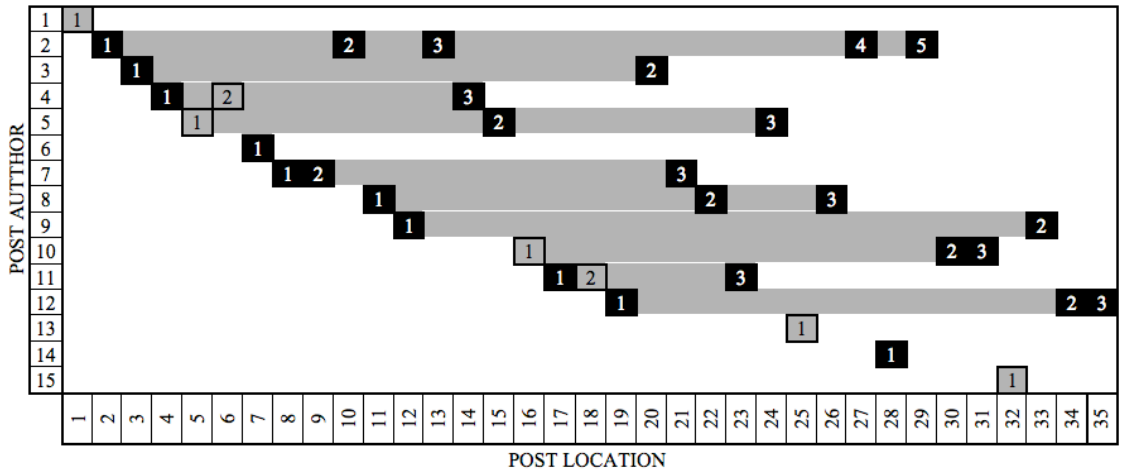

Figure C9. Post author sequence map, S9.

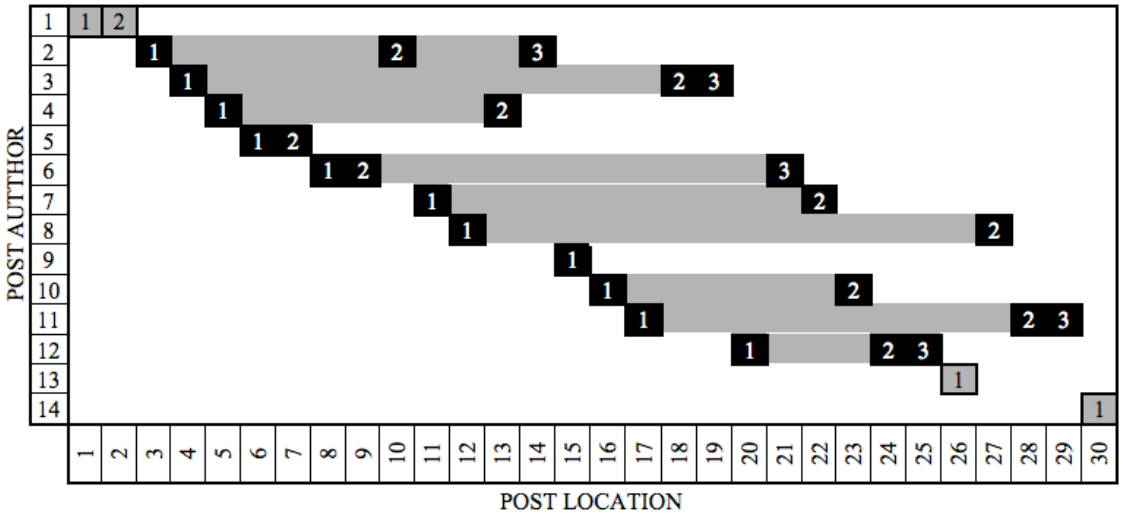

Figure C10. Post author sequence map, S10.

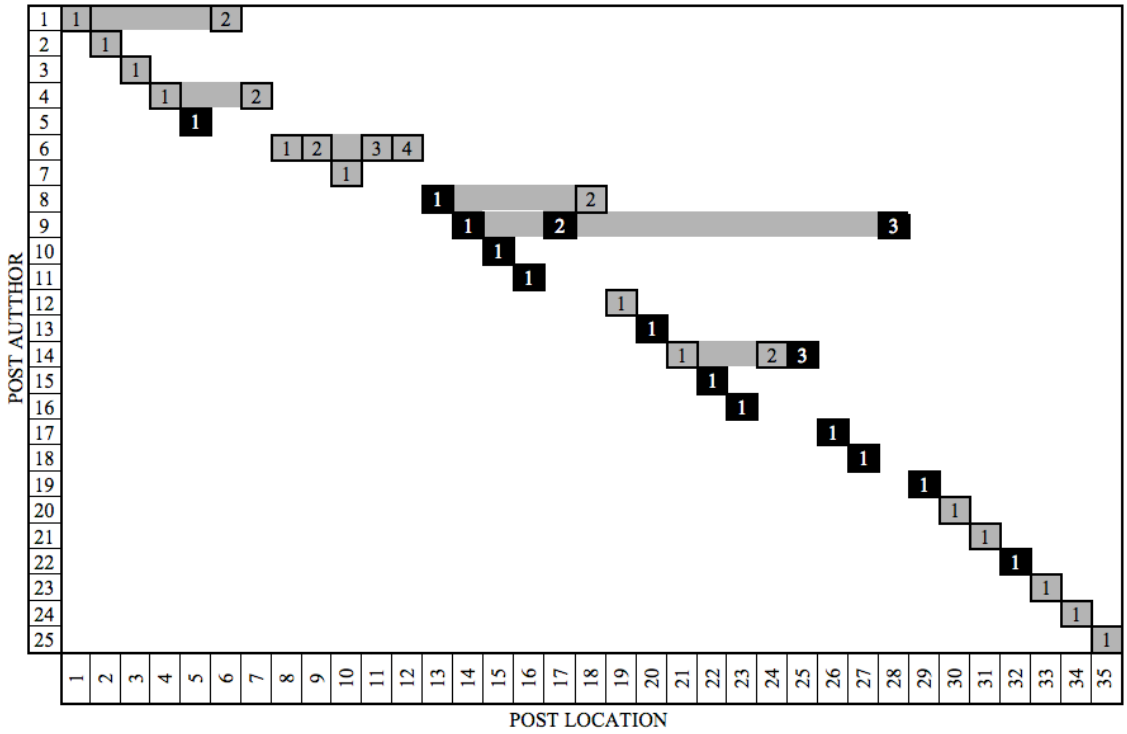

Figure C11. Post author sequence map, F1.

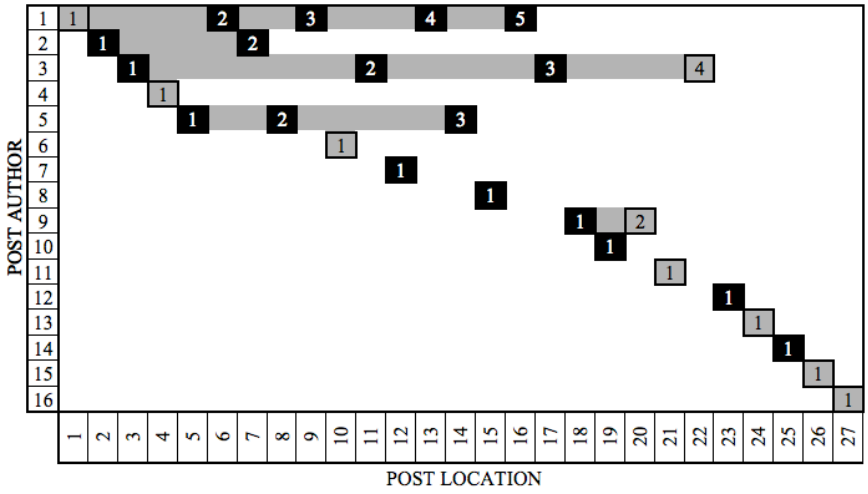

Figure C12. Post author sequence map, F2.

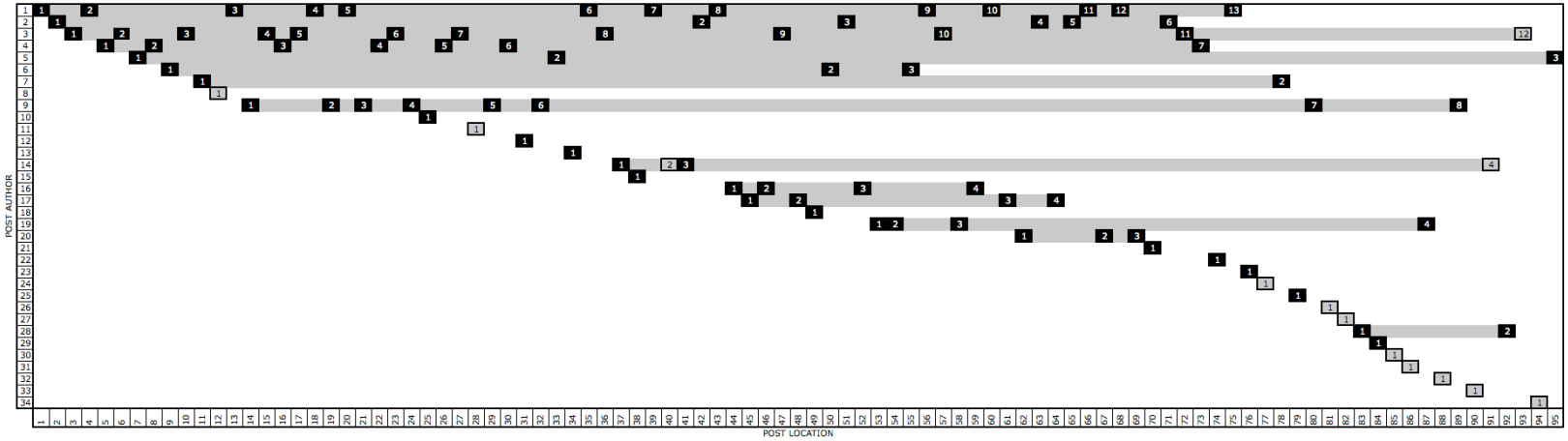

Figure C13. Post author sequence map, F3.

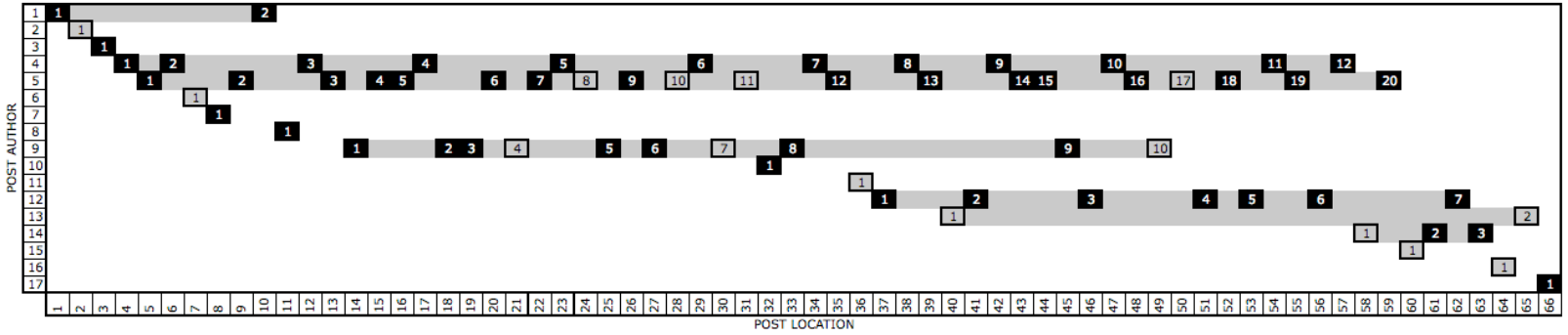

Figure C14. Post author sequence map, F4.

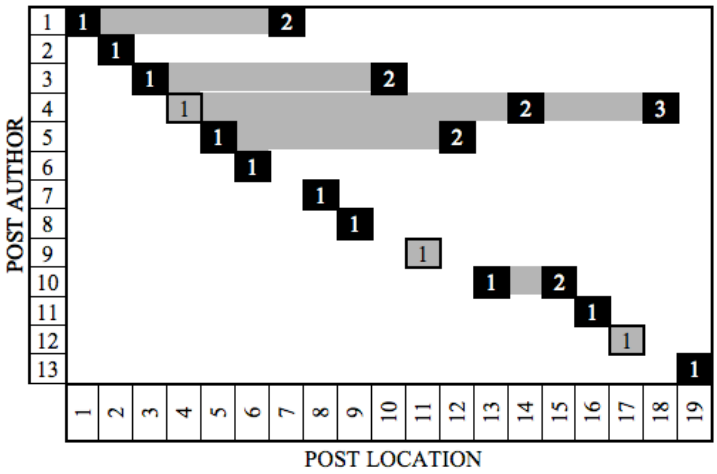

Figure C15. Post author sequence map, F5.

APPENDIX D: POST TIME SEQUENCE MAPS, ALL POST SETS

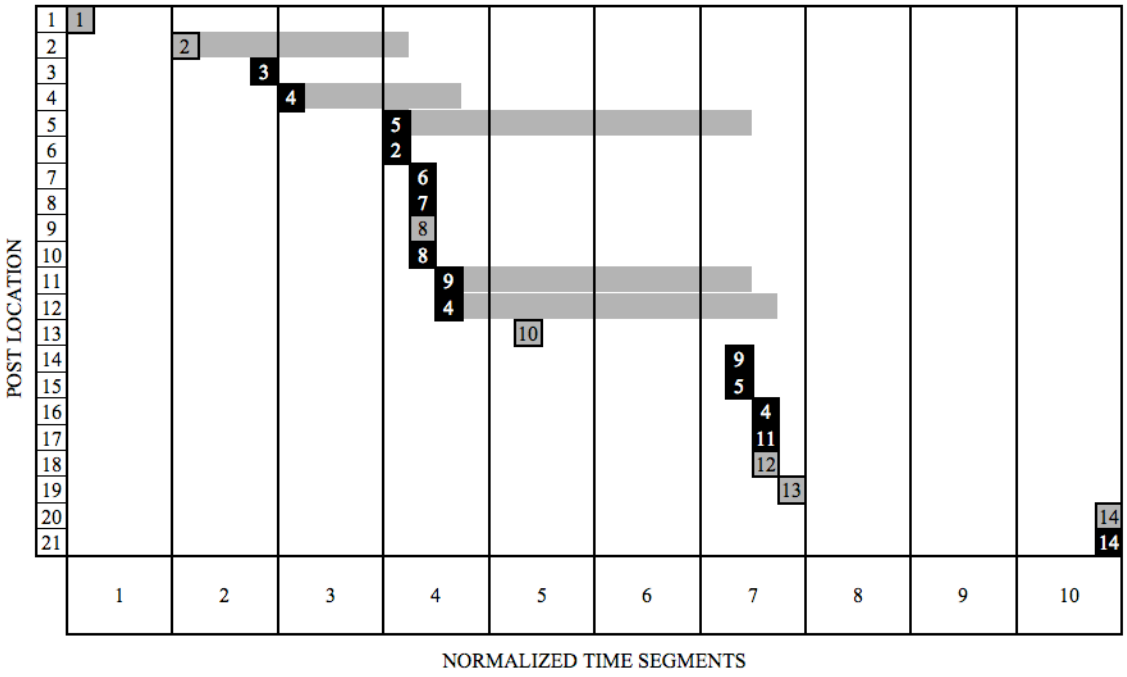

Figure D1. Post time sequence map, S1.

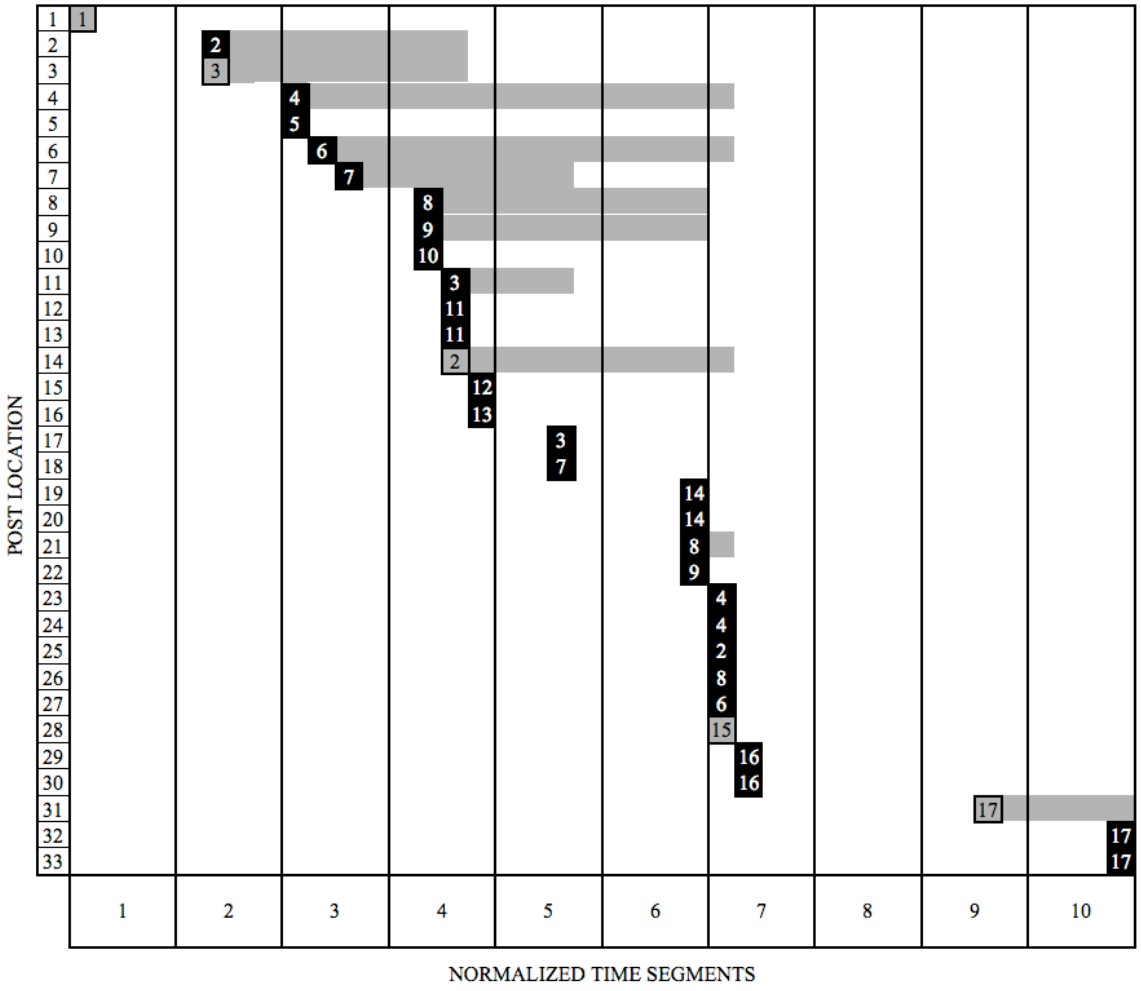

Figure D2. Post time sequence map, S2.

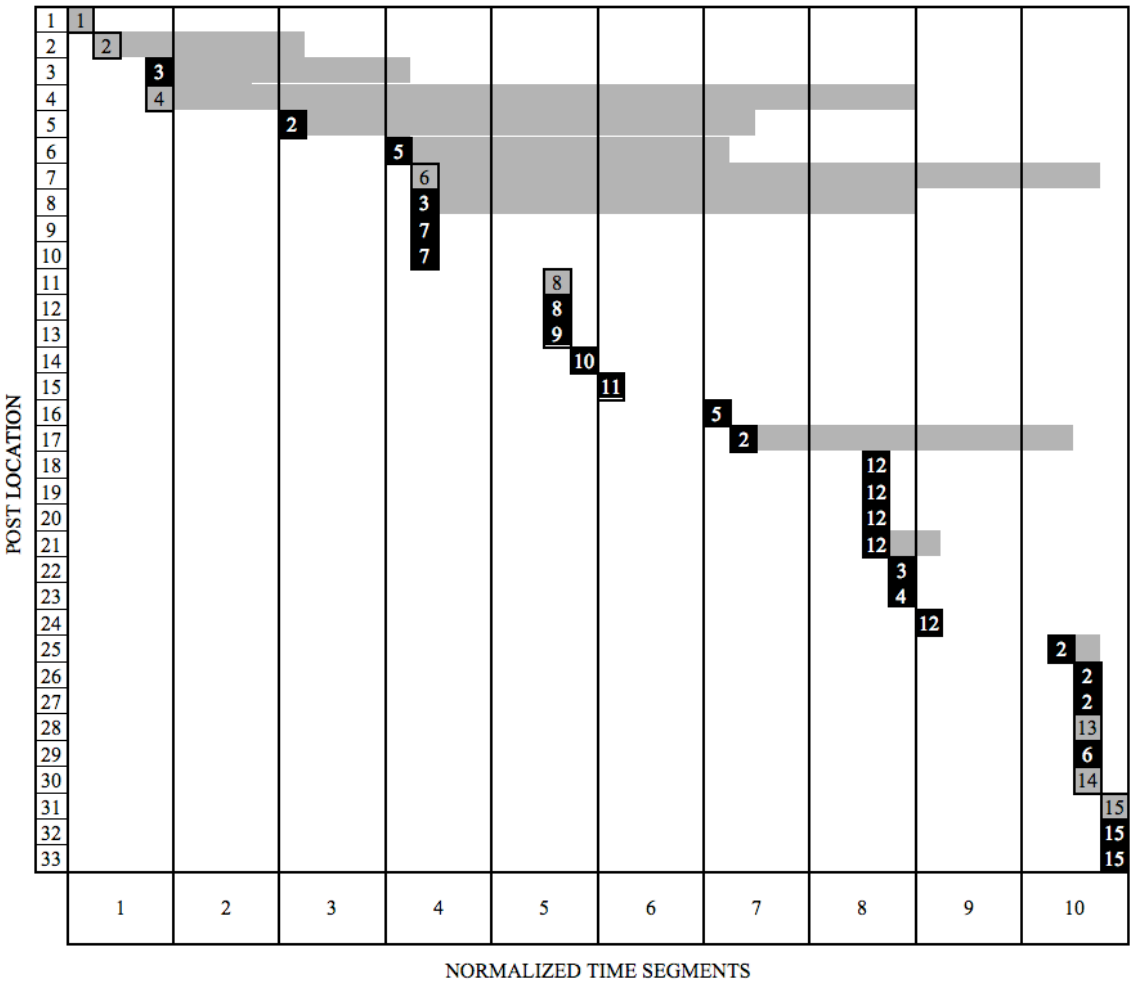

Figure D3. Post time sequence map, S3.

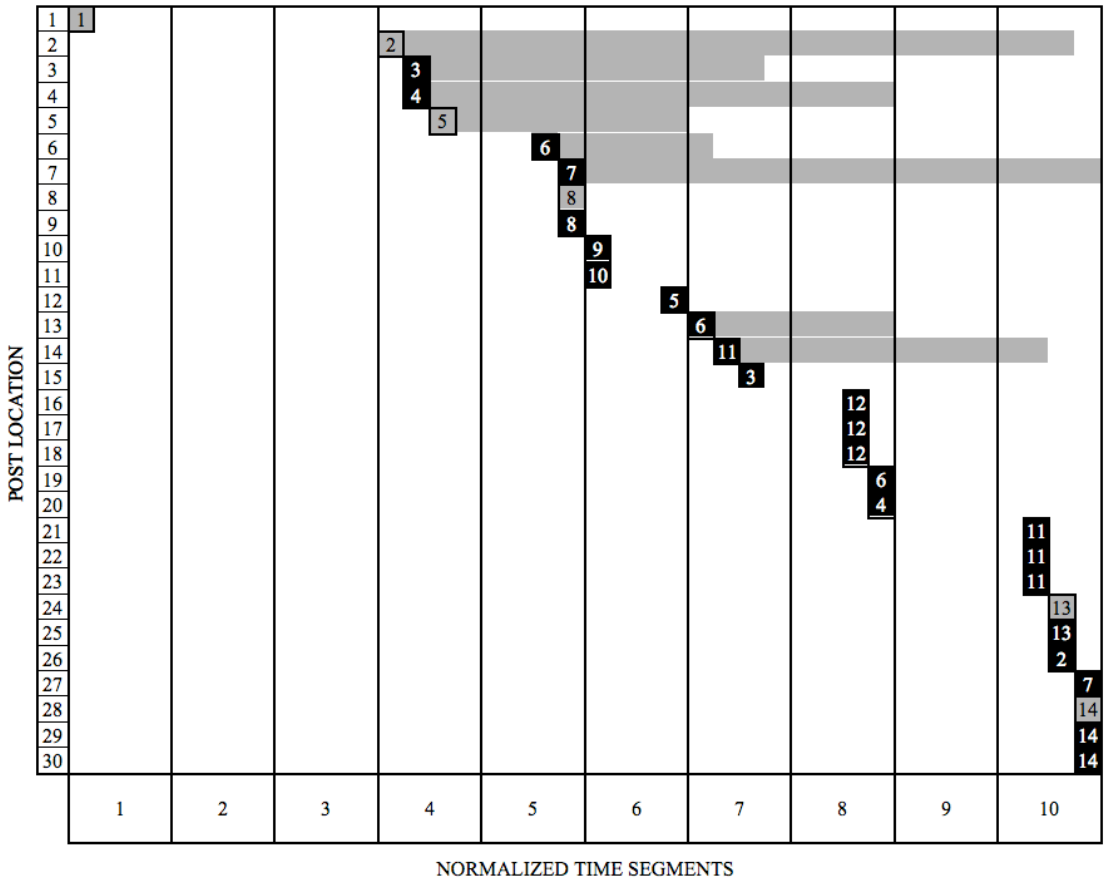

Figure D4. Post time sequence map, S4.

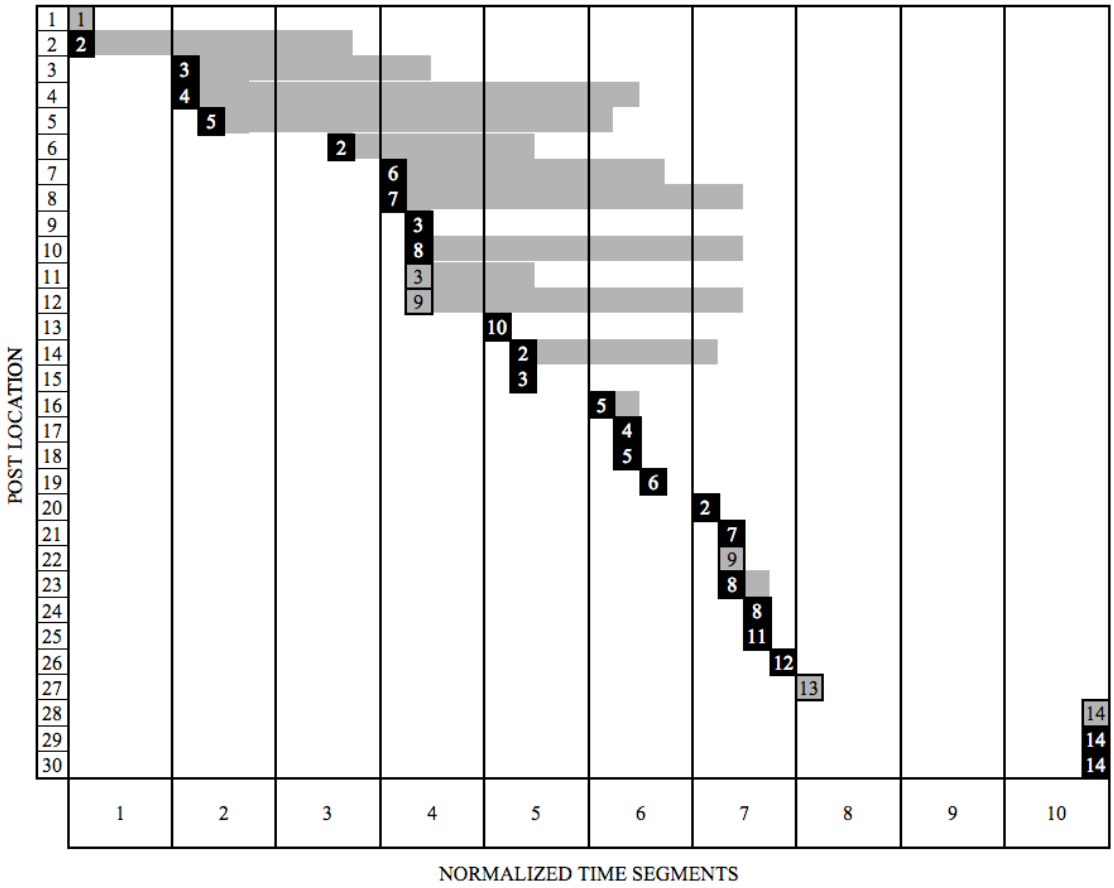

Figure D5. Post time sequence map, S5.

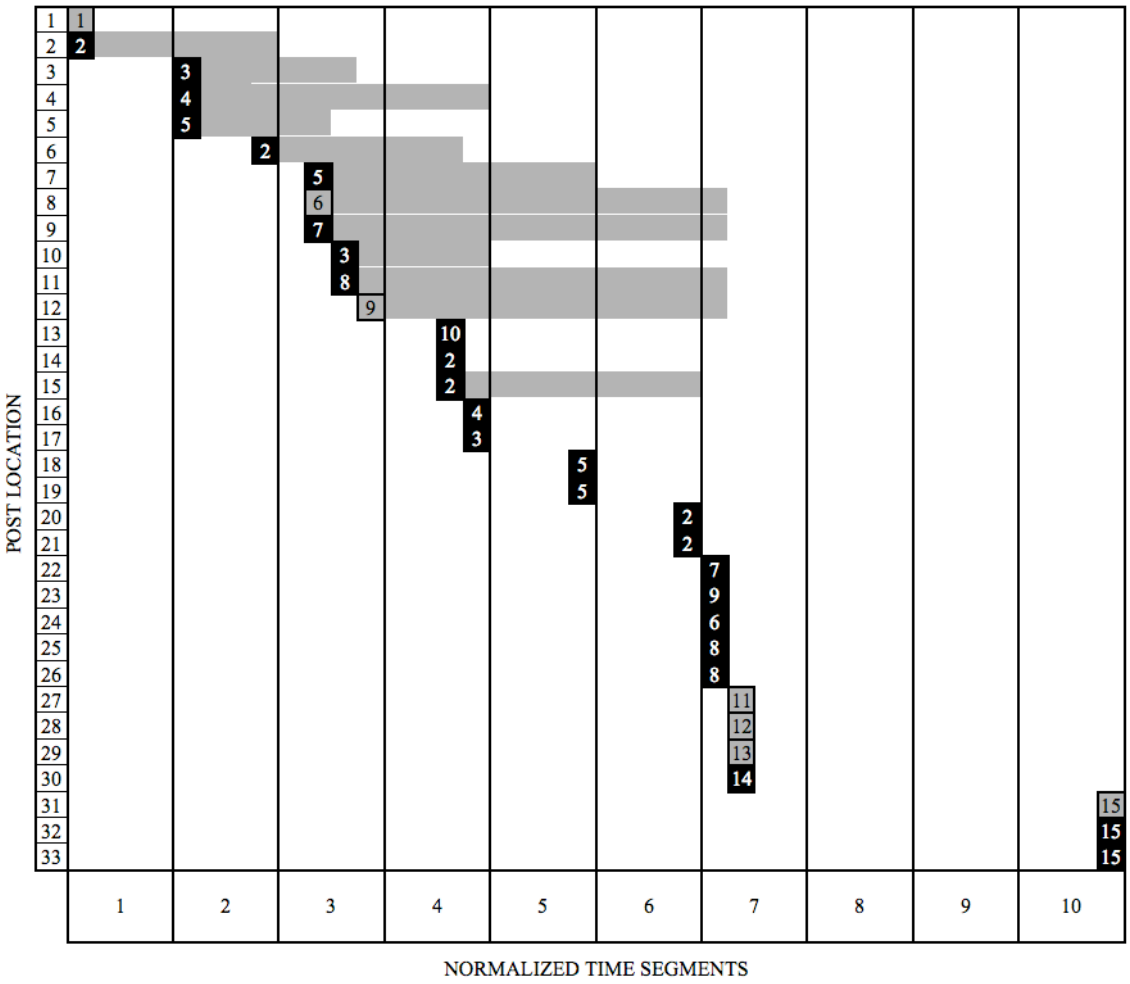

Figure D6. Post time sequence map, S6.

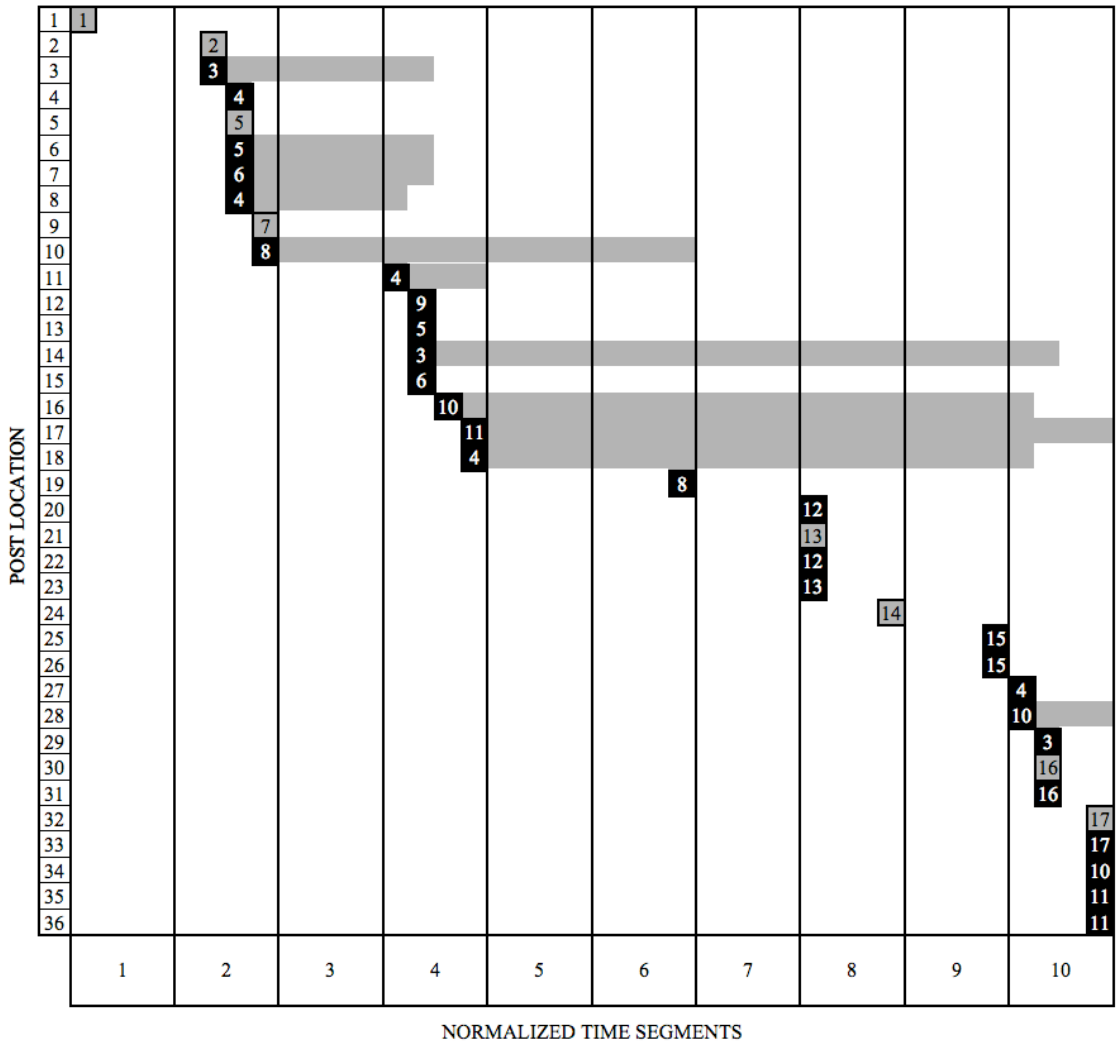

Figure D7. Post time sequence map, S7.

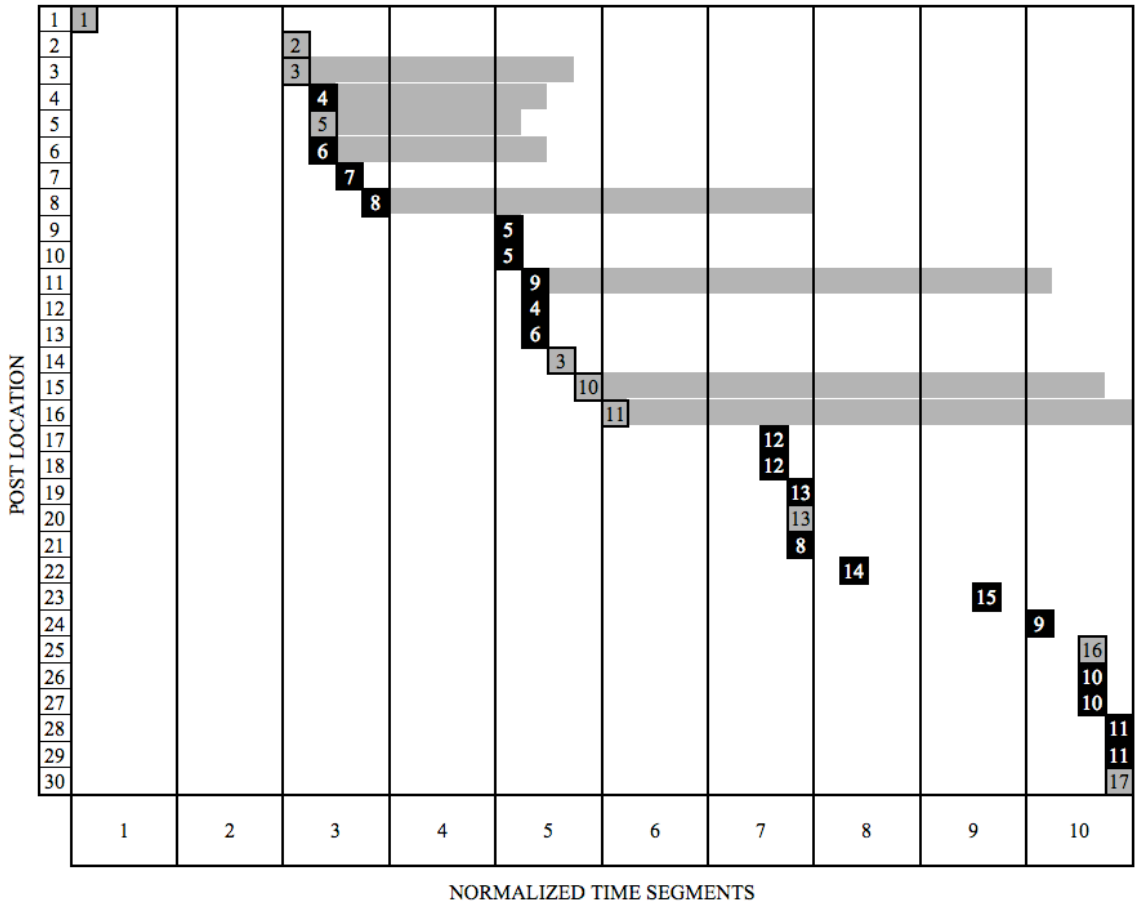

Figure D8. Post time sequence map, S8.

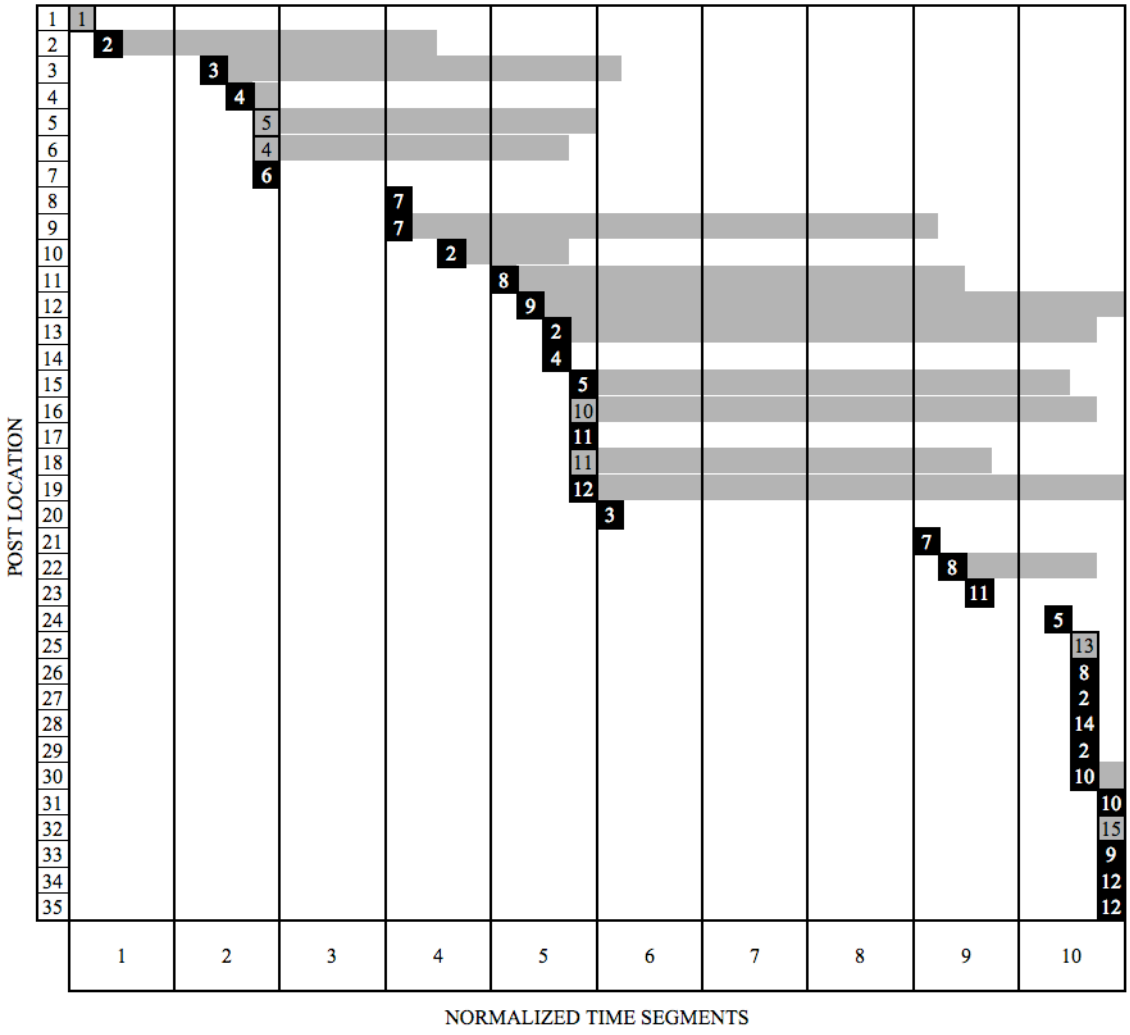

Figure D9. Post time sequence map, S9.

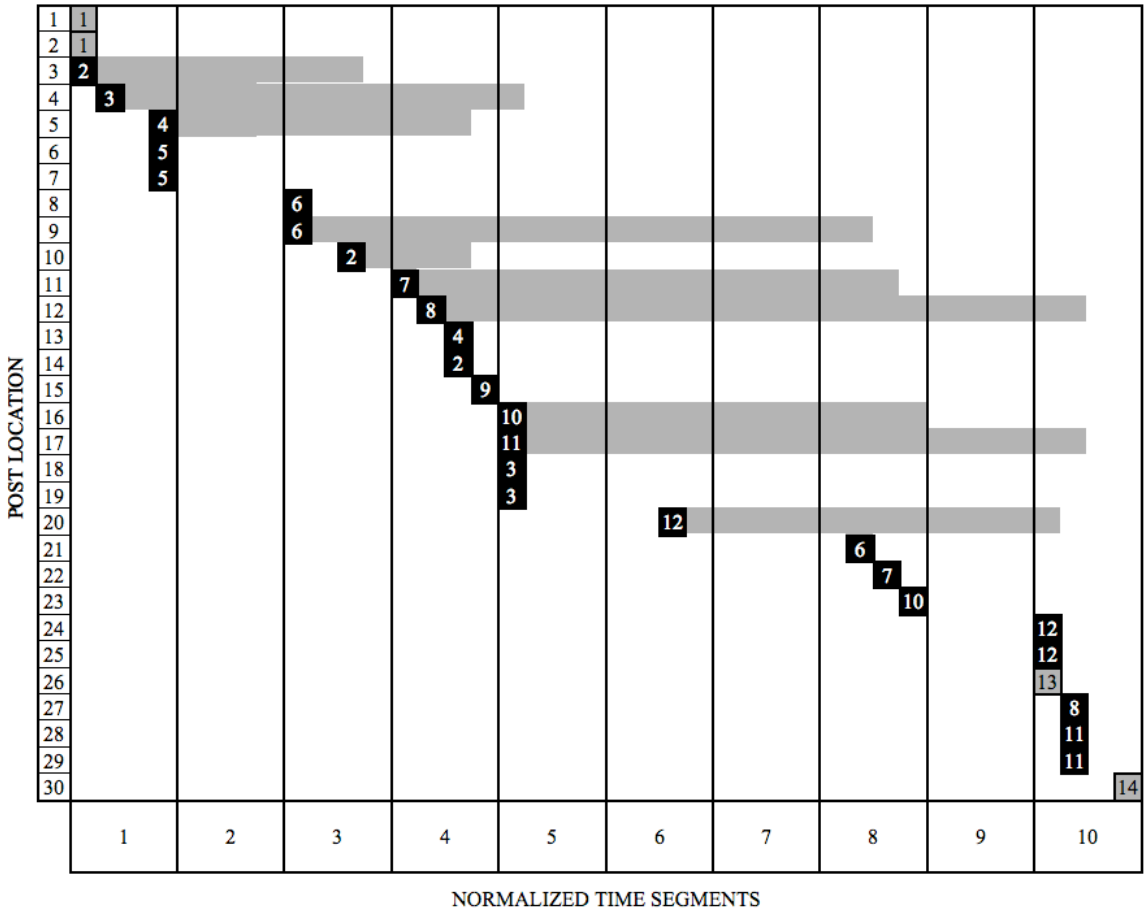

Figure D10. Post time sequence map, S10.

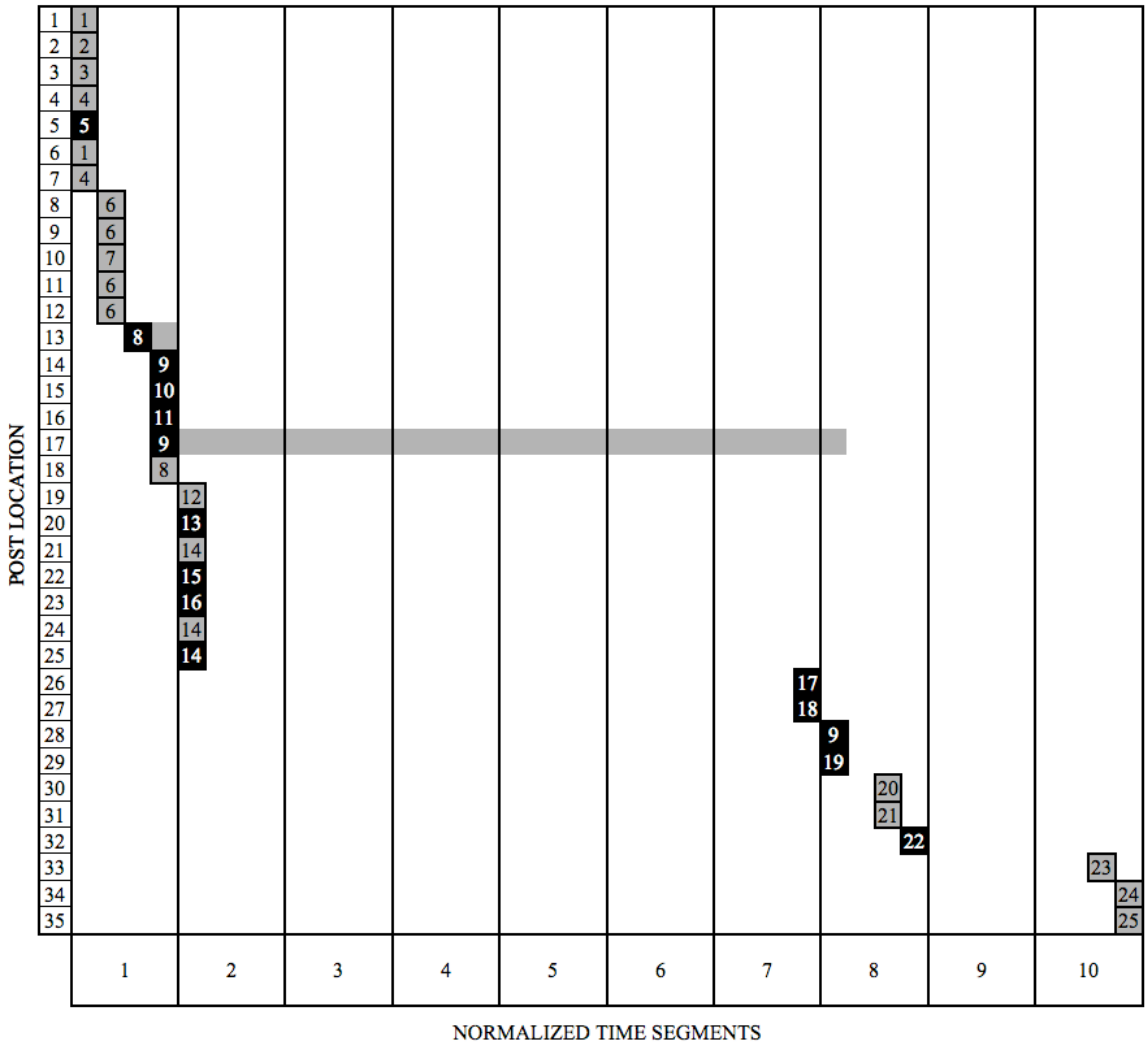

Figure D11. Post time sequence map, F1.

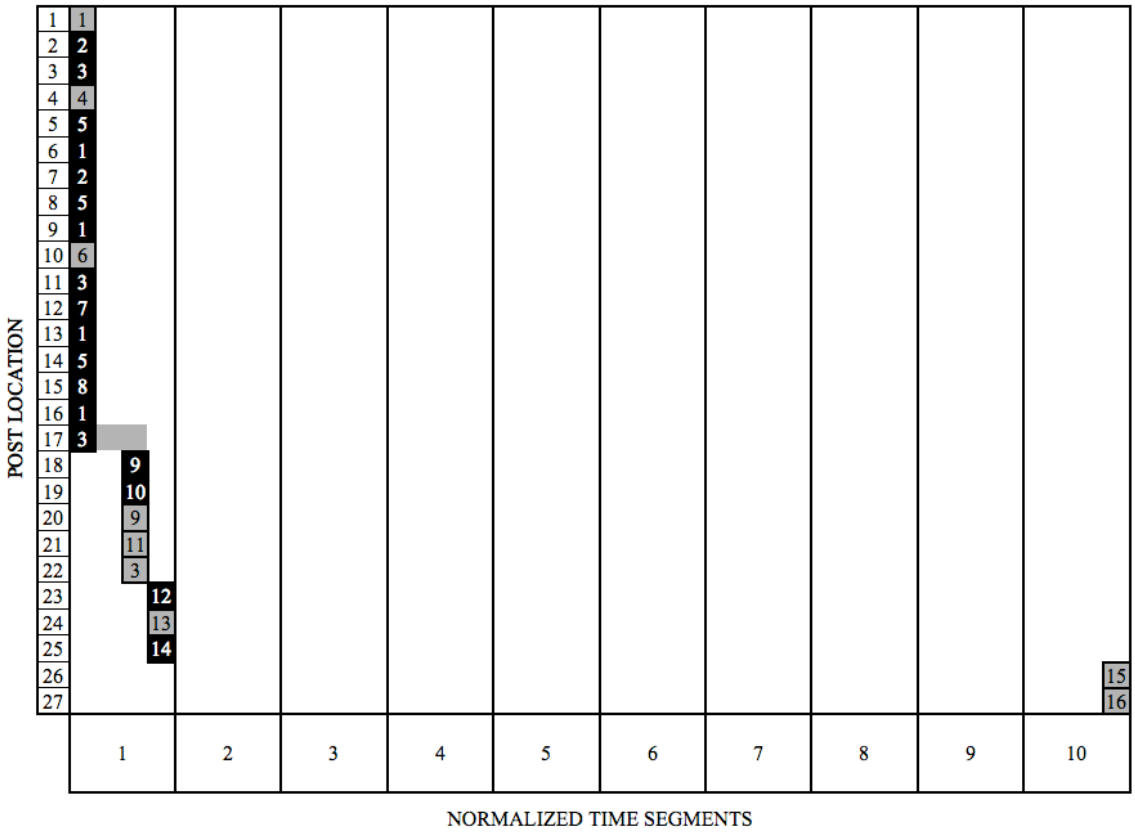

Figure D12. Post time sequence map, F2.

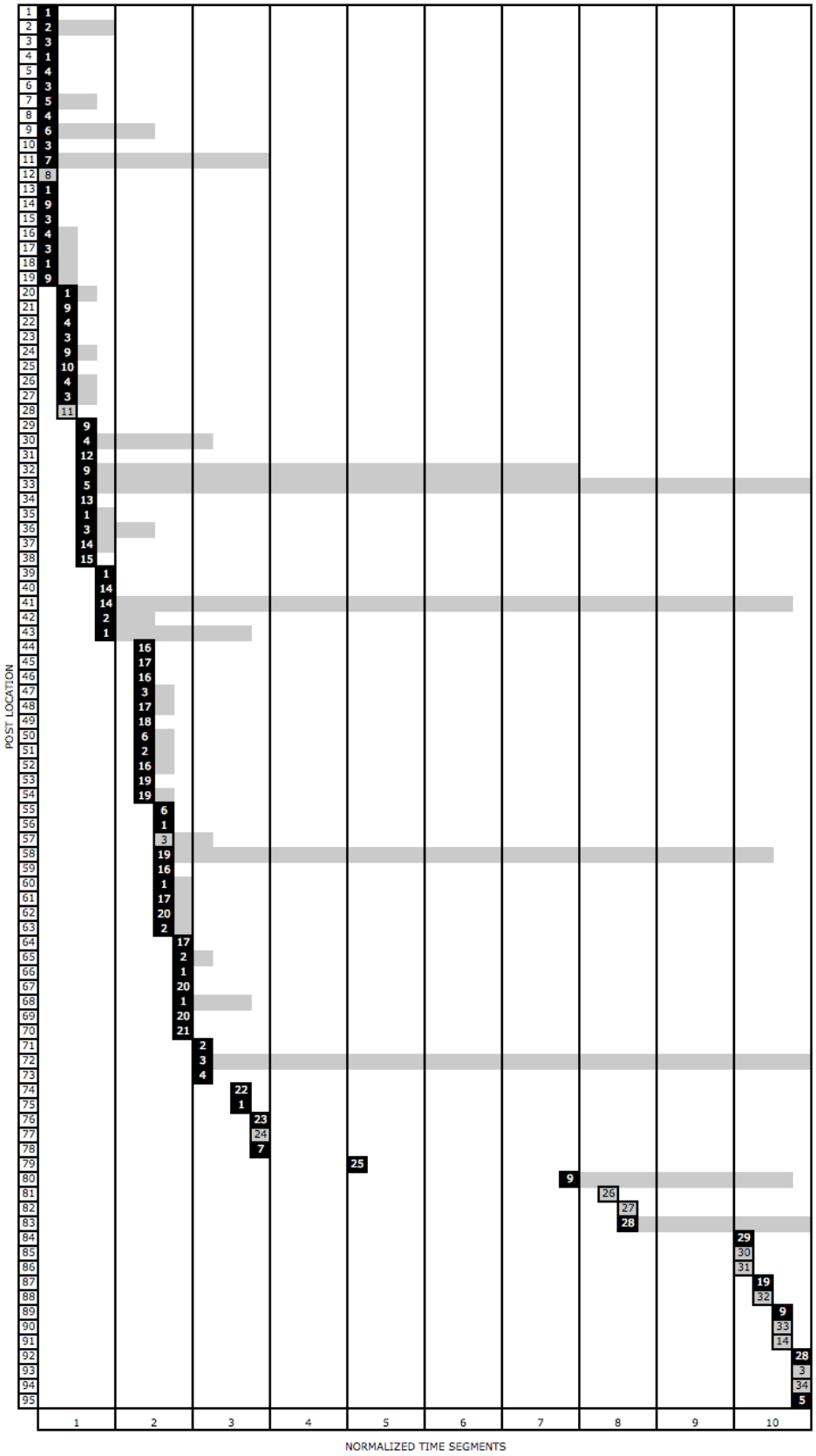

Figure D13. Post time sequence map, F3.

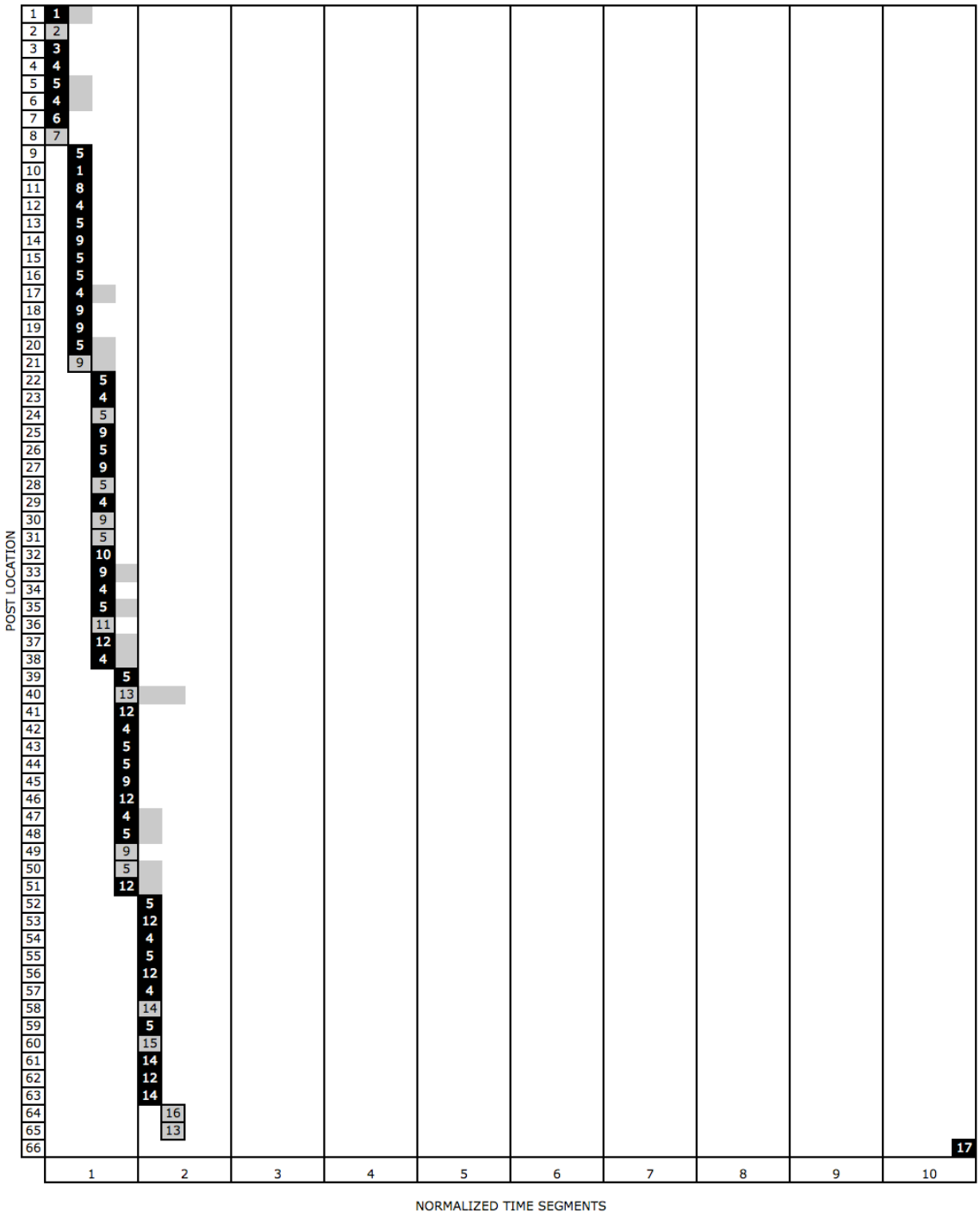

Figure D14. Post time sequence map, F4.

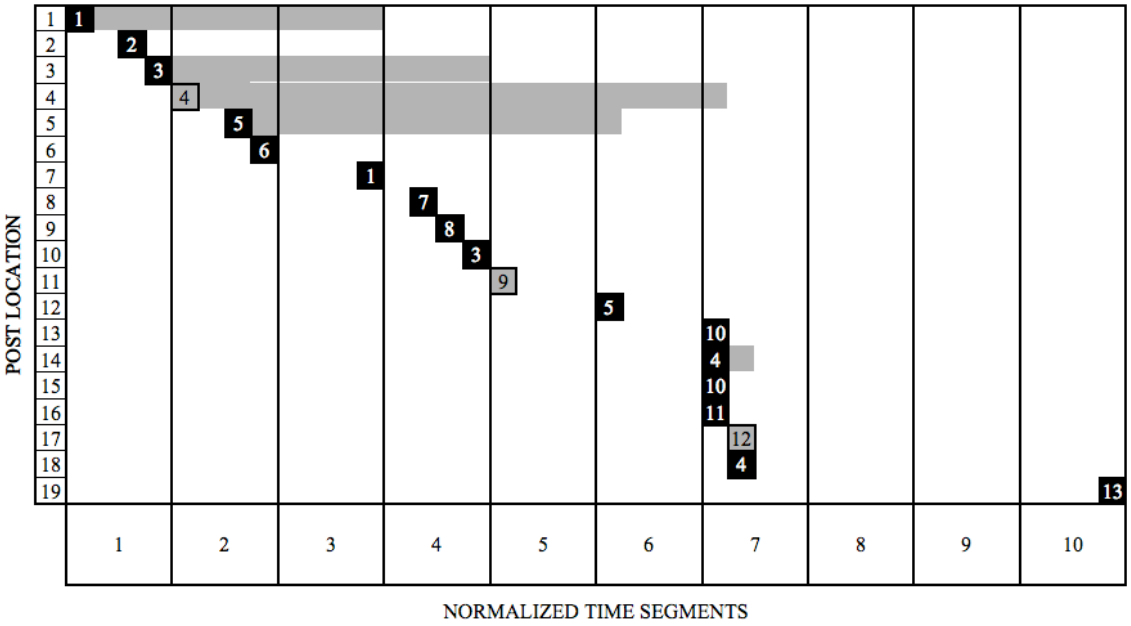

Figure D15. Post time sequence map, F5.

APPENDIX E: LINKED CHAIN MAPS, ALL POST SETS

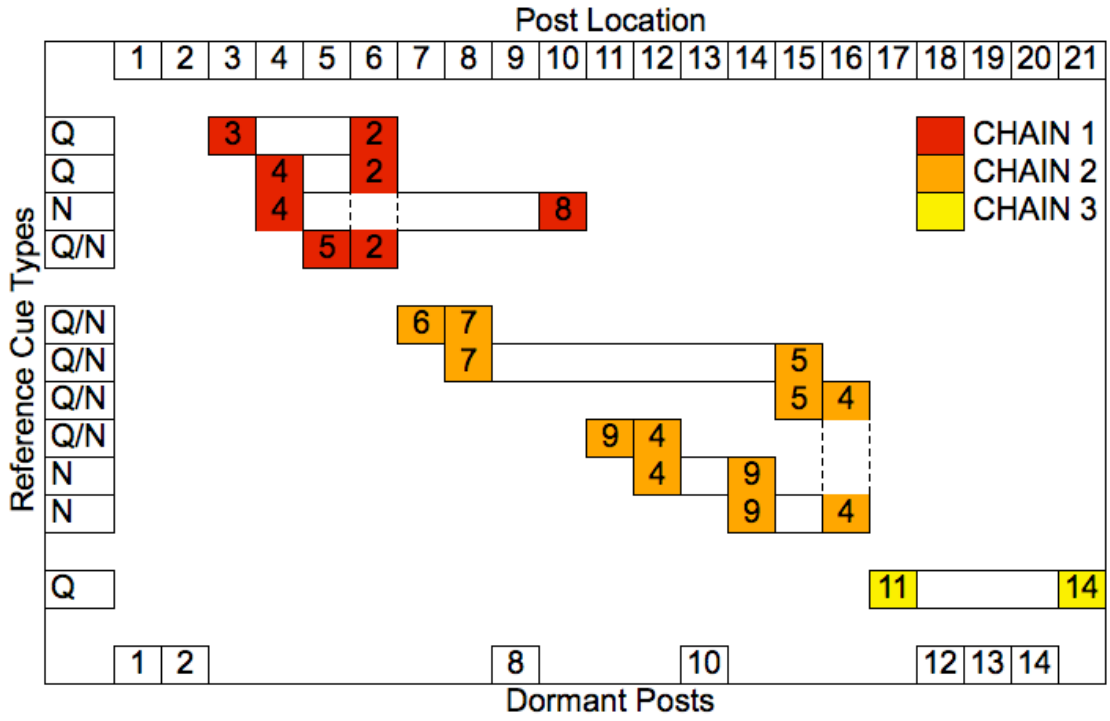

Figure E1. Linked chain map, S1.

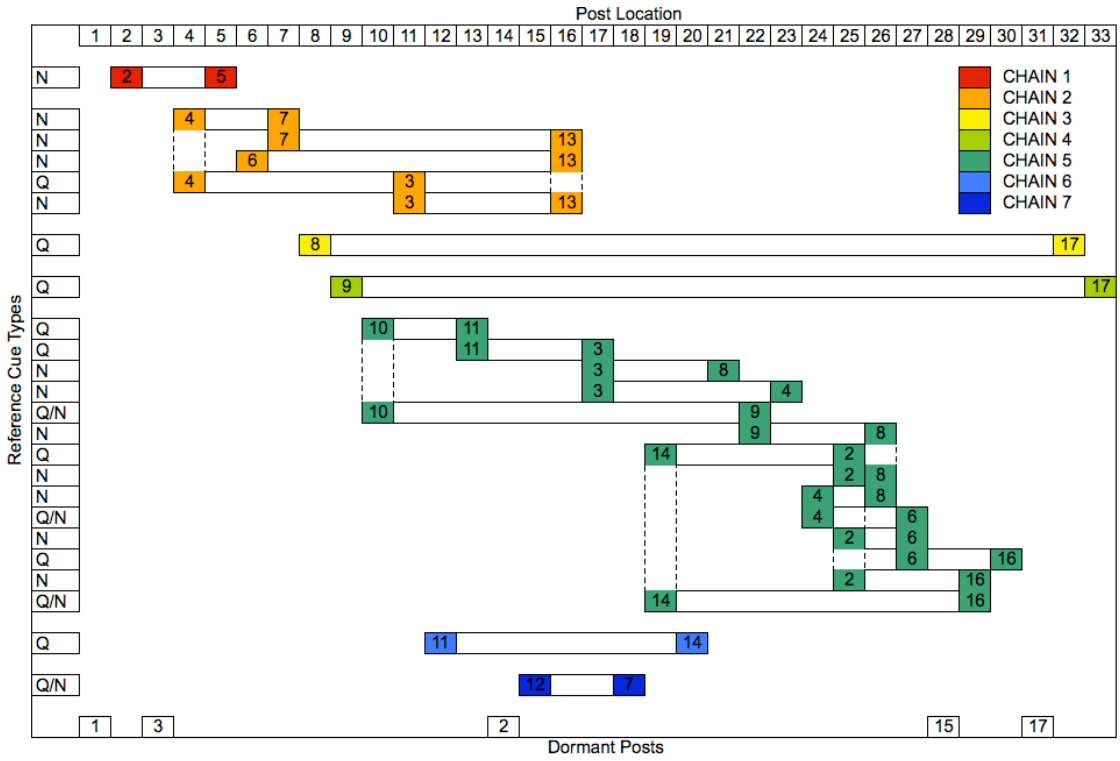

Figure E2. Linked chain map, S2.

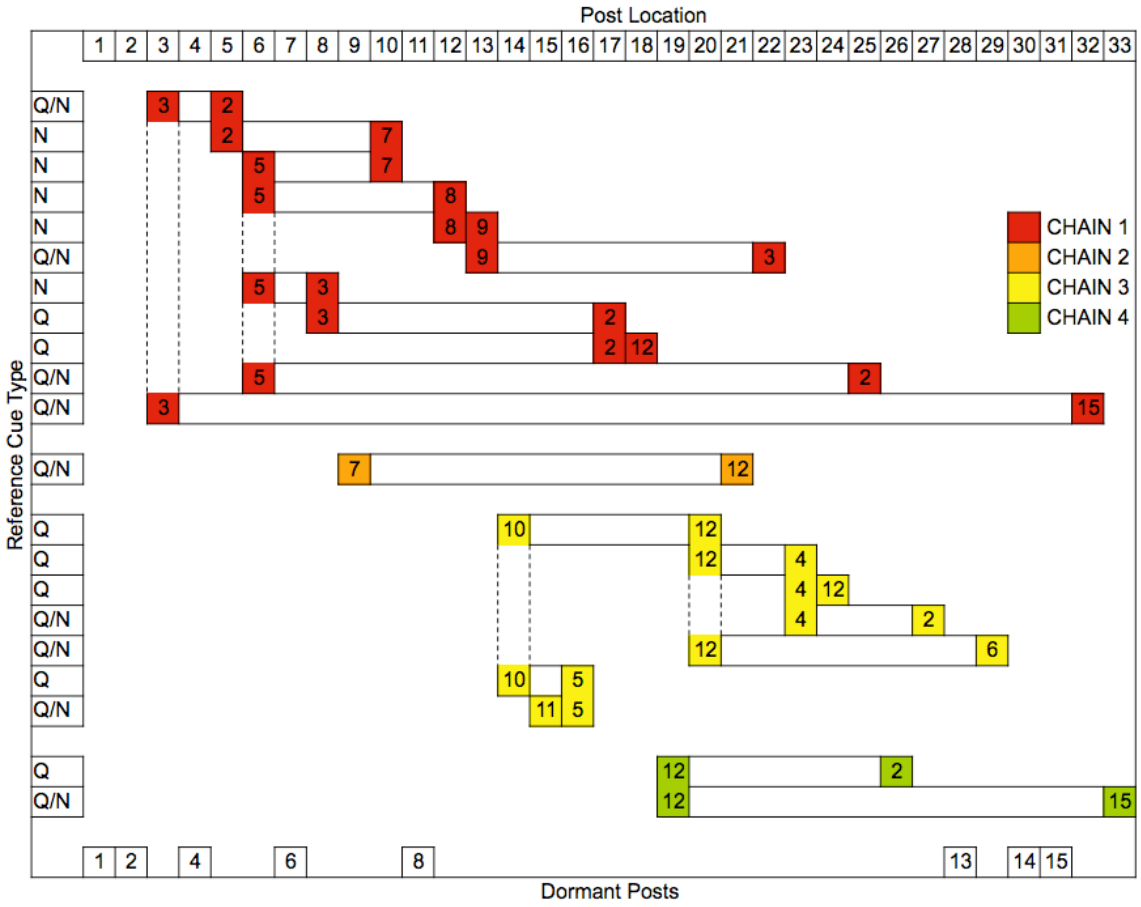

Figure E3. Linked chain map, S3.

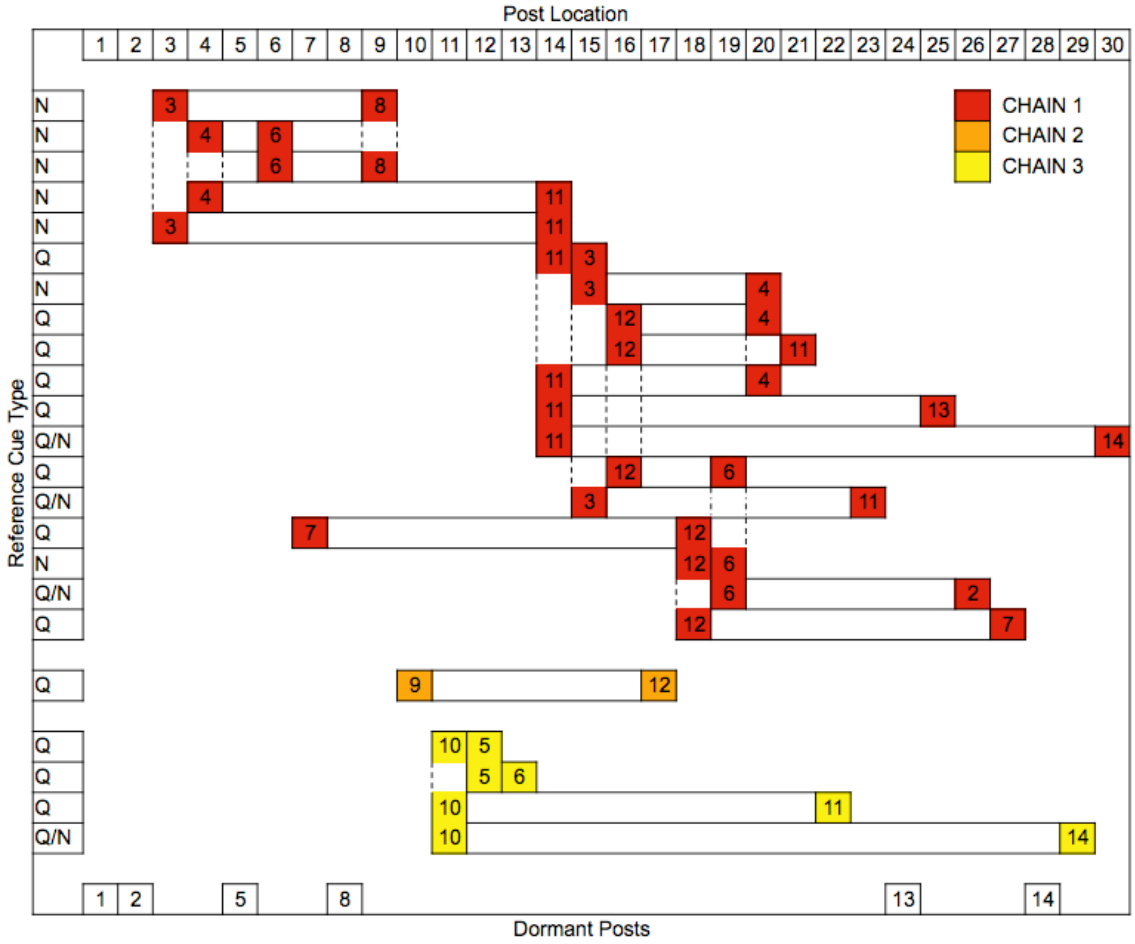

Figure E4. Linked chain map, S4.

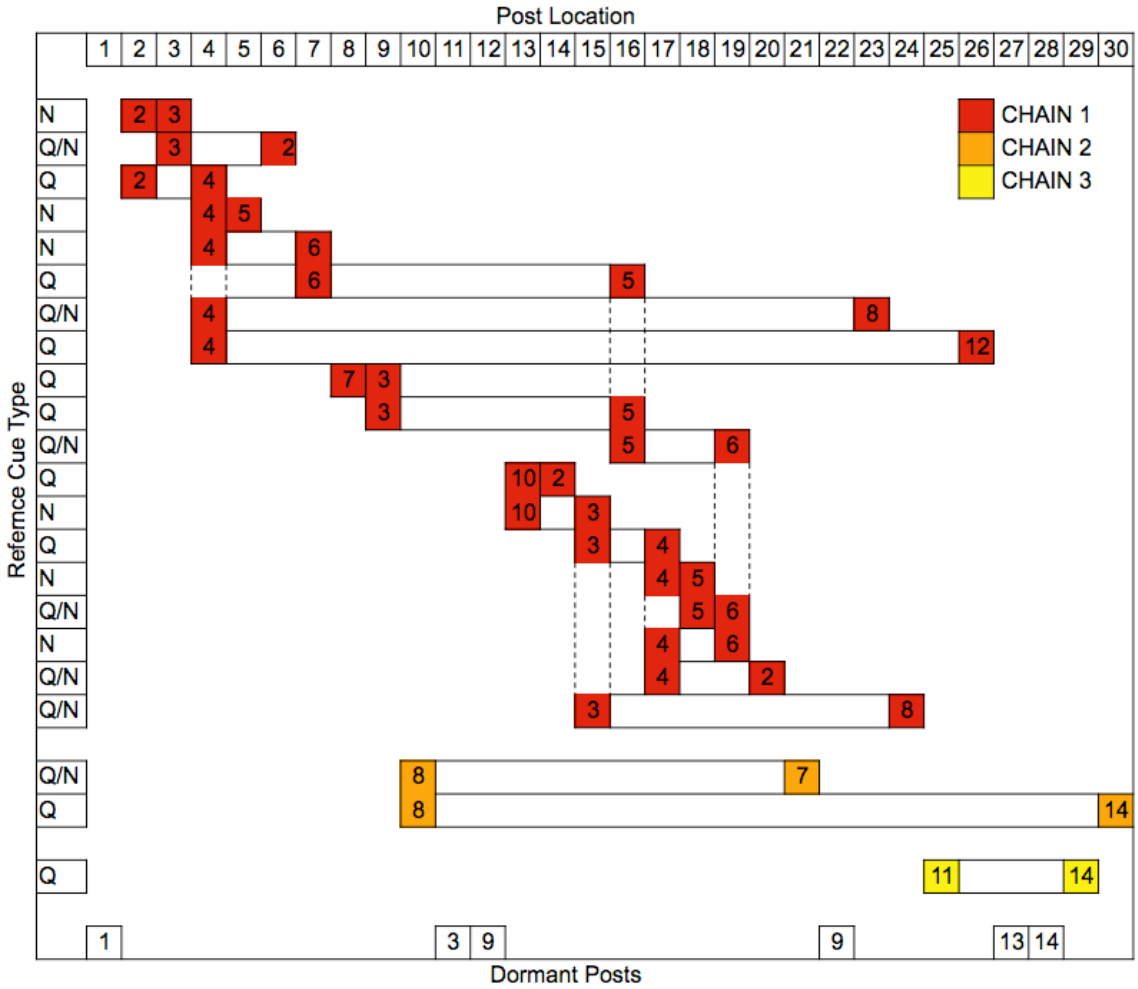

Figure E5. Linked chain map, S5.

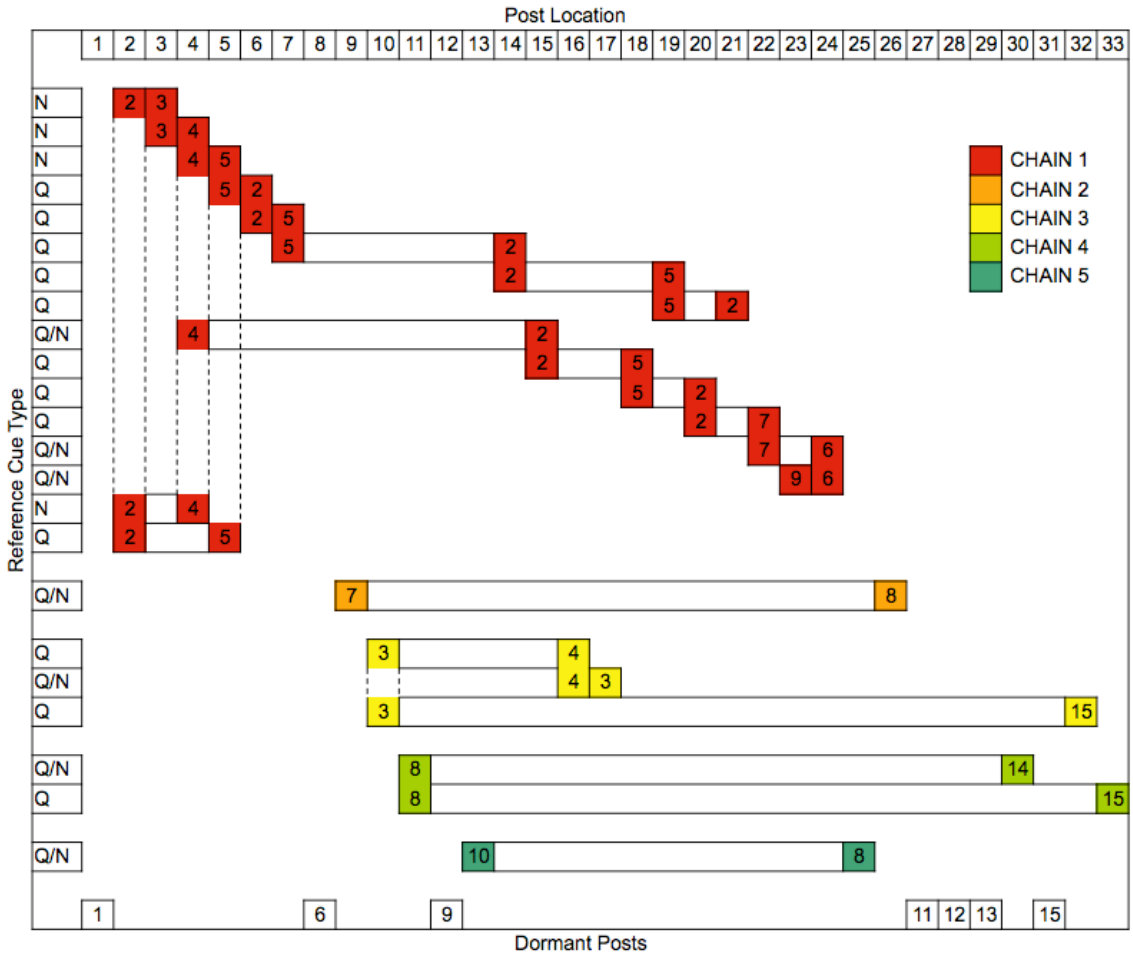

Figure E6. Linked chain map, S6.

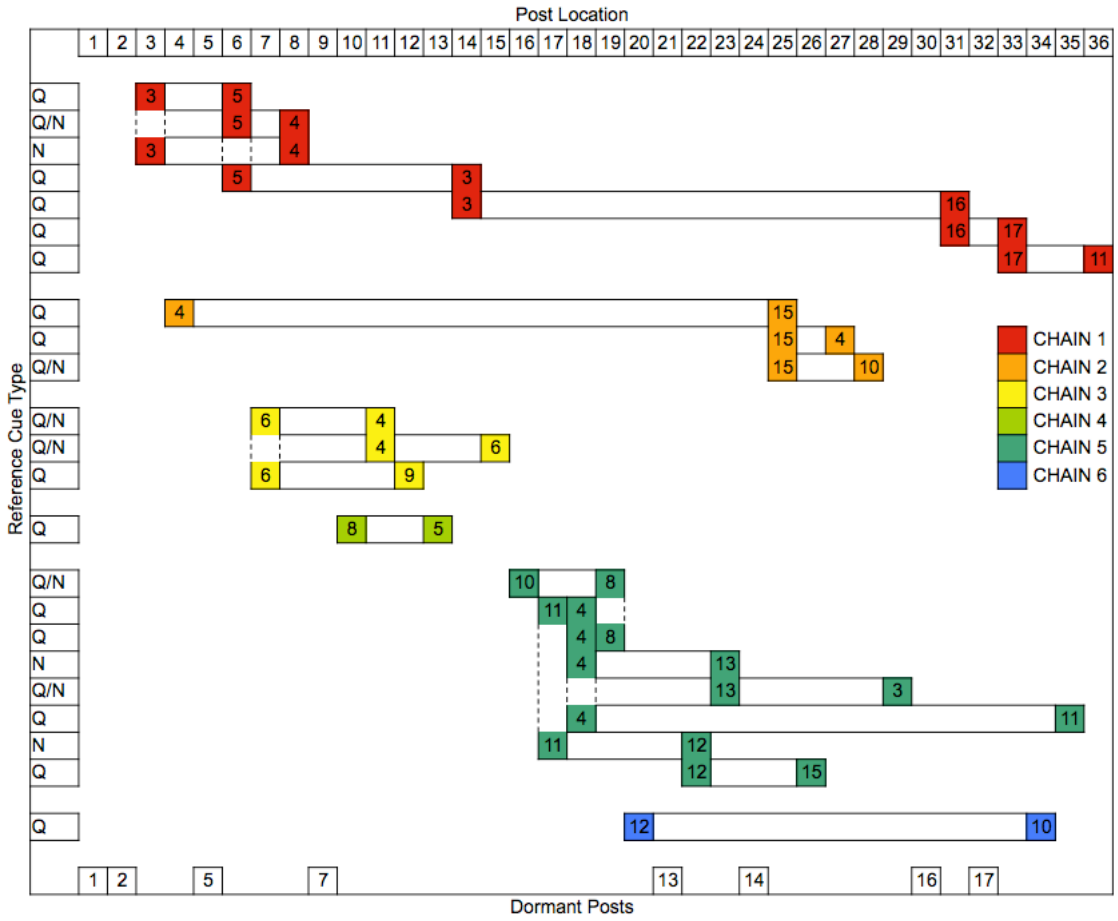

Figure E7. Linked chain map, S7.

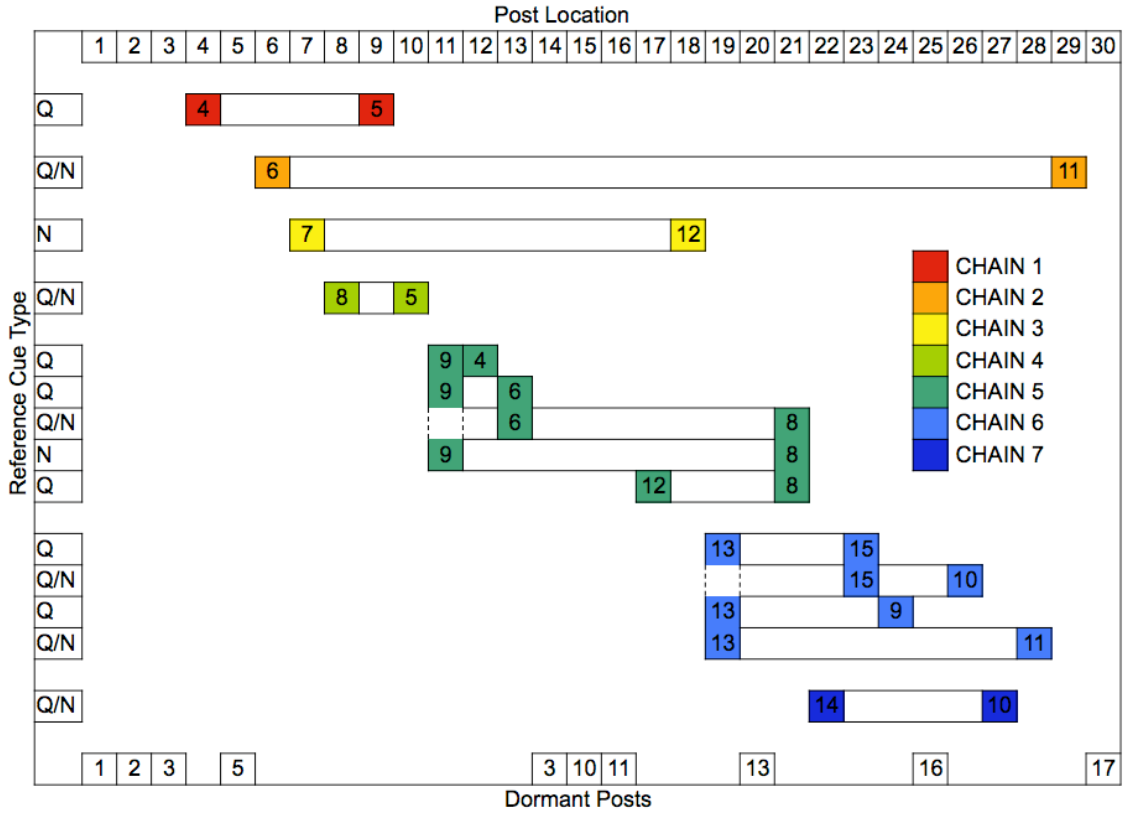

Figure E8. Linked chain map, S8.

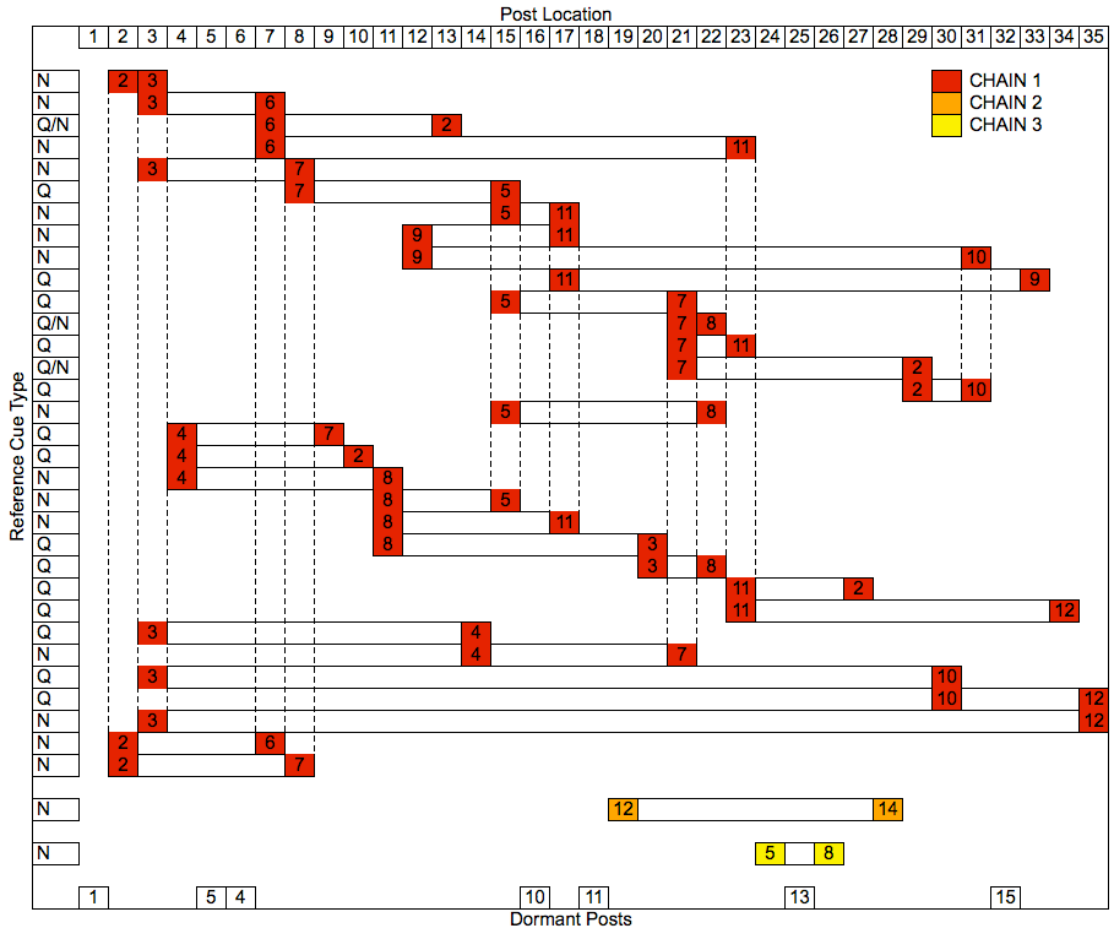

Figure E9. Linked chain map, S9.

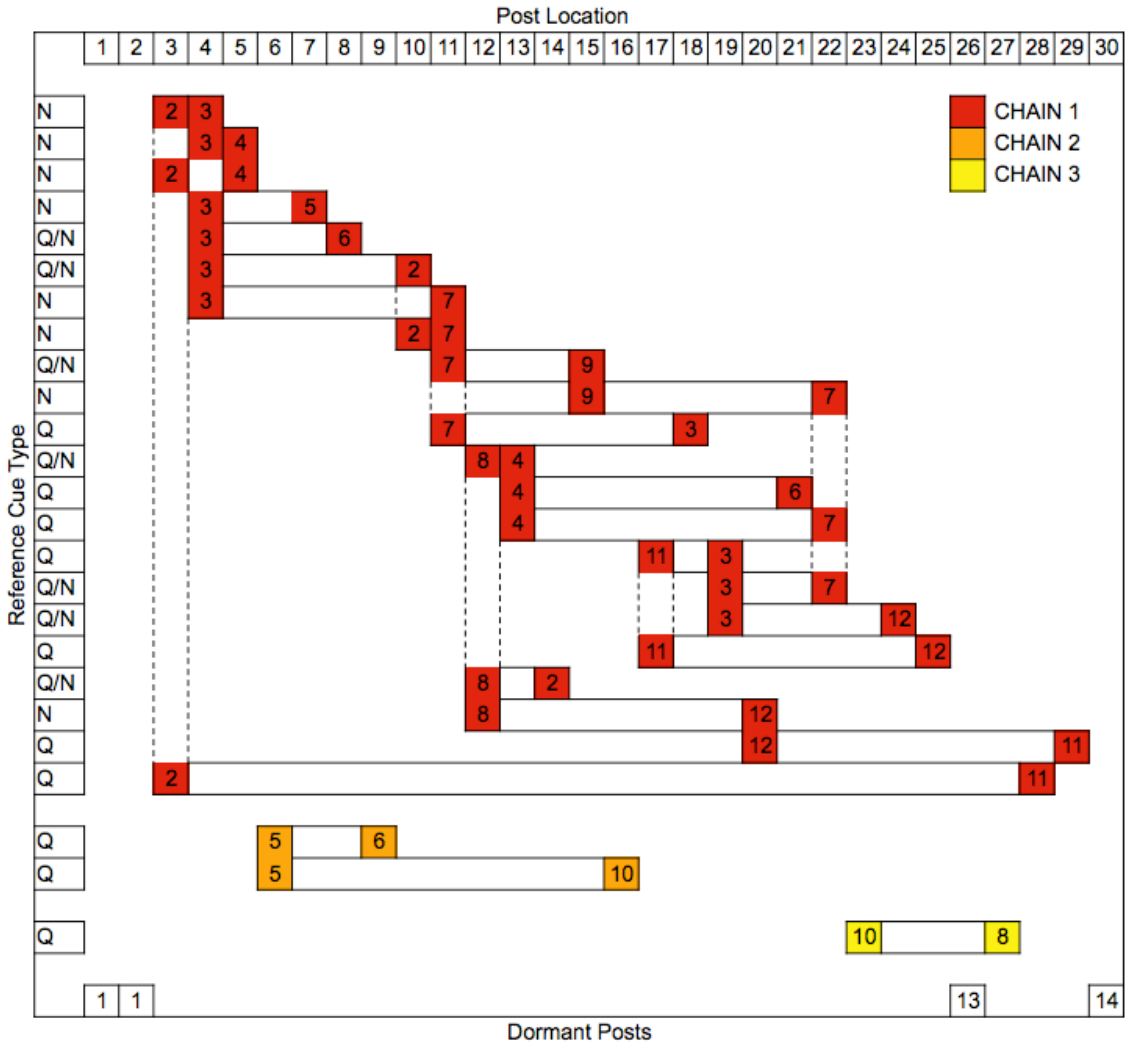

Figure E10. Linked chain map, S10.

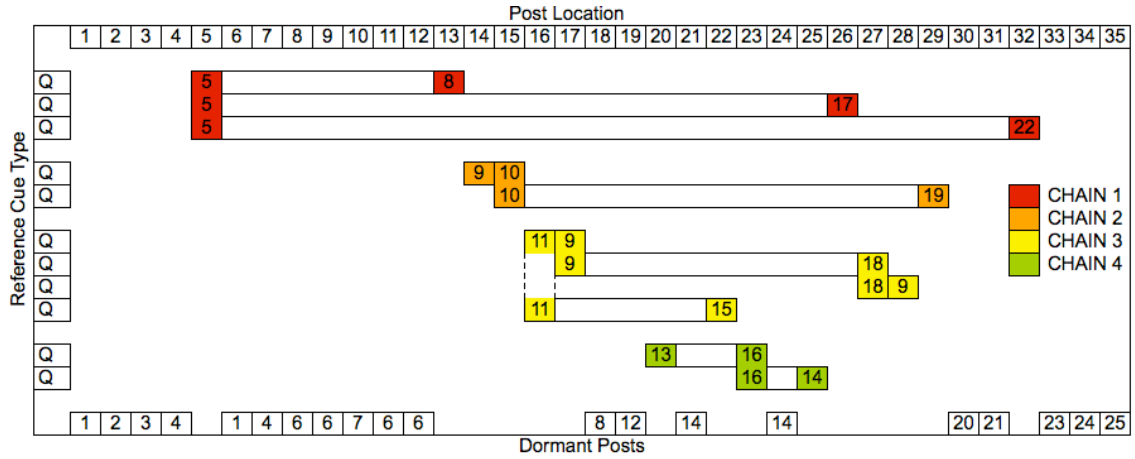

Figure E11. Linked chain map, F1.

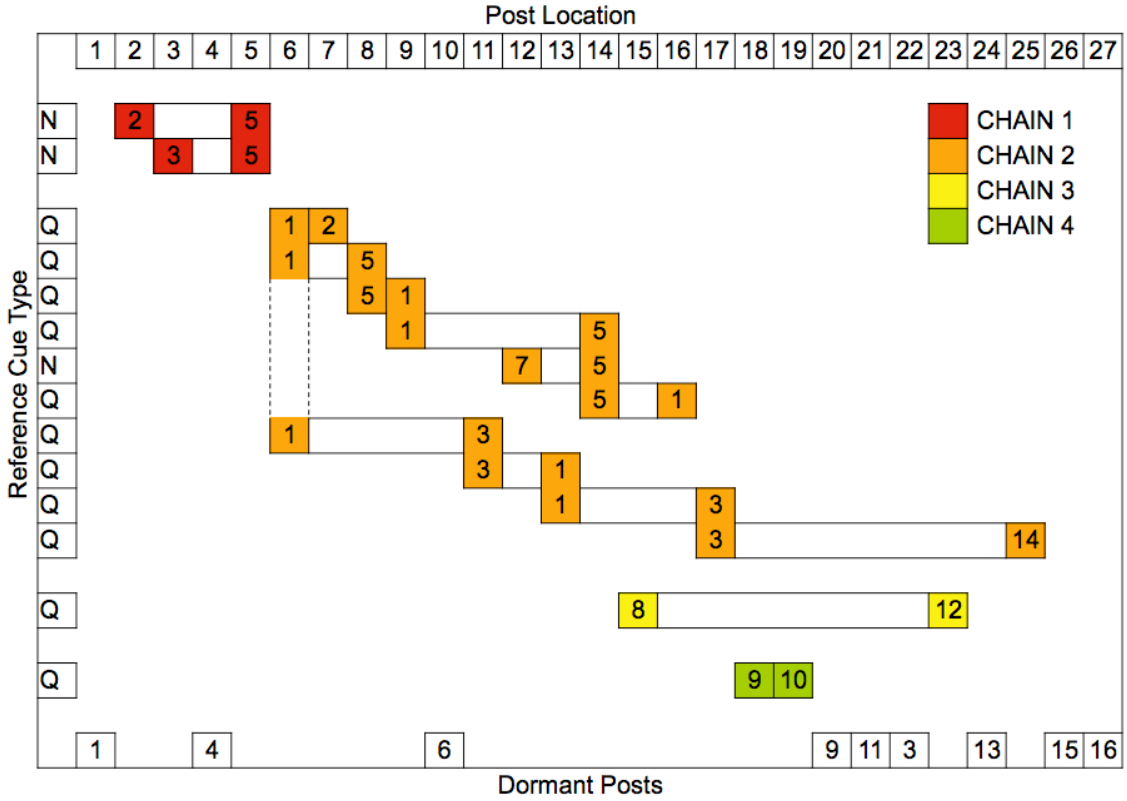

Figure E12. Linked chain map, F2.

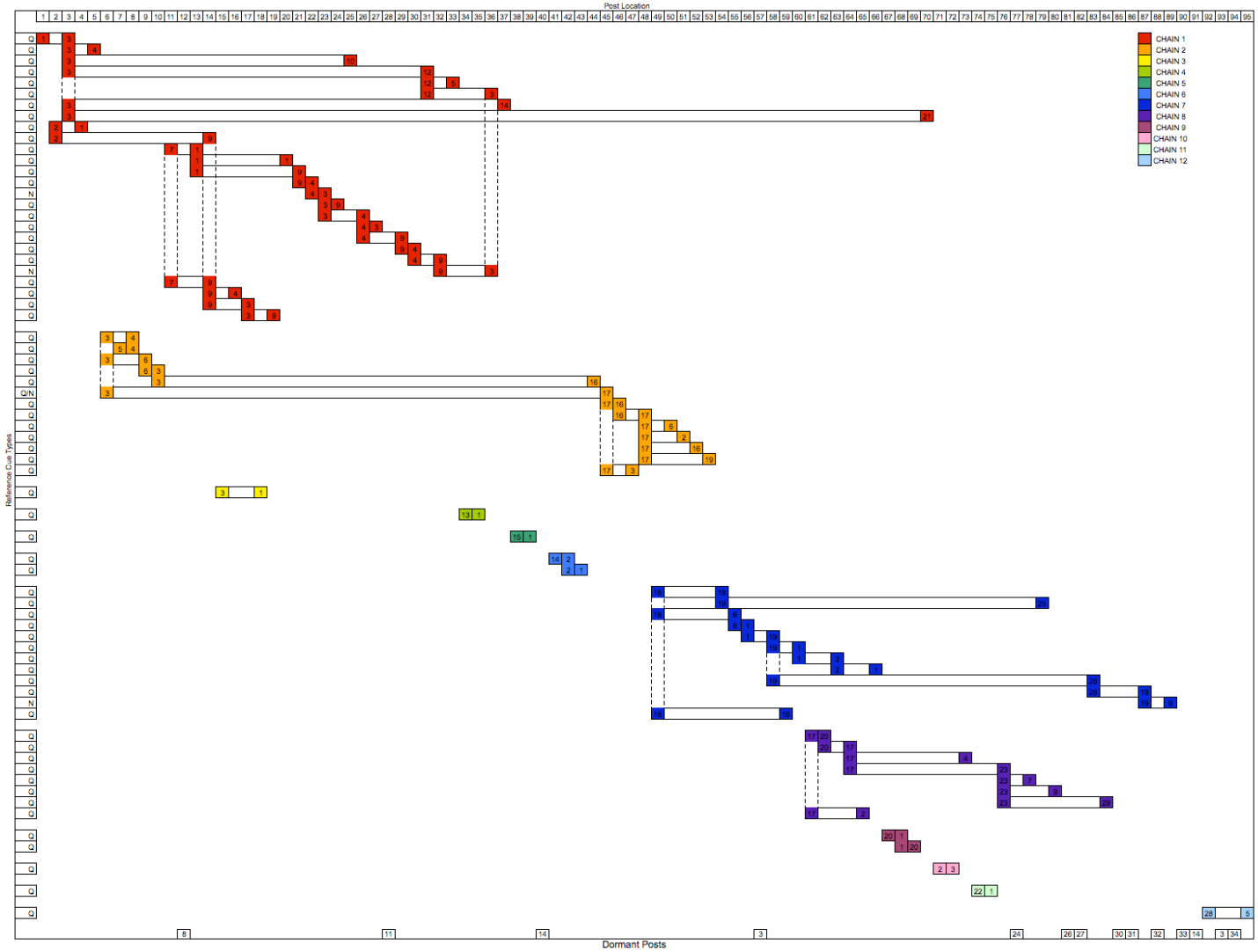

Figure E13. Linked chain map, F3.

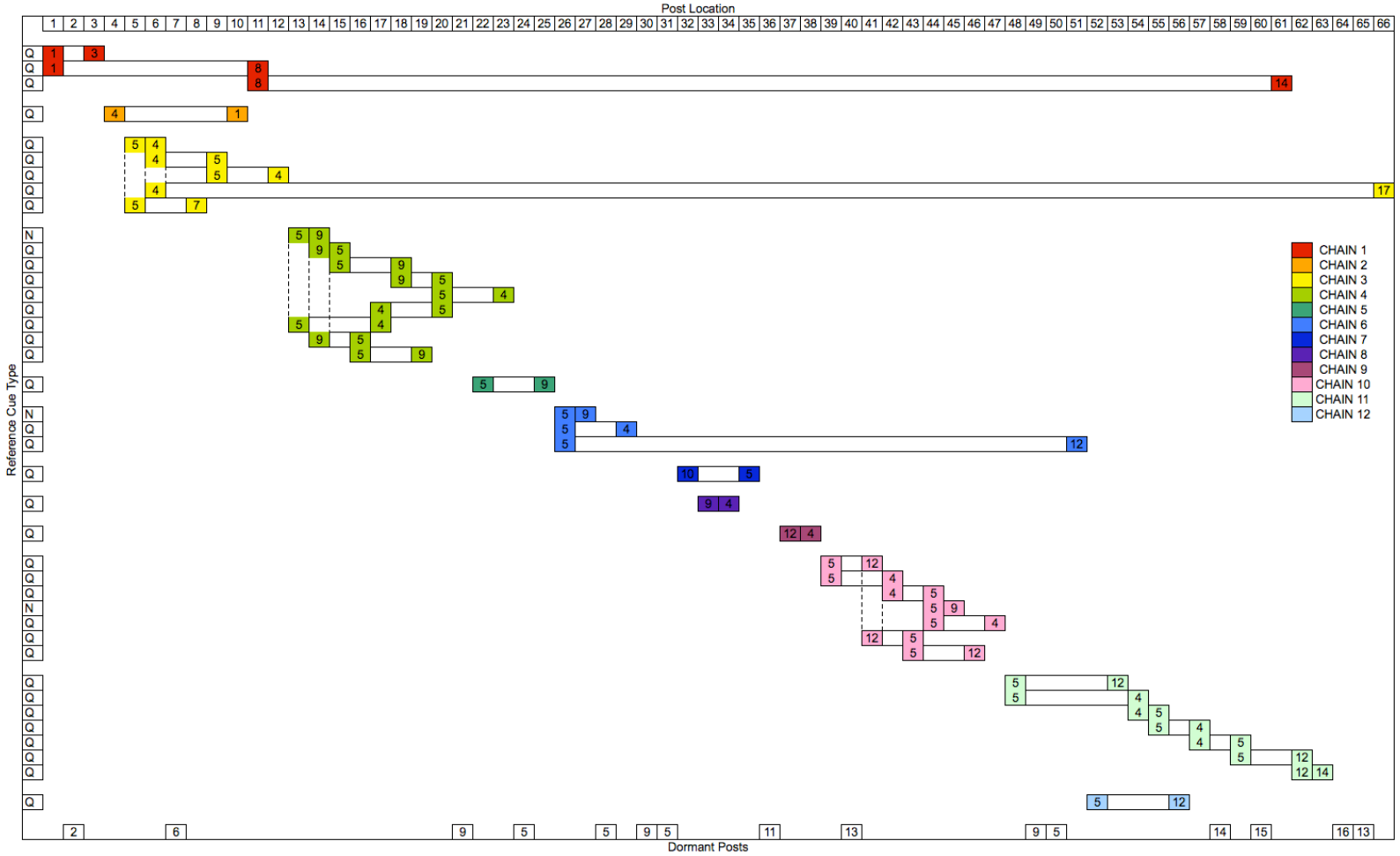

Figure E14. Linked chain map, F4.

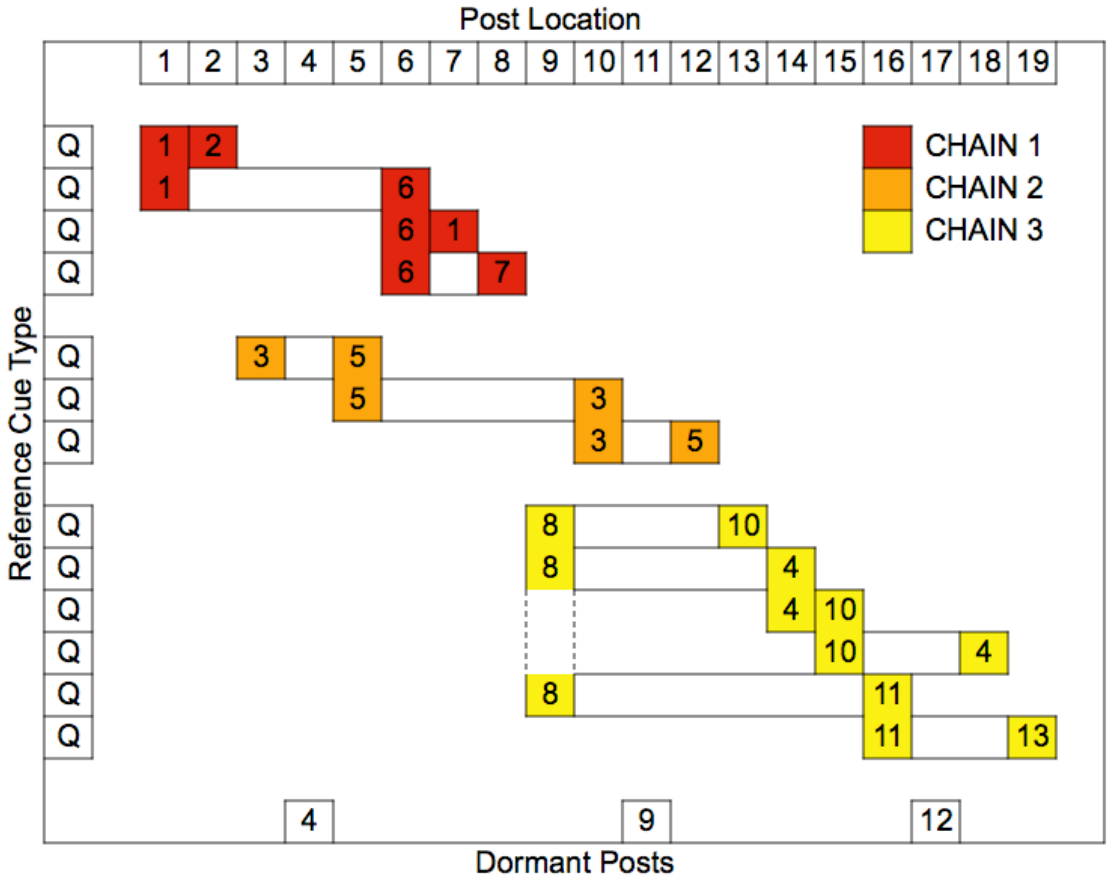

Figure E15. Linked chain map, F5.

APPENDIX F: CHAIN CHARACTERISTICS, ALL POST SETS

Table F1

Characteristics of Chains in Student Data Post Sets

Post set	Chain	# posts	# of post authors	PA values of post authors
S1	CH1	5	5	2, 3, 4, 5, 8
	CH2	7	5	4, 5, 6, 7, 9
	CH3	2	2	11, 14
S2	CH1	2	2	2, 5
	CH2	5	5	3, 4, 6, 7, 13
	CH3	2	2	8, 17
	CH4	2	2	9, 17
	CH5	13	10	2, 3, 4, 6, 8, 9, 10, 11, 14, 16
	CH6	2	2	11, 14
	CH7	2	2	7, 12
S3	CH1	12	8	2, 3, 5, 7, 8, 9, 12, 15
	CH2	2	2	7, 12
	CH3	8	6	2, 4, 5, 10, 11, 12
	CH4	3	3	2, 12, 15
S4	CH1	17	9	2, 3, 4, 6, 8, 11, 12, 13, 14
	CH2	2	2	9, 12
	CH3	5	5	5, 6, 10, 11, 14
S5	CH1	19	9	2, 3, 4, 5, 6, 7, 8, 10, 12
	CH2	3	3	7, 8, 14
	CH3	2	2	11, 14
S6	CH1	15	7	2, 3, 4, 5, 6, 7, 9
	CH2	2	2	7, 8
	CH3	4	3	3, 4, 15
	CH4	3	3	8, 14, 15
	CH5	2	2	8, 10
S7	CH1	7	6	3, 4, 5, 11, 16, 17
	CH2	4	3	4, 10, 15
	CH3	4	3	4, 6, 9
	CH4	2	2	5, 8
	CH5	9	8	3, 4, 8, 10, 11, 12, 13, 15
	CH6	2	2	10, 12
S8	CH1	2	2	4, 5
	CH2	2	2	6, 11
	CH3	2	2	7, 12
	CH4	2	2	5, 8
	CH5	5	5	4, 6, 8, 9, 12
	CH6	5	5	9, 10, 11, 13, 15
	CH7	2	2	10, 14
S9	CH1	24	11	2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12
	CH2	2	2	12, 14
	CH3	2	2	5, 8
S10	CH1	19	10	2, 3, 4, 5, 6, 7, 8, 9, 11, 12
	CH2	3	3	5, 6, 10
	CH3	2	2	8, 10

Table F2

Characteristics of Chains in Fan Data Post Sets

Post set	Chain	# posts	# of post authors	PA values of post authors
F1	CH1	4	4	5, 8, 17, 22
	CH2	3	3	9, 10, 19
	CH3	5	4	9, 11, 15, 18
	CH4	3	3	13, 14, 16
F2	CH1	3	3	2, 3, 5
	CH2	11	6	1, 2, 3, 5, 7, 14
	CH3	2	2	8, 12
	CH4	2	2	9, 10
F3	CH1	27	10	1, 2, 3, 4, 5, 9, 10, 12, 14, 21
	CH2	14	7	2, 4, 5, 6, 16, 17, 19
	CH3	2	1	3
	CH4	2	2	1, 13
	CH5	2	2	1, 15
	CH6	3	3	1, 2, 14
	CH7	13	9	1, 2, 6, 9, 16, 19, 24, 25, 28
	CH8	9	8	2, 4, 7, 9, 17, 20, 23, 29
	CH9	3	2	1, 20
	CH10	2	2	2, 3
	CH11	2	2	1, 22
	CH12	2	2	5, 28
F4	CH1	4	4	1, 3, 8, 14
	CH2	2	2	1, 4
	CH3	6	4	4, 5, 7, 17
	CH4	9	3	4, 5, 9
	CH5	2	2	5, 9
	CH6	4	4	4, 5, 9, 12
	CH7	2	2	5, 10
	CH8	2	2	4, 9
	CH9	2	2	4, 12
	CH10	8	4	4, 5, 9, 12
	CH11	8	4	4, 5, 12, 14
	CH12	2	2	5, 12
F5	CH1	5	4	1, 2, 6, 7
	CH2	4	2	3, 5
	CH3	7	5	4, 8, 10, 11, 13

APPENDIX G: Sequential lag maps, all post sets

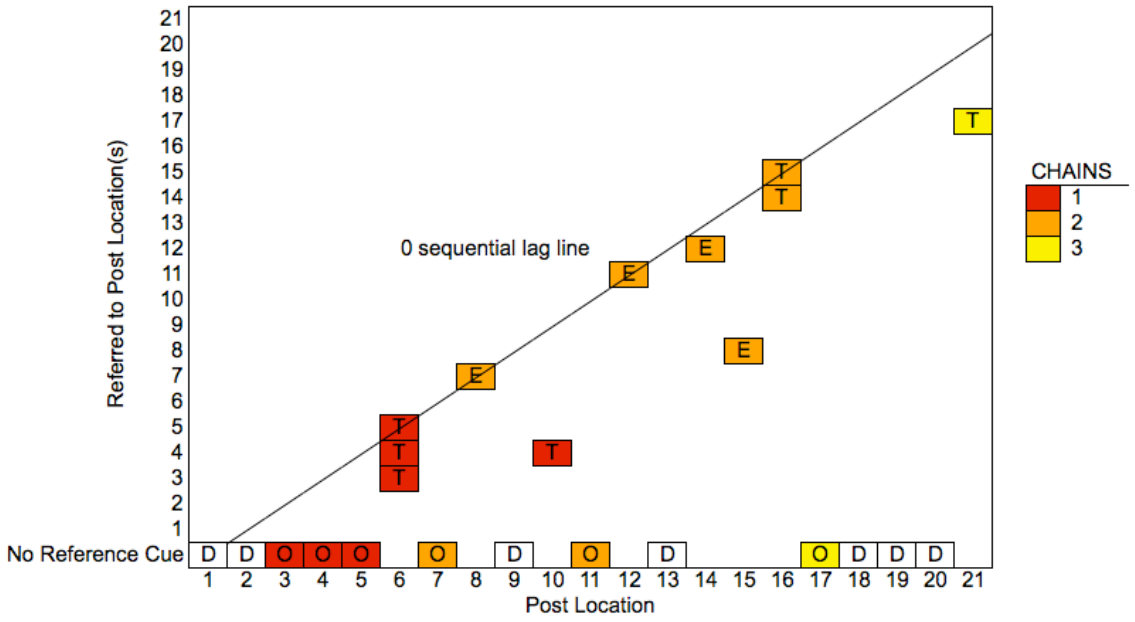

Figure G1. Sequential lag map, S1.

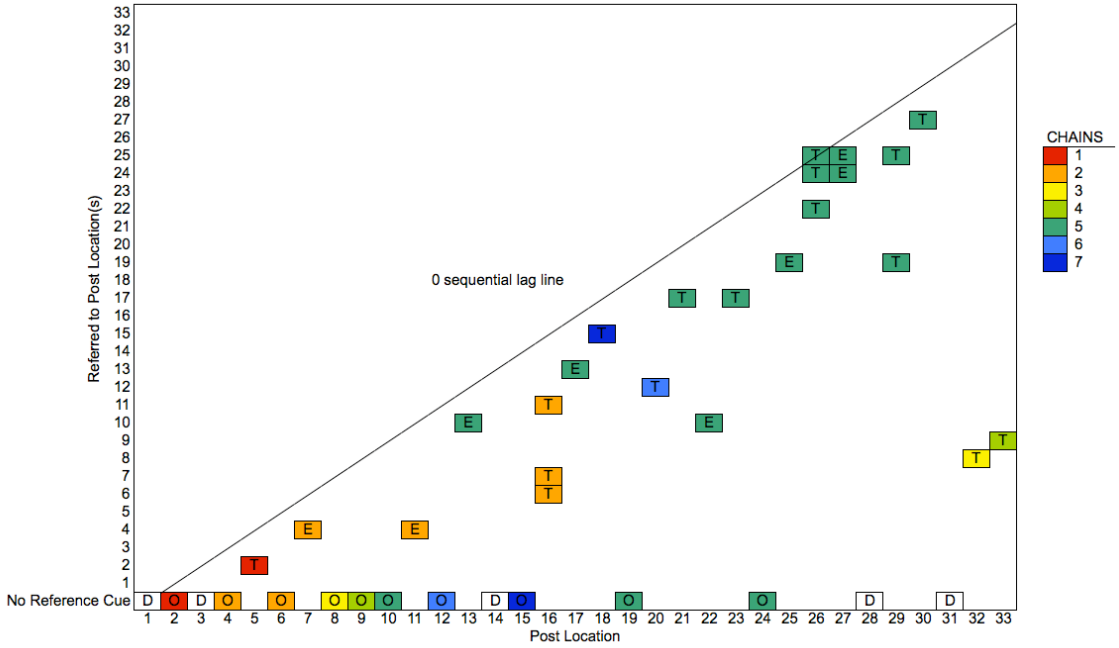

Figure G2. Sequential lag map, S2.

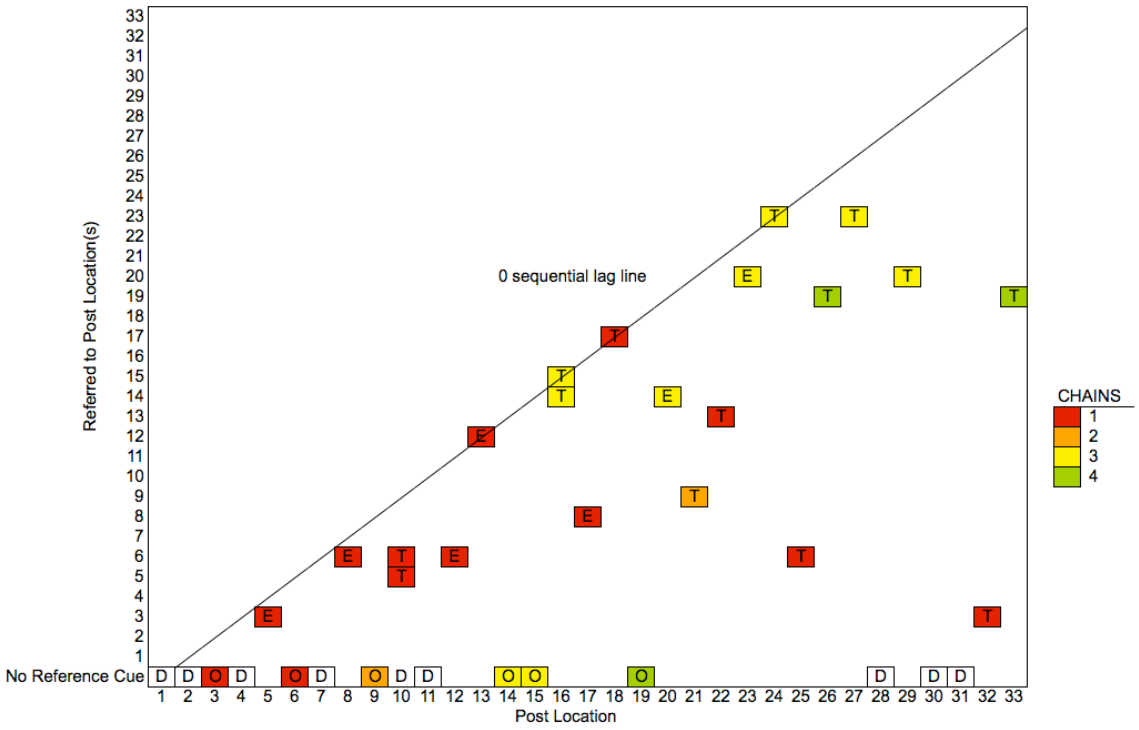

Figure G3. Sequential lag map, S3.

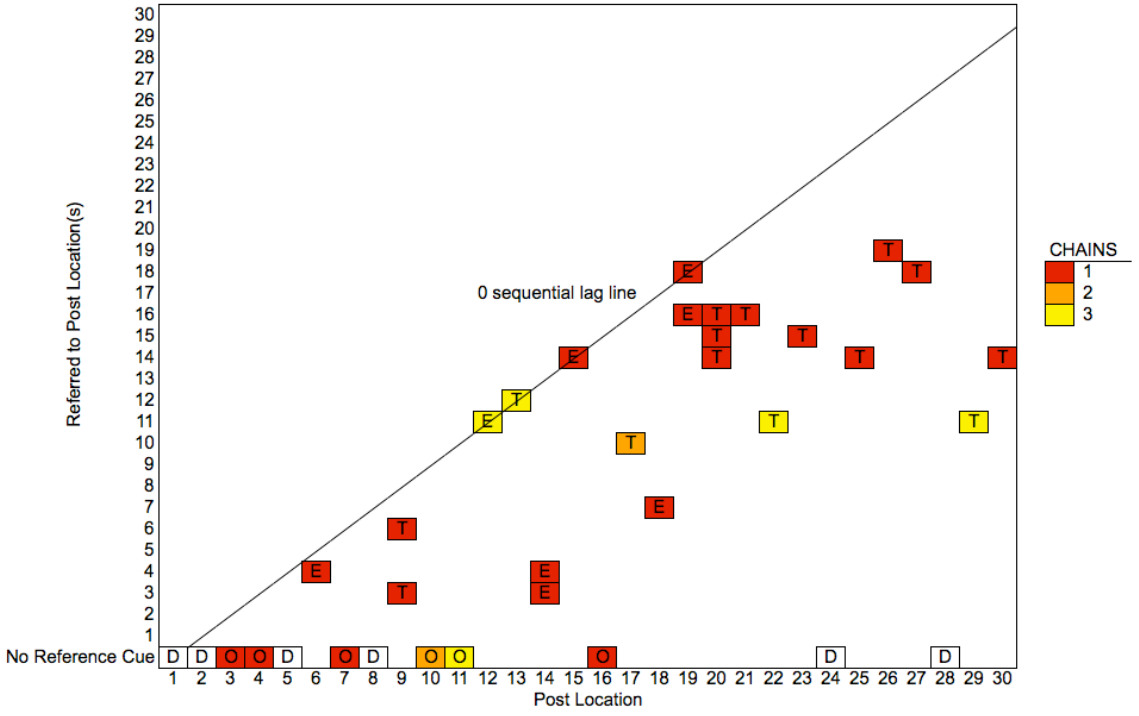

Figure G4. Sequential lag map, S4.

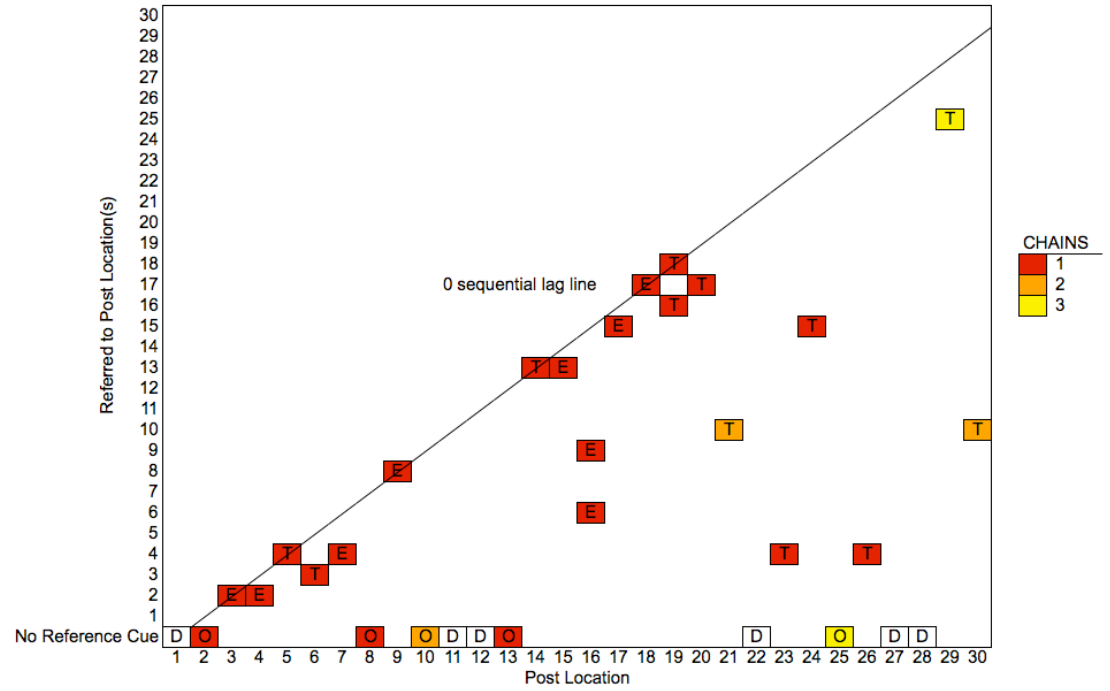

Figure G5. Sequential lag map, S5.

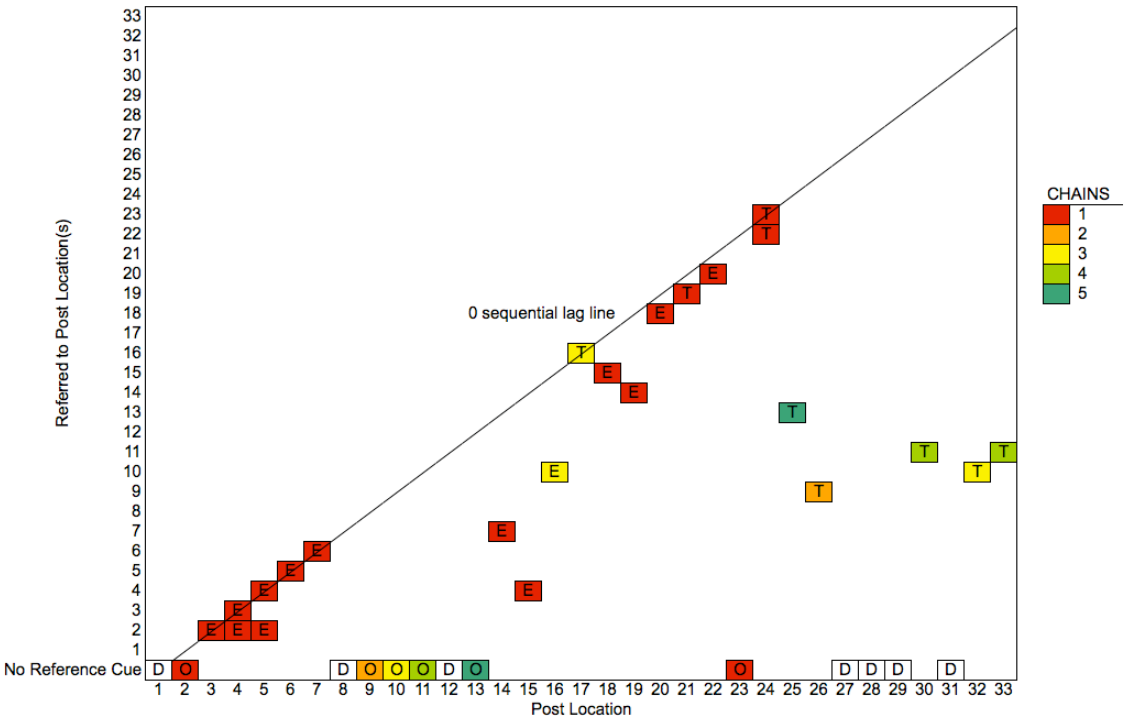

Figure G6. Sequential lag map, S6.

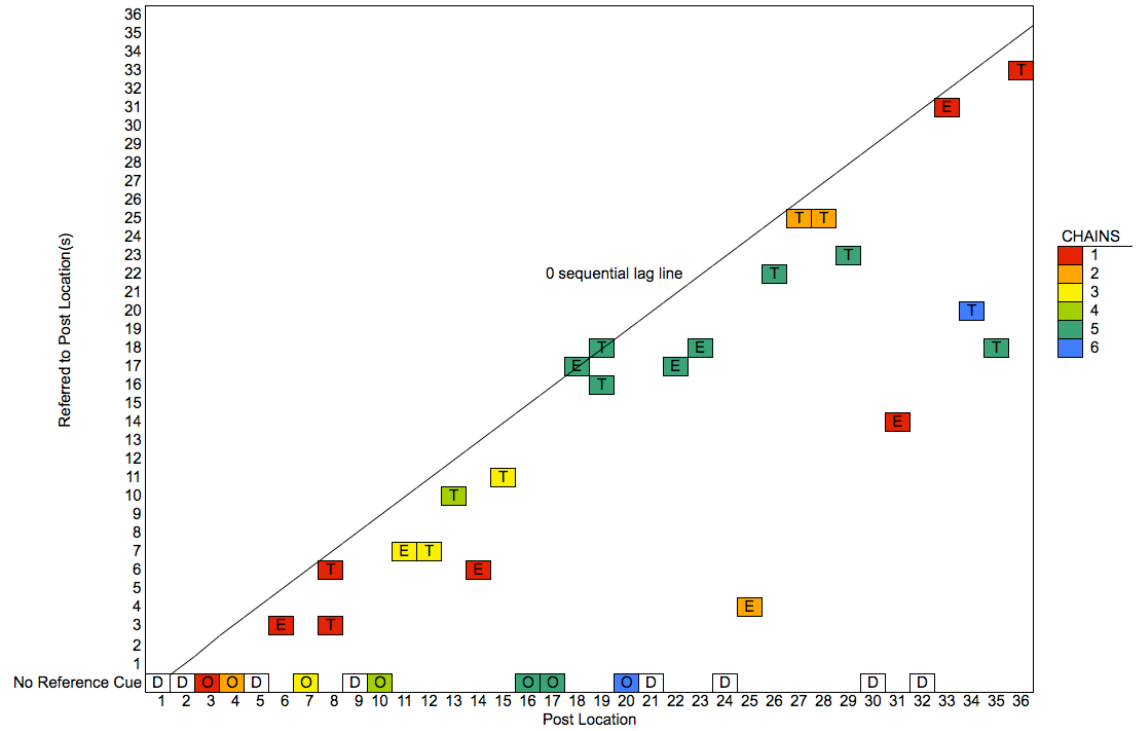

Figure G7. Sequential lag map, S7.

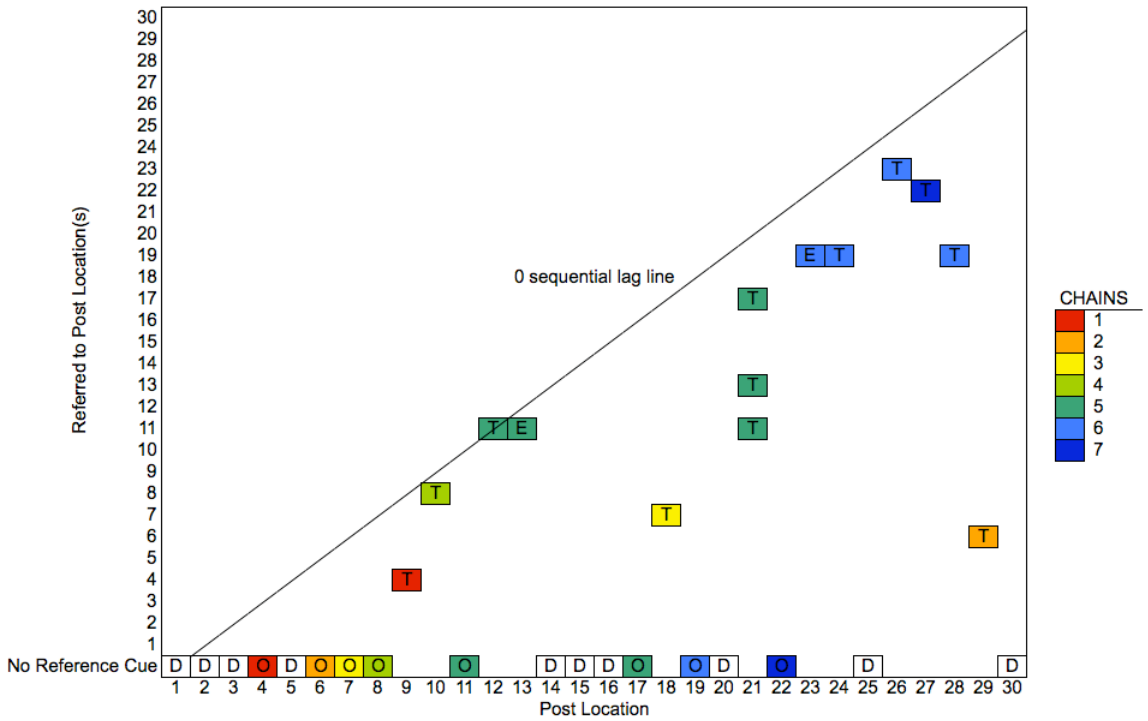

Figure G8. Sequential lag map, S8.

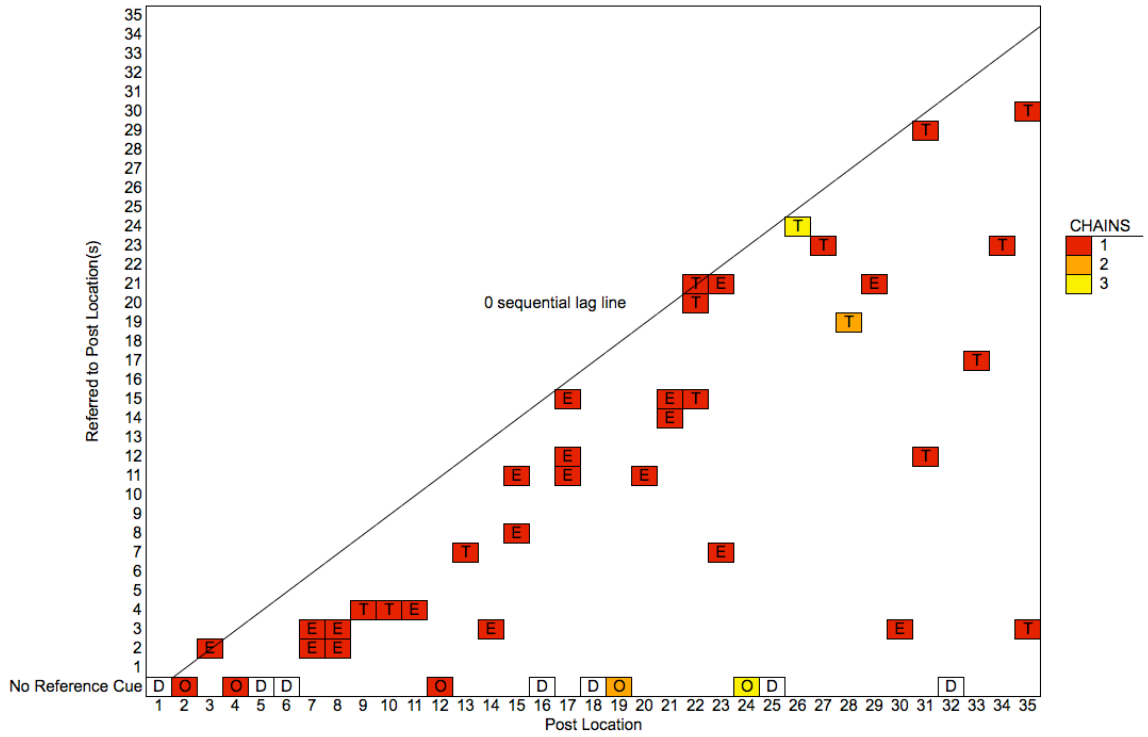

Figure G9. Sequential lag map, S9.

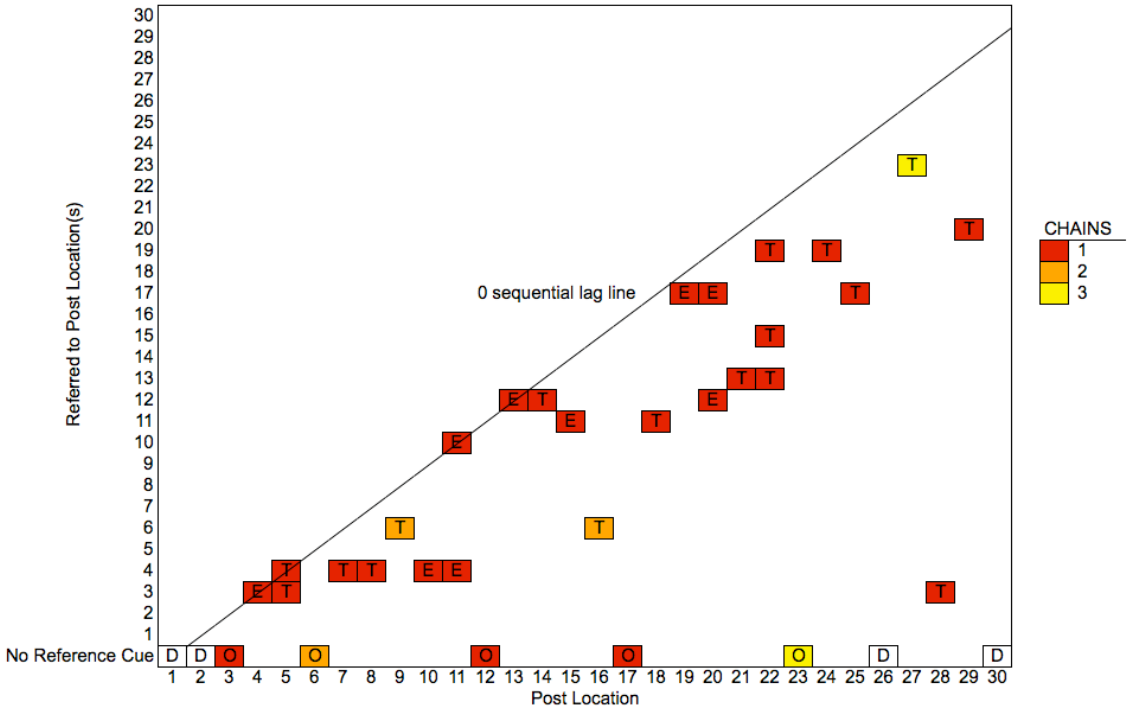

Figure G10. Sequential lag map, S10.

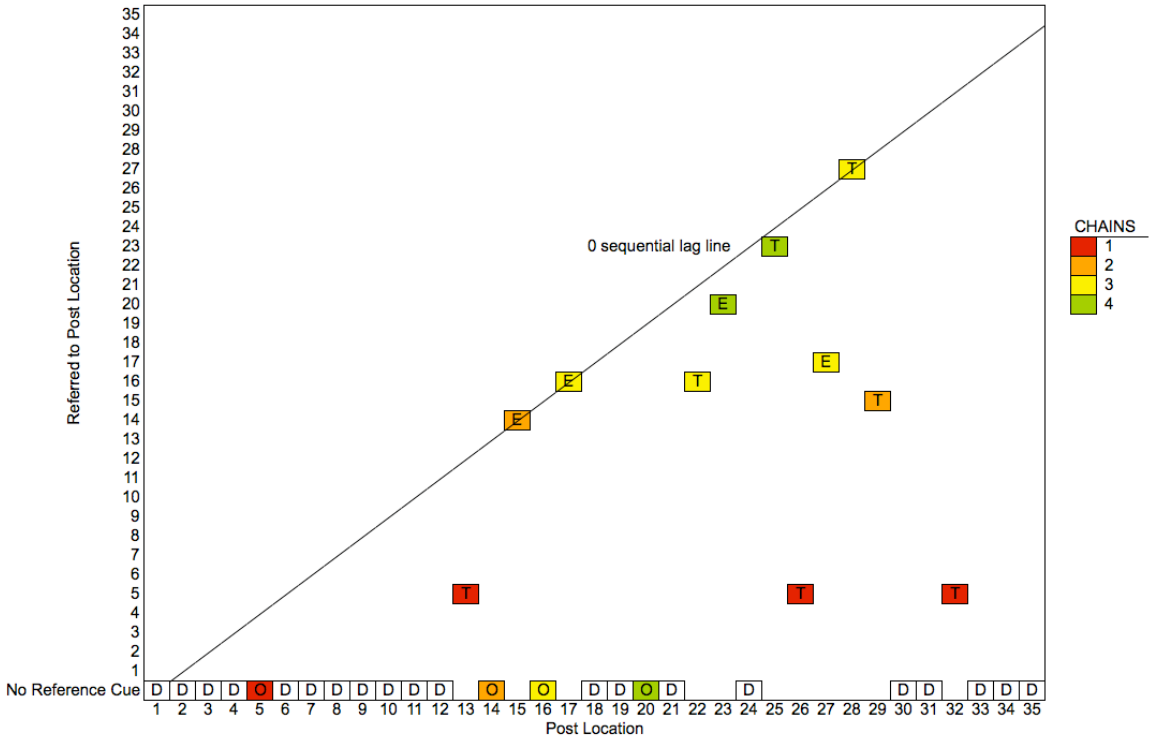

Figure G11. Sequential lag map, F1.

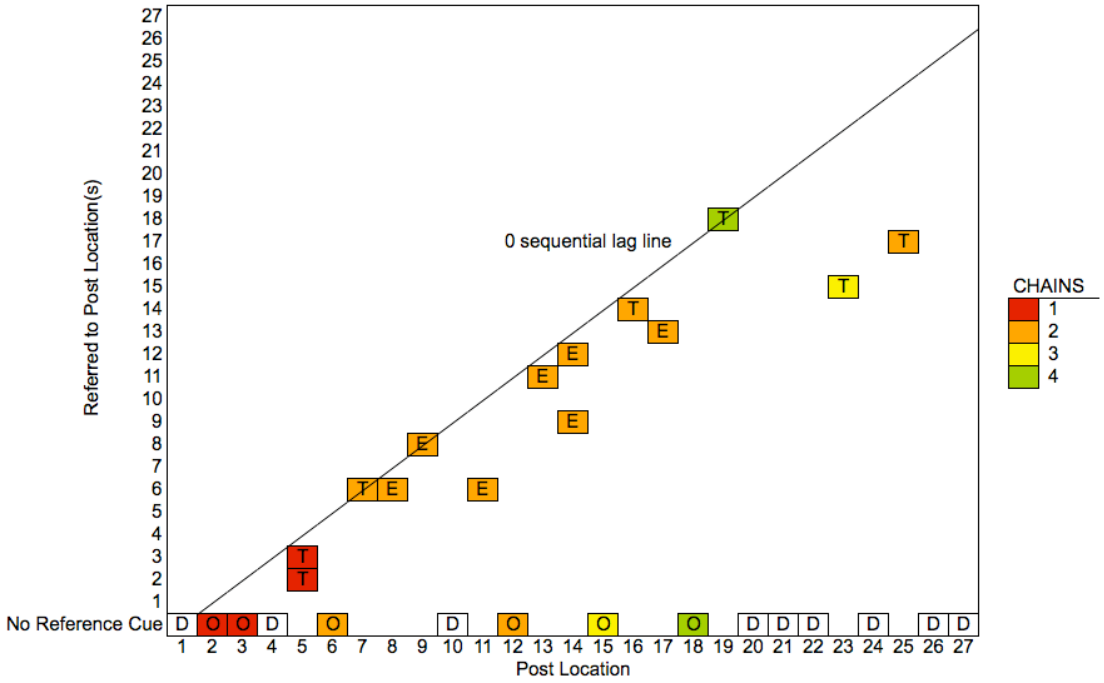

Figure G12. Sequential lag map, F2.

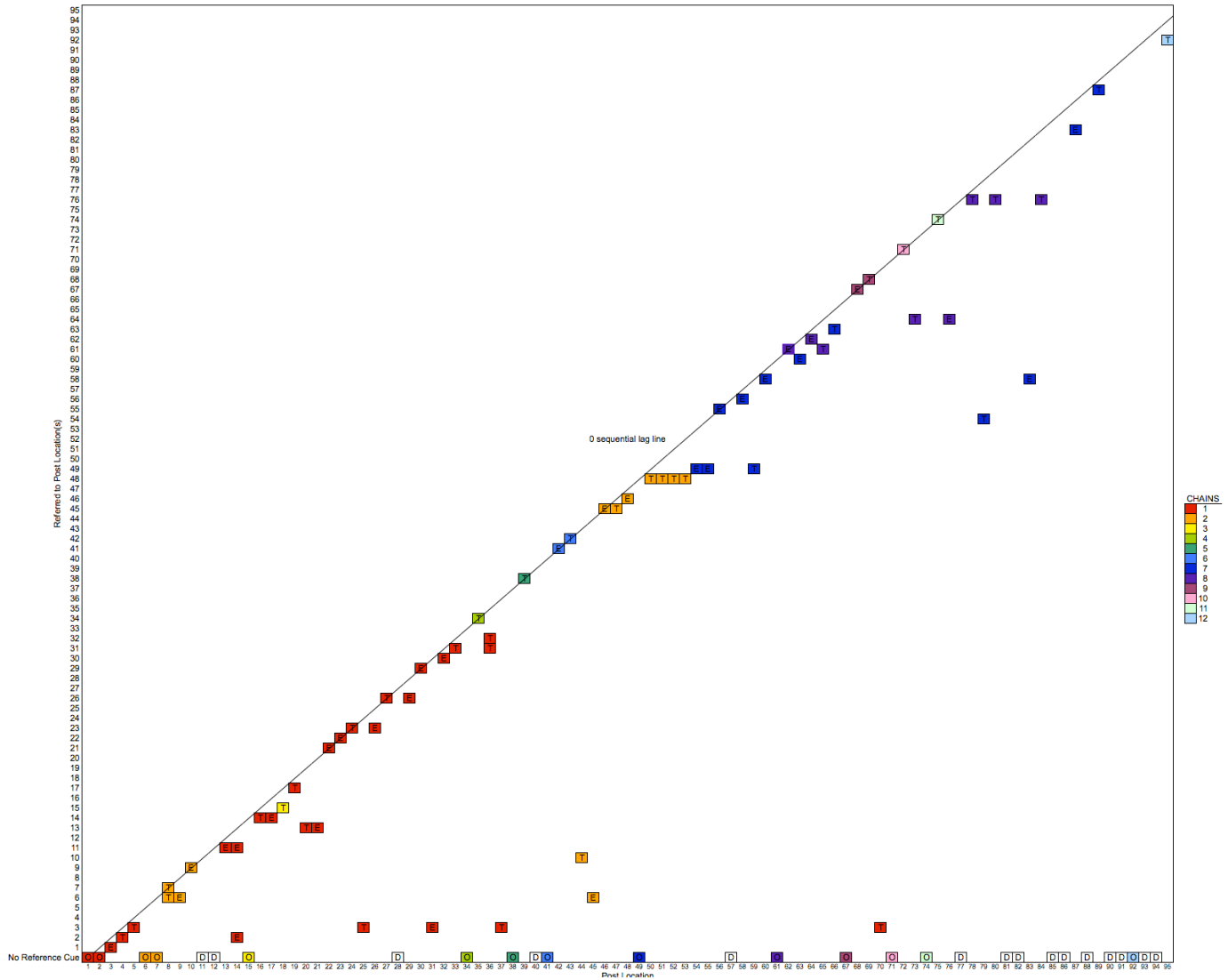

Figure G13. Sequential lag map, F3.

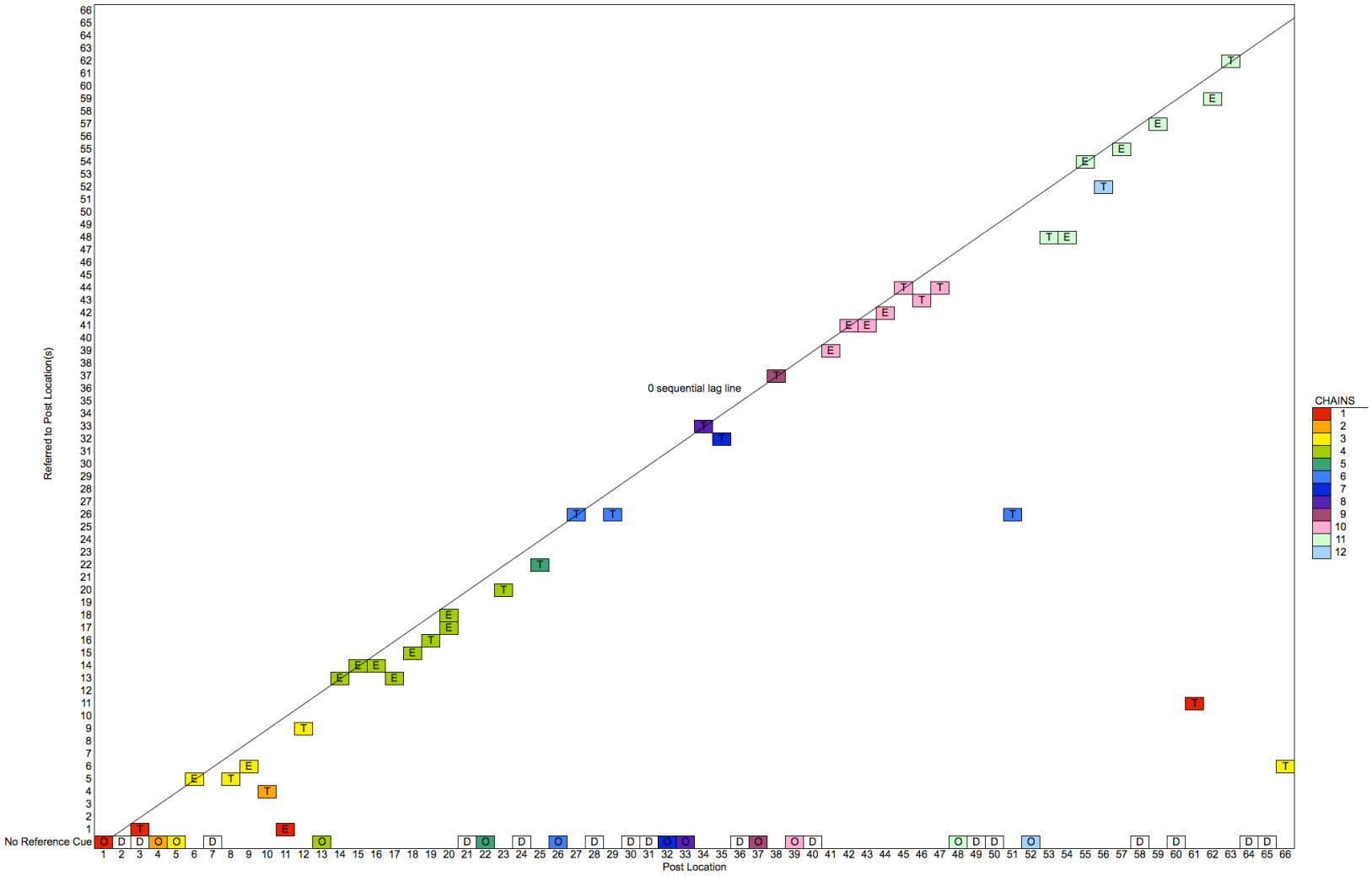

Figure G14. Sequential lag map, F4.

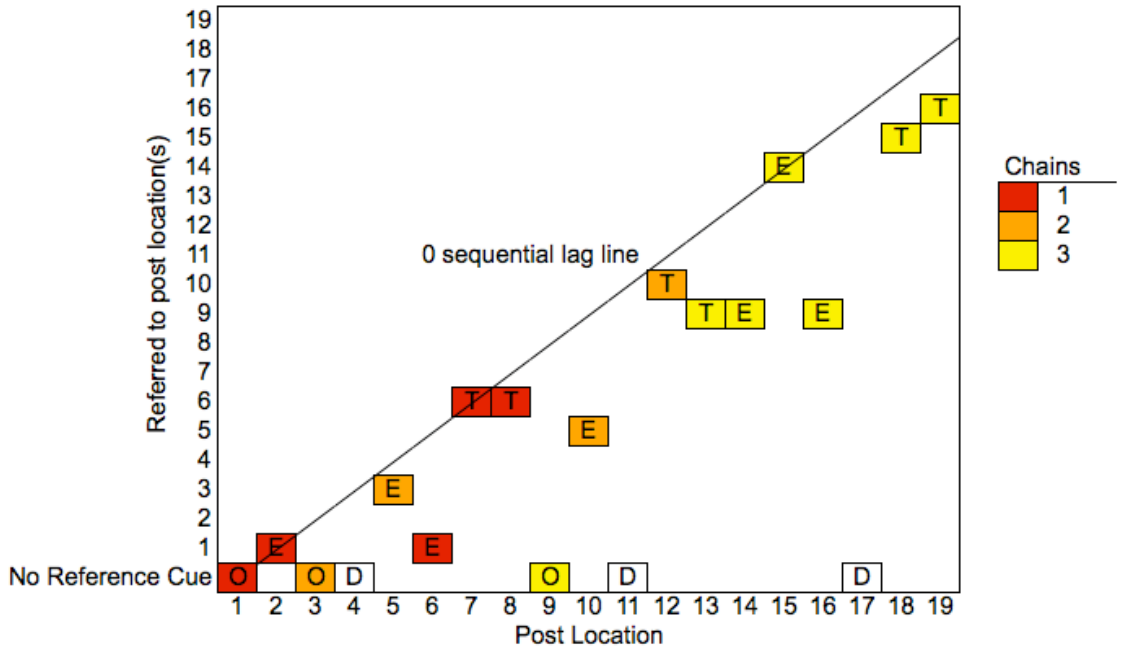

Figure G15. Sequential lag map, F5.

REFERENCES

Alexander, J.O. (1999). Collaborative design, constructivist learning, information technology immersion, and electronic communities: A case study. *Interpersonal Computing and Technology: An Electronic Journal for the 21st Century, 7*(1/2).

Alvesson, M. and Kärreman, D. (2000). Taking the linguistic turn in organizational research. *The Journal of Applied Behavioral Science, 36*(2), 136-158.

Aviv, R. Erlich, Z., Ravid, M. A. and Geva, A. (2003). Network analysis of knowledge construction in asynchronous learning networks. *Journal for Asynchronous Learning Networks, 7*(3), 1-23.

Aviv, R., Erlich, Z. and Ravid, G. (2004). Design and architecture of collaborative online communities: A quantitative analysis. Proceeding of the 27th *Association for Educational Communications & Technology Conference,* Chicago, IL.

An, H., Shin, S., and Lim, K. (2009). The effects of different instructor facilitation approaches on students' interactions during asynchronous online discussions. *Computers & Education, 53*, 749-760.

Austin, J. L. (1962). *How to do things with words.* Cambridge, MA: Harvard University Press.

Bakeman, R. (1974). The size of informal groups in public. *Environment and Behavior, 6*(3), 378-390.

Bateson, G. (1996). Communication. In H. B. Mokros (Ed.), *Interaction and identity: Information and behavior* (pp. 45-70). New Brunswick, NJ: Transaction Books.

Baym, N. K. (1995). The Performance of Humor in Computer-Mediated Communication. *Journal of Computer-Mediated Communication, 1*(2).

Baym, N. K. (1996). Agreements and disagreements in a computer-mediated discussion. *Research on Language and Social Interaction, 29*, 315-345.

Baym, N. K.(1998). The emergence of on-line community. In S. G. Jones (Ed.) *Cybersociety 2.0: Revisiting computer-mediated communication and community* (pp. 35-68). Thousand Oaks, CA: Sage.

Benoit, W. L., Klyukovski, A. A., McHale, J. P., & Airne, D. (2001). A fantasy theme analysis of political cartoons on the Clinton-Lewinsky-Starr affair. *Critical Studies in Media Communication, 18(*4), 377-394.

Blau, P.M., McHugh-Falbe, C., McKinley, W. and Tracy, P.K. (1976). Technology and organization in manufacturing. *Administrative Science Quarterly, 21*, 20-40.

Bormann, E. G. (1972). Fantasy and rhetorical vision: The rhetorical criticism of social reality. *Quarterly Journal of Speech, 58*(4), 396-407.

Browning, L. D. & Stephens, K. K. (2004). Giddens' structuration theory and ICTs. In L. D. Browning, Saetre, A. S., Stephens, K. K., & Sornes, J. O. (Eds.), *Information & communication technology in action: Linking theory and narratives of practice* (pp. 47-55). Herndon, VA: Abstrakt.

Burnett, G. (2000). Information exchange in virtual communities: A typology. *Information Research, 5*(4).

Chester, A. and Gwynne, G. (1998). Online teaching: Encouraging collaboration through anonymity. *Journal of Computer-Mediated Communication, 4*(2).

Chiu, C. M., Hsu, H. M., & Wang, E. T. G. (2006). Understanding knowledge sharing in virtual communities: An integration of social capital and social cognitive theories. *Decision Support Systems, 42*, 1872-1888.

Cockett, L. S. (2000). Self, other and situation in collaborative contexts (Doctoral dissertation, Rutgers University, 2000).

Crystal, D. (2006). *Language and the Internet*. New York: Cambridge University Press.

Dholakia, U. M., Bagozzi, R. P., & Pearo, L. K. (2004). A social influence model of consumer participation in network- and small-group-based virtual communities. *International Journal of Research in Marketing, 21*, 241-263.

Dorfman, P. W. and Stephan, W. G. (1984). The effects of group performance on cognitions, satisfaction, and behavior: A process model. *Journal of Management*, *2*, 173-192.

Duncan, S. (1969). Nonverbal communication. *Psychological Bulletin, 72*(2), 118-137.

Duncan, S. and Fiske, D. W. (1977). *Face-to-face interaction: Research, methods and theory*. Hillsdale, NJ: Erlbaum.

Duncan, S. and Fiske, D. W. (1985). *Interaction structure and strategy*. New York: Cambridge University Press.

Dustdar, S. and Gall, H. (2003). Architectural concerns in distributed and mobile collaborative systems. *Journal of Systems Architecture, 49*, 457-473.

Farooq, U. Schank, P. Harris, A. Fusco, J. and Schlager, M. (2007). Sustaining a community computing infrastructure for online teacher professional development: A case study of designing *Tapped In*. *Computer Supported Cooperative Work, 16*, 397-429.

Ferrara, K. Brunner, H. and Whittemore, G. (1991). Interactive written discourse as an emergent register. *Written Communication, 8*(1), 8-34.

Garton, L. Haythornethwaite, C., and Wellman, B. (1997). Studying online social networks. *Journal of Computer-Mediated Communication, 3*(1).

Goffman, E. (1983). The interaction order: American Sociological Association, 1982 presidential address. *American Sociological Review, 48*, 1-17.

Grint, K. and Woolgar, S. (1997). *The machine at work*. Cambridge: Polity Press.

Herring, S. C. (2002). Computer-mediated communication on the Internet. *Annual Review of Information Science and Technology,* 109-168.

Herring, S. C. (2004). Slouching toward the ordinary: Current trends in computer-mediated communication. *New Media & Society, 6*, 26-36.

Hine, C. (2000). *Virtual Ethnography*. New York: Sage.

Hratsinski, S. and Keller, C. (2007). Computer-mediated communication in education: A review of recent research. *Educational Media International, 44*(1), 61-77.

Hutchby, I. (2005). *Conversation & technology: From the telephone to the Internet*. Malden, MA: Polity Press.

Johnson, C. (2001). A survey of current research on online communities of practice. *The Internet and Higher Education, 4*, 45-60.

Karetnick, R. D. (2000). Identity in cyberspace: An ethnographic and microanalytic study of participation in a virtual community. (Doctoral dissertation, Rutgers University, 2000).

King, W. R., Marks, P. V. and McCoy, S. (2002). The most important issues in knowledge management. *Communications of the ACM, 45*(9), 93-97.

Kurabacak, G. (2005). Learners and faculty creative partnerships through building a virtual community: Their experiences toward sharing knowledge online. In C.

Crawford et al. (Eds.), *Proceedings of Society for Information Technology and Teacher Education International Conference 2005* (pp. 2271-2278). Chesapeake, VA: AACE.

Lindlof, T. R. & Taylor, B. C. (2002). Qualitative Communication Research Methods: Second Edition. London: Sage Publications.

Looi, C-K. and Ang, D. (2000). A multimedia-enhanced collaborative learning environment. *Journal of Computer Assisted Learning, 16*, 2-13.

Maltz, (1996). Customary Law & Power in Internet Communities. *Journal of Computer-Mediated Communication, 2*.

Marcoccia, M. (2004). On-line polylogues: Conversation structure and participation framework in Internet newsgroups. *Journal of Pragmatics, 36*, 115-145.

McQuown, N.E. (Ed.). (1971). The Natural History of an Interview. *Microfilm Collection of Manuscripts in Cultural Anthropology*. Chicago: Regenstein Library.

Mennecke, B. E. and Valacich, J. S. (1998). Information is what you make of it: The influence of group history and computer support on information sharing, decision quality, and member perceptions. *Journal of Management Information Systems, 15*, 173-197.

Meyers, R. Davis, H., and Botti, J. (2002). Professional development: Building effective virtual communities through cooperative learning. *NECC 2002: National Educational Computing Conference Proceedings*.

Mokros, H. B. (1984). Patterns of persistence and change in the sequencing of nonverbal actions (Doctoral dissertation, The University of Chicago, 1984).

Mokros, H. (1985). Pattern of persistence and change in action sequences. In Starkey & Duncan (Eds.) *Interaction structure and strategy* (pp. 175-232). New York: Cambridge University Press.

Mokros, H. (2003). *Identity matters: Communications-based explorations and explanations.* Cresskill, NJ: Hampton Press, Inc.

Mokros, H.B. (2006). Composing relationships at work: Three minutes at a wholesale produce market. In J. Wood & S. Duck (Eds.), *Composing relationships*, (pp. 175-185). Belmont, CA: Wadsworth.

Orlikowski, W. J. (1992). The duality of technology: Rethinking the concept of technology in organizations. *Organization Science, 3*, 498-427.

Paccagnella, L. (1997). Getting the seats of your pants dirty: Strategies for ethnographic research on virtual communities. *Journal of Computer-Mediated Communication, 3*(1).

Paulus, T. M. (2007). CMC modes for learning tasks at a distance. *Journal of Computer-Mediated Communication, 12*, 1322-1345.

Peirce, C. S. (1955). Abduction and induction. In J. Buchler (Ed.), *Philosophical writings of Peirce* (pp.150-156). New York: Dover.

Poole, M. S. & DeSanctis, G. (1990). Understanding the use of group decision support systems: The theory of adaptive structuration. In J. Fulk & C. Steinfeld (Eds.), *Organizations and communication technology* (pp. 173-193). Thousand Oaks, CA: Sage Publications.

Preece, J. (2000). *Online Communities*. Chichester: Wiley.

Rheingold, H. (2000). *The virtual community: Homesteading on the electronic frontier.* Cambridge, MA: MIT Press.

Rice, R. E. (1987). New patterns of social structure in an information society. In J. Schement & L. Lievrouw (Eds.), *Competing visions, complex realities: Social aspects of the information society* (pp. 107-120). Norwood, NJ: Ablex.

Ridings, C. M., Gefen, D. & Arinze, B. (2002). Some antecedents and effects of trust in virtual communities. *Journal of Strategic Information Systems, 11*, 271-295.

Ridings, C. M., & Gefen, D. (2004). Virtual community attraction: Why people hang out online. *Journal of Computer-Mediated Communication, 10*(1), April 7th, 2006 from http://jcmc.indiana.edu/vol10/issue1/ridings_gefen.html

Rumsey, E. L. (2001). Making sense of health and illness online: A study of patterns of participation and use on one computer-mediated cancer support site (Doctoral dissertation, Rutgers University, 2001). *Dissertation Abstracts International, 62,* 27.

Sawyer, R.K. and Berson, S. (2004). Study group discourse: How external representations affect collaborative conversation. *Linguistics and Education, 15*, 387-412.

Stephenson, H. C. (1998). Guiding becoming: A microanalytic study of one instructor's approach (Doctoral dissertation, Rutgers University, 1998).

Sun, C. T., Lin, H. & Ho, C. H. (2006). Sharing tips with strangers: Exploiting gift culture in computer gaming. *CyberPsychology & Behavior, 9*(5), 560-570.

Tellent-Runnels, M. K., Thomas, J. A., Lan, W. Y., Cooper, S., Ahern, T. C., Shaw, S. M., and Liu, X. (2006). Teaching courses online: A review of the research. *Review of Educational Research, 76*(1), 93-135.

Toffler, A. (1981). *The third wave*. London: Collins.

Tremblay, D-G. (2005). Virtual communities of practice: Explaining different effects in two organizational contexts. *Canadian Journal of Communication, 30*, 367-382.

Walther, J. B. & Burgoon, J. K. (1992). Relational communication in computer-mediated interaction. *Human Communication Research, 19*(1), 50-88.

Wang, Y. & Fenemaeir, D. R. (2004). Towards understanding members' general participation in and active contribution to an online travel community. *Tourism Management, 25*, 709-722.

Ward, K. J. (1999). Cyber-ethnography and the emergence of the virtually new community. *Journal of Information Technology, 14*, 95-105.

Wasko, M. M. & Faraj, S. (2000). "It is what one does": Why people participate and help others in electronic communities of practice. *Journal of Strategic Information Systems, 9*, 155-173.

Wasko, M. N., Teigland, R. and Faraj, S. (2009). The provision of online public goods: Examining social structure in an electronic network of practice. *Decision Support Systems, 47*, 254-265.

Wheeler, S. (2006). Learner support needs in online problem-based learning. *The Quarterly Review of Distance Education, 7*, 175-184.

Winzelberg, A. (1997). The analysis of an electronic support group for individuals with eating disorders. *Computers in Human Behavior, 13*(3), 393-407.

Ye, S., Chen, H. & Jin, X. (2006). An empirical study of what drives users to share knowledge in virtual communities. *KSEM*.

CURRICULUM VITAE

Education

2005 – 2010 Rutgers, The State University of New Jersey
School of Communication and Information
Ph.D., January 2010

2003 – 2005 Rutgers, The State University of New Jersey
School of Communication and Information
Masters in Communication and Information, May 2005

1997 – 2001 Rutgers, The State University of New Jersey
BA in English; Psychology Minor, May 2001

Positions Held

2007 – 2009 Graduate Assistant, Institute of Marine and Coastal Sciences, Education and Outreach

2002 – 2007 Communication Specialist, Institute of Marine and Coastal Sciences, Education and Outreach

2001 – 2002 Administrative Assistant, Thomas Publishing Company, Thomas Global Register