7-6-2010

The Effect of cross-chain discussion by eyewitnesses and non-eyewitnesses on rumor accuracy

Samuel Brougher

THE EFFECT OF CROSS-CHAIN DISCUSSION BY EYEWITNESSES AND NON-EYEWITNESSES ON RUMOR ACCURACY

Samuel J. Brougher

Rochester Institute of Technology

Author Note

Samuel J. Brougher, Master of Science student in Applied Experimental and Engineering Psychology in the Psychology Department, Rochester Institute of Technology.

Submitted in partial fulfillment of the requirements for the degree of Master of Science in Applied Experimental and Engineering Psychology in the Graduate College of the Rochester Institute of Technology, Rochester, NY. Thesis approved: 07/06/2010. This material was based upon work supported by the National Science Foundation under Grant BCS-0527371. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation. Thanks to Nicholas DiFonzo (thesis advisor), Esa M. Rantanen (thesis committee member), and Kirsten Condry (thesis committee member) for help with the planning, implementation, and writing of this research.

Correspondence concerning this paper should be addressed to Samuel J. Brougher,
E-mail: sambrougher@gmail.com

Abstract

As rumors travel from person to person, information is lost during the transmissions. Discussion, however, increases retention of the original account and organizes later transmissions. This organization then decreases the attentional load of the listeners, thus allowing them to attend to and retain a more complex message. Further, the information that is lost tends to be explanatory statements (instead of event statements). Such explanatory statements make ratings of a negative event more neutral, but are lost when the listener's attentional load is increased, and the central (event-based) message is more likely to be retained. With this in mind, the effects of discussion among eyewitnesses and subsequent listeners (non-eyewitnesses) were examined. Participants were exposed to the negative events of an actor's drunken night mitigated by such things as he did not know he was drunk. Half the participants discussed, the other half did not. All participants made audio recordings of their versions of the story and rated the actor. A second-generation then got those recordings and followed the same procedure. Discussion increased accuracy of individual reports (except in the second-generation following first-generation discussion). Nevertheless, this advantage was dampened by looking at the full amount of information transmitted by a pair (their pooled accuracy) which showed no difference between discussers and non-discussers. This reduces validity of arguments both for and against eyewitnesses discussing with each other; it does not seem to affect the amount of information available to, or used by, the jury. Ordering of subsequent transmissions was always highly similar to the original and accounted for a significant portion of the central-peripheral bias.

Keywords: communication, discussion, eyewitness, rumor, serial chain, word of mouth

The Effect of Cross-Chain Discussion by Eyewitnesses and Non-Eyewitnesses on Rumor Accuracy

Rumors are “unverified account[s] or explanation[s] of events, circulating from person to person and pertaining to an object, event, or issue of public concern” (Peterson & Gist, 1951, p. 159) or “unverified and instrumentally relevant information statements in circulation that arise in contexts of ambiguity, danger, or potential threat and that function to help people make sense and manage risk” (DiFonzo & Bordia, 2007, p. 13). The common thread between these definitions is that rumors are unsubstantiated, important statements that are transmitted from one person to another. To be classified as a rumor, the statement must be unverified, since there must be some uncertainty as to the verity of the statement. The statement must be important; otherwise people would not be motivated to talk about it. Finally, the statement must be transmitted, since a statement within a single person’s head is not a rumor. This thesis focused on the accuracy of rumor statements as they relate to discussion before and after transmission.

Transmission occurs when one person tells something to another person (in this instance, a rumor; Allport & Postman, 1945; 1946; DiFonzo & Bordia, 2007). The person telling the rumor is called the transmitter, teller, source, or parent. The person being told the rumor is called the listener, target, or child. A transmission pair involves two people: one teller and one listener. The same person could be a listener in one instance (when they are told the rumor) and a teller in another (when they later tell the rumor to someone else). The first transmission of a rumor is called the origin, original, or original source. Except for the original teller, a person must be a listener before they can become a teller. The final transmission of a rumor is called the terminal report.

Most rumor research of transmission focuses on serial chains (e.g., Allport & Postman, 1945). A serial chain is formed by having an original source given to a participant. That person then transmits the rumor to one listener. The listener then becomes a teller and transmits the rumor to another person. This chain of transmissions (similar to the telephone game) continues until the experiment stops with the terminal report. Serial chains can either be one-way (i.e., the teller speaks and the listener does not have a chance to say anything) or with discussion (i.e., the listener can speak and ask questions of the teller). In the serial chain, the first participant in a chain is said to be the first-generation. That participant's listener is the second-generation, followed by the third-generation, etc., until the terminal report is reached.

Discussion is equivalent to "interaction" as described by DiFonzo and Bordia (2007): Back and forth communication that allows for "clarification, comparison, and interpretation" (p. 175). While DiFonzo and Bordia limited interaction to transmission pairs, the same basic definition is adapted here to two listeners (who could potentially become tellers). Research involving discussion typically uses the implicit definition, "Whatever participants do when asked to talk about, discuss, collaborate, or cooperate to remember the stimulus," as their definition for discussion (Andersson & Rönnberg, 1995; Hope, Ost, Gabbert, Healey, & Lenton, 2008; Leavitt & Mueller, 1951; Paterson & Kemp, 2005; Weldon & Bellinger, 1997; Yuker, 1955).

Historically, experiments examined the accuracy of rumor transmission in a serial chain (e.g., Allport & Postman, 1945; Lyons & Kashima, 2003; 2006) or of rumors in the public (e.g., Peterson & Gist, 1951). Few experiments have studied discussion during rumor transmission in a serial chain (e.g., McAdam, 1962) and the author knows of no studies that examine the effects of earlier discussion in a chain affecting the accuracy of the rumor later in the chain.

Within eyewitness and memory research, however, many studies have examined the effects of discussion on individual and overall accuracy of later recall (e.g., Andersson & Rönnberg, 1995; Gabbert, Memon, & Wright, 2006; Hope et al., 2008; Luus & Wells, 1994; Skagerberg, 2007; Weldon & Bellinger, 1997; Yuker, 1955); as have studies of decision making (e.g., Henningsen & Henningsen, 2004; Stasser & Titus, 1985); and education (Leavitt & Mueller, 1951). Unfortunately, these studies never examined transmission of the discussed recollections, only accuracy following initial discussion.

Therefore, the current body of research does not examine how discussion influences the accuracy of a message being transmitted multiple times. This study examined the accuracy of a rumor that was transmitted, discussed or not discussed, transmitted again, then (again) discussed or not.

For our purposes, accuracy was defined as DiFonzo and Bordia (2007) describe precision: “The degree to which the rumor corresponds with an original perception or message” (p. 142), focusing particularly on completeness of the transmission, as opposed to distortion. Implications for rumor research and eyewitness testimony are discussed.

Three propositions are proposed: through leveling, information is lost during transmissions; discussion increases information retained in subsequent tellings; and inaccurate or incomplete tellings will result in subsequent tellings being more inaccurate and incomplete.

Proposition 1: Leveling

Leveling. When a statement (such as a rumor) is transmitted from one person to another, without discussion, the statement becomes more concise and less elaborate (Allport & Postman, 1945; 1946). Details are discarded because of memory limitations. As such, most information is lost at the beginning of the rumor's life, as the first people cut the story down to a size manageable by human working memory. Over a series of five transmissions, as many as 14 of 20 details can be lost (Allport & Postman, 1945). The remaining details are likely to be the most important, relevant, or memorable. I will refer to these as the central details. In the case of a story, the *central* details are the ones that are “fundamental to a story line or plot” on the main causal-temporal chain (Lyons & Kashima, 2003, p. 992); the other details that serve as background or explanatory details are *peripheral* details (Lyons & Kashima, 2003). This data reduction is called leveling; the accenting and highlighting of the central details is called sharpening (Allport & Postman, 1945; 1946).

Despite the evidence for information loss over the course of a rumor's life during serial transmission, Peterson and Gist (1951) argued that rumors do not function naturally in a serial transmission paradigm. They said when rumors occur naturally, adding occurs, not leveling. Adding is the integration of completely new and previously untransmitted information to the account. Nevertheless, their evidence for adding only analyzes the total number of statements within a population as time goes on, as opposed to the total number of statements exhibited during an individual transmission compared to a previous transmission, so little is known about how and where these additions occur. The best experiments of short-term rumor transmission are those of serial transmission.

What is lost? Leveling is not random. While the information loss results from memory economization, the information that is retained is guided by what Allport and Postman (1945; 1946) called assimilation. That is, the content of later transmissions will fit the general theme, or stereotype better than earlier transmissions (cf. Lyons & Kashima, 2003; 2006). This change occurs because inconsistencies with the story's theme or stereotype are often eliminated from one generation to another. That is, stereotype-consistent information is kept, and stereotype-inconsistent information is discarded so the entire story fits the central theme or stereotype. For instance, Allport and Postman (1945) found that, during rumor transmission, a church with a clock tower in a war story was reported as just a clock or clock tower, eliminating the inconsistent idea of "church" from the stereotype of the war scene.

During the Teller-Listener Extremity experiments (Baron et al., 1997; Gilovich, 1987; Inman et al., 1993), participants were presented with the verbal account (an actor on video or audio) of negative (or positive) events (e.g., "I drove drunk and got in a car accident"). This event information (which was stereotype-consistent; Lyons & Kashima, 2003; 2006) was mitigated by extenuating circumstances that caused participants to rate the actor more neutrally (e.g., "My drink was spiked without my knowledge and the other driver ran a stop sign"). These mitigating statements were stereotype-inconsistent (Lyons & Kashima, 2003; 2006). Note that, while Gilovich, (1987), Inman et al., (1993) and Baron et al. (1997) used the word "event," their events were equivalent to Lyons and Kashima's (2003) "central" statements; the events were the central statements in the story. Further, the "mitigating" statements of Gilovich, (1987), Inman et al., (1993) and Baron et al. (1997) were equivalent to Lyons and Kashima's (2003) "peripheral" statements; the mitigating statements were peripheral to the story. This does not have to be the case (e.g., it is possible to construct a story with peripheral events), but that was how these studies were designed.

When participants transmitted the story to the next generation of listeners the mitigating circumstances tended to be lost, resulting in more negative (or positive) ratings of the original actor. This effect was named the "Teller-Listener Extremity" effect; between the transitions from the tellers (first-generation) to the listeners (second-generation), the ratings of the actor became more extreme. Inman et al. (1993) discovered that the first-generation tellers included the peripheral information along with the central information, but the second-generation was not receiving the peripheral information. Based on similar research by Gilbert, Pelham, and Krull (1988), Baron et al. (1997) showed that when the first-generation's account was less clear than the original, and this increased the attentional load on participants in the second-generation. Baron et al. (1997) found that manipulating the temporal order of statements (i.e., disorganizing them) or increasing attentional load in other ways (e.g., adding noise) caused the second-generation participants to ignore the peripheral information and only remember the central parts of the story.

This research leads to the first proposition: When transmitting information serially from one person to another without discussion, information is lost (called leveling). Furthermore, central information is remembered better than peripheral information. This bias toward remembering central information is because the account received by the second-generation is less clear than the original account which causes a loss of peripheral information. That is, the information loss is not random.

Proposition 2: Discussion

With discussion between two or three people, a more accurate recall will be developed than any single individual's recall, but less accurate than a group of individuals working independently with their results pooled (Andersson & Rönnberg, 1995; Weldon & Bellinger, 1997; Yuker, 1955). For instance, Andersson and Rönnberg (1995) split people into two kinds of recall pairs: those who discussed the event and those who did not. All unique responses from the non-discussing pair were pooled together to create the nominal score, which was consistently higher than the score of the discussing pair. So, if four participants (A, B, C, and D) were presented with five words (cat, dog, toy, fish, goat) and split into two pairs (AB and CD), one pair (AB) discussed while recalling, and the other pair (CD) each recalled individually, then AB might remember four words (cat, dog, toy, fish), C might remember three words (cat, dog, fish), and D might remember three words (toy, fish, goat). The non-unique responses of C and D would be combined and result in five words (cat, dog, toy, fish, goat). So, while the discussers (AB) would have more responses than an individual (C or D), the combined efforts of individuals (CD) would be more than the efforts of discussers.

Nevertheless, these same studies and Basden, Basden, and Henry (2000) showed that individual recall following group discussion was better than an individual's recall without group discussion. This shows that, while two individuals working alone are better than two individuals working together, two individuals working alone following collaboration are better than two individuals working alone. Since the serially transmitted account after discussion is more complete, more peripheral information should be acquired by subsequent generations, lessening the Teller-Listener Extremity effect compared to recall without discussion (Baron et al., 1997).

The first proposition is: Information is lost during transmission, but this second proposition is that discussion can increase information retention of a mutually experienced event, and organization of subsequent transmissions, if given information is accurate. Thus, while information is typically lost, some of it can be retained with discussion.

Proposition 3: Inaccurate or Incomplete Information

The previous section showed that discussion when complete information is available results in better retention of information. When information given to listeners is inaccurate (changed or added) or incomplete (information is lost), what happens to those listeners' later transmissions?

In cases of the introduction of misinformation, co-witnesses with different information decreased the accuracy of their partners' individual report (e.g., Hope et al., 2008; Luus & Wells, 1994; Paterson & Kemp, 2006a; Skagerberg, 2007; Wright, Self, & Justice, 2000). The term "co-witnesses" refers to witnesses of the same event (Luus & Wells, 1994). These studies use deception (e.g., telling participants they witnessed the same event, but they did not) or confederates to introduce the different information. Essentially, the experimenters added the extraneous information.

Nevertheless, adding is unlikely to occur naturally at the first-generation, or even at the second-generation. Allport and Postman (1945; 1946) never found adding (only leveling and distortion), even after 7 generations; Lyons & Kashima (2003, experiment 1) found few added items spread evenly across participants and excluded the additions from analyses; and Lyons and Kashima (2006) found very few additions *or* distortions (26 of 328 total statements across participants in the communication condition). Instead, leveling occurred from the first- to second-generations. In a situation where two first-generation tellers witness the same event and each transmits it serially to a separate person, participants in the second-generation will receive two different, correct, but incomplete, versions of the story with very little added or changed information.

One study, Hope et al. (2008), presented two co-witnesses with different camera angles for the same event, thus giving them accounts that were both technically accurate, but different. One participant saw a culprit steal money out of a wallet, but the view the other participant had of the culprit was blocked at that time, so the other participant did not see this event. After discussion, the participant who did not actually see the culprit steal money was more likely to report that he or she saw the culprit steal money even though he or she did not personally witness it. While it was technically inaccurate from the individual’s viewpoint, the report gave a better, more complete picture of what actually occurred (given an omniscient viewpoint).

Hope et al. (2008), however, only analyzed a short series of events (two events from a short video) with no peripheral information. Since peripheral information is much more likely to be lost between generations (Baron et al., 1997; Gilovich, 1987; Inman et al., 1993), it is likely that neither second-generation discussant will have access to much of this peripheral information, creating differing incomplete (but correct) accounts.

Stasser and Titus (1985) found, when given incomplete, different, but correct sets of information, people were much more likely to talk about information that everybody knew, as opposed to information that only one person knew. In their study, participants were given a profile of a candidate for student government. This profile contained many pieces of information, but every participant's profile was incomplete in different ways, giving all participants different, correct, but incomplete versions of the profile. During discussion, the pieces of information that were common to several participants were much more likely to be discussed than pieces of information of which only one participant had access. This is likely to result in second-generation discussers talking only about common information. Since peripheral information is lost more, participants will only talk about central information, thus decreasing the overall completeness of their transmissions even more.

This leads to the final proposition: Adding would severely decrease later accuracy, but does not occur in significant amounts. Further, if given information is not accurate or complete (which it will be after the first-generation due to leveling), discussion will cause a greater loss of information compared to no discussion.

Purpose of the Research

This research examined the effects of introducing discussion before and after serial, one-way transmission. Specifically, serial chain transmission design (e.g., Allport & Postman, 1945; 1946; Baron et al., 1997; Gilovich, 1987; Inman et al., 1993; Lyons & Kashima, 2003; 2006) and co-witness discussion design (cf. Andersson & Rönnberg, 1995; Basden et al., 2000; Gabbert et al., 2006; Hope et al., 2008; Paterson & Kemp, 2006a; Weldon & Bellinger, 1997; Wright et al., 2000) were combined. In a serial chain transmission design, one participant is exposed to the original stimulus. This participant then describes the stimulus from memory, as best he or she can, to another participant, who then explains the stimulus to the next participant, and so on until the terminal report (forming a chain of participants from the first to last). In co-witness designs, two participants are exposed to slightly different stimuli (they do not know the stimuli are different) and then discuss what they saw and respond to the experimenters' memory tests. In this study, cross-chain discussion was explored; that is, participants in a serial chain were sometimes allowed to discuss with participants of an equivalent generation before transmitting the stimuli.

Participants heard a first-person report of a series of negative events, discussed the report (or did not discuss it) with another eyewitness, then transmitted the story into a digital audio recorder. The next generation of participants listened to those transmissions, discussed it (or did not discuss it) with another listener, and transmitted the terminal story into a second audio recorder. *Figure 1* shows an information flow diagram. Accuracy was determined from all audio transmissions, and all participants rated the original actor for how good, self-centered (reverse scored), hurtful (reverse scored), sensitive, responsible, aggressive (reverse scored), and sorry he was.

The following thesis was tested: Accuracy is reduced with serial one-way transmission. Peripheral information is much more likely to be lost than central information. This reduction is

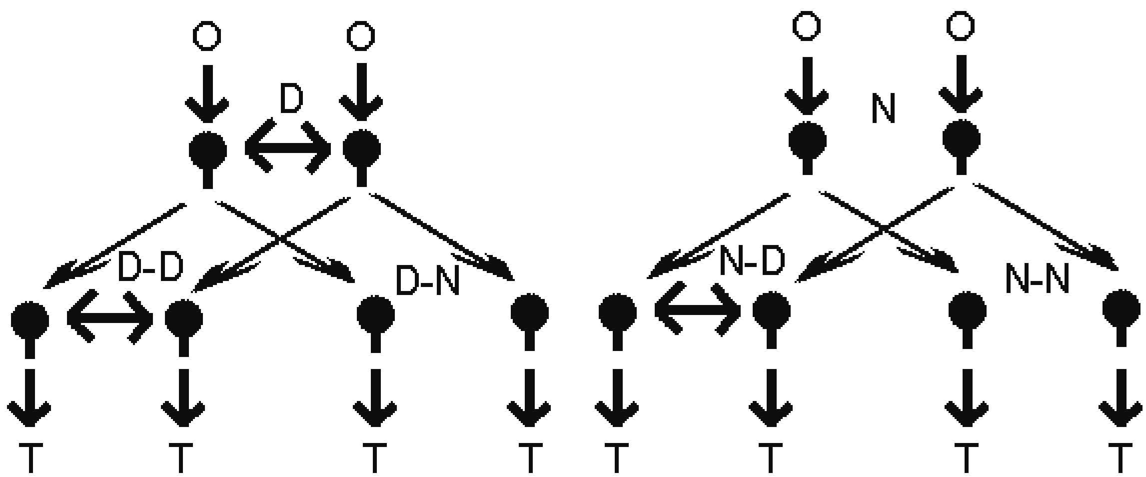

Figure 1. Information flow diagram. Each ball on a stick represents a person. Arrows represent directions of communication; double-headed arrows are discussions, while single-headed arrows are serial transmissions through audio recordings. O = original story. D = discussion. N = non-discussion. D-D = discussion following discussion. D-N = non-discussion following discussion. N-N = non-discussion following non-discussion. N-D = discussion following non-discussion. T = terminal report.

attenuated by discussion. When discussion has not occurred in the first-generation (i.e., significant amounts of information are already lost), discussion further decreases the accuracy of the account compared to no discussion. Transmissions that are more organized allow more peripheral information to be retained than allowed by more poorly organized transmissions. Finally, ratings of the actor are tied to the amount of peripheral information retained.

Hypotheses. To test this thesis, a number of hypotheses about accuracy and ratings of the original storyteller were examined (shown graphically in **Error! Reference source not found.**). The stories presented to participants were negative. Therefore, when participants were predicted to have "more neutral" ratings of an actor, they were predicted to rate the actor higher, as in Baron et al.'s (1997) research.

1. First-generation discussers were predicted to transmit more accurate accounts (and have more neutral, i.e., higher, ratings of the actor) than first-generation non-discussers, as shown by co-witness research. When two witnesses discussed an event, both were predicted to be better able to accurately transmit the event later.
2. Second-generation participants following first-generation discussion were predicted to transmit more accurate accounts (and have more neutral, i.e., higher, ratings of the actor) than second-generation participants following first-generation non-discussion. This was supported by Baron et al.'s (1997) research that showed more complete accounts were better remembered than less complete accounts.
3. After first-generation discussion, second-generation discussers were predicted to transmit more accurate accounts (and have more neutral, i.e., higher, ratings of the actor) than second-generation non-discussers. This extended from hypothesis 2: Discussion after discussion was predicted to be better than non-discussion after discussion.

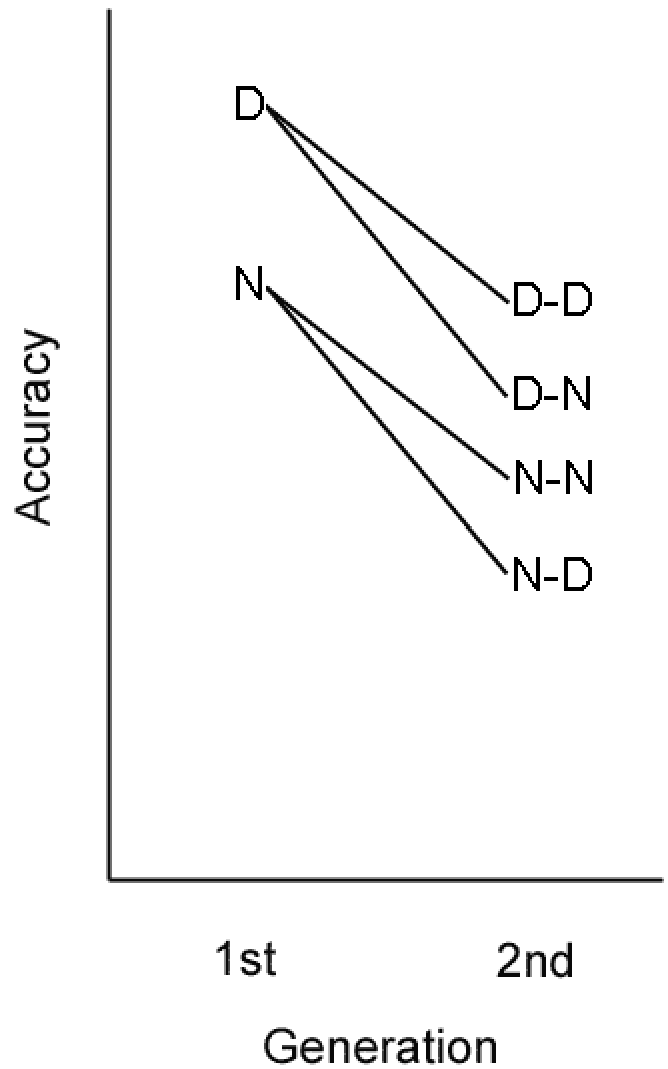

Figure 2. Predictions of transmission accuracy for each group of storytellers. D = discussion. N = non-discussion. D-D = discussion following discussion. D-N = non-discussion following discussion. N-N = non-discussion following non-discussion. N-D = discussion following non-discussion.

4. After first-generation non-discussion, second-generation non-discussers were predicted to transmit more accurate accounts (and have more neutral, i.e., higher, ratings of the actor) than second-generation discussers. This was shown by first-generation accounts without discussion losing most of the peripheral information (Baron et al., 1997) and Stasser and Titus's (1985) research which showed that participants only talk about information both

participants know. Since the pair of second-generation discussers would only share central information, they would further lose whatever peripheral information had been transmitted.

Additionally, the following hypotheses about pooled accuracy will be examined:

5. Pooled responses from first-generation discussers were predicted to be more accurate than pooled responses from first-generation non-discussers. While Andersson and Rönnberg (1995) found that pooled individuals were better than groups, they also found that individual responses from individuals who had been part of a group produced more complete responses than individuals who had not been part of a group. Therefore, pooled responses from individuals who had been part of a group should have been greater than pooled responses from individuals who had not been part of a group.
6. Pooled responses from second-generation discussers following first-generation discussion were predicted to be more accurate than pooled responses from second-generation non-discussers following first-generation non-discussion. This would show that the discussion system results in more accurate responses than the non-discussion system.
7. Pooled responses from second-generation discussers following first-generation discussion were predicted to be more accurate than pooled responses from second-generation non-discussers following first-generation discussion. This was a combination of hypotheses 3 and 5.

The Central-Peripheral Bias (CPB) is a statistic that measures the bias between the number of central statements and the number of peripheral statements in a story. The higher the CPB, the more central statements (compared to peripheral statements) were transmitted in a story. The following hypotheses regarding the CPB will be examined:

8. First-generation discussers were predicted to have lower CPB scores than first-generation

non-discussers (i.e., more peripheral information will be retained). This peripheral retention was predicted to be result of discussion helping to increase organization.

9. Second-generation participants following first-generation discussion were predicted to have lower CPB scores than second-generation participants following first-generation non-discussion. That is, second–generation participants were predicted to receive more peripheral information if they followed first-generation discussion.
10. After first-generation discussion, second-generation discussers were predicted to have lower CPB scores than second-generation non-discussers. After more accurate accounts were transmitted, further discussion was predicted to help retain more peripheral information.
11. After first-generation non-discussion, second-generation non-discussers were predicted to have lower CPB than second-generation discussion. After inaccurate accounts with little peripheral information were transmitted, discussion was predicted to further lose peripheral information.
12. An individual's CPB was predicted to be negatively correlated with his or her rating of the actor.

Further, the following prediction regarding statement ordering is made:

13. First-generation accounts were predicted to be highly organized and it was predicted that there would not be a difference between discussers and non-discussers (cf. Weldon & Bellinger, 1997).

And about the effects of order on CPB:

14. The more closely a first-generation account followed the order of the original story, the lower the CPB scores of the subsequent second-generation participants were predicted to be. Baron et al. (1997) showed that decreased organization caused peripheral information

loss, so the more organized a transmission was, the more peripheral information should have been retained.

Applications

There are two primary areas of application for this research: rumors (within the larger framework of informal communication) and eyewitness accounts.

Rumors. Experimental rumor research has most often used serial chain transmission designs (e.g., Allport & Postman, 1945; 1946; Baron et al., 1997; Gilovich, 1987; Inman et al., 1993; Lyons & Kashima, 2003; 2006). These experiments have been criticized for their lack of external validity and realism (Peterson & Gist, 1951). Discussion as transmission experiments have occurred (e.g., McAdam, 1962), but are rare. To the author's knowledge, discussion before and after transmission has never been examined. By adding the discussion dimension to the serial chain transmission design, realism and external validity were increased.

Eyewitnesses. It is standard legal procedure (Australia: Paterson & Kemp, 2005; United States of America: Stephenson, 1990; England: Warnick & Sanders, 1980) and scientific recommendation (e.g., Wright, Memon, Skagerberg, & Gabbert, 2009) to discourage witness collaboration. Nevertheless, most collaborative witness studies examine the effects of introducing misinformation via a collaborative witness (e.g., Gabbert, Memon, & Allan, 2003; Gabbert, Memon, Allan, & Wright, 2004; Hope et al., 2008; Kanematsu, Mori, & Mori, 2003; Luus & Wells, 1994; Paterson & Kemp, 2006a; Skagerberg, 2007; Wright et al., 2000). As pointed out by Paterson and Kemp (2005), the beneficial effects of discussion (as shown by Warnick & Sanders, 1980) have been ignored while focusing on misinformation.

Since the current research gave all participants the same (i.e., accurate) information in the first-generation, the effects of discussion versus independent report can be examined without being contaminated by the introduction of misinformation. Further, co-witness versus independent witness effects on a jury (i.e., second-generation discussers) can be examined.

Method

Participants

Participants were 181 RIT students, some of whom participated to earn extra credit in psychology undergraduate courses; the rest received monetary compensation. All participants were hearing (otherwise, an interpreter would add an intermediary teller/listener) and spoke English as a first language (since internal translation would increase attentional load, thereby decreasing stereotype-inconsistent information recall further; Baron et al., 1997) and did not have a heavy accent (since the attentional load of subsequent listeners would be increased; as determined by the experimenter). One participant (a non-discusser) was unclear on their audio transmission and replaced in the rumor chain. This participant was excluded from all analyses. The analyzed participants consisted of 106 males. They had a mean age of 20.74 years (*SD* = 2.747), and a mean college year level of 3.16 years (*SD* = 1.468).

Design

Independent variables. There were three independent variables. The number of eyewitnesses (either 0 or 2) was the number of people involved who experienced the original stimulus. Discussion (yes or no) was whether or not the pair discussed. Finally, transmission source (original, from first-generation discussers, or from first-generation non-discussers) was from where the participants received their transmission. Number of eyewitnesses was confounded with transmission source, since 2 eyewitnesses meant the transmission source must be original, and 0 eyewitnesses meant the transmission source must not be original (see the configuration section below).

Dependent variables. Absolute accuracy and relative accuracy were the primary dependent variables. Absolute accuracy was measured as the proportion of statements in the original report remaining in the transmitted report. Relative accuracy was the proportion of statements in the previous generation's story (that were absolutely accurate) that remained in the transmitted report. To find absolute accuracy, one rater compared the audio recordings to a coding sheet with each statement. The statements included in the transmitted report represented the overall amount of accurate information transmitted. These sheets were then compared to each other to find the relative accuracy. Absolute accuracy was used for comparisons where the groups being compared had different parents. This tested the differences from one system to another (e.g., D-D vs. N-N). Relative accuracy was used for comparisons where the groups being compared had the same parent (e.g., D vs. N or D-D vs. D-N) because it reduced the variability of the parent.

For Pooled Accuracy (PA), coding sheets for pairs were combined without duplication to represent the overall amount of information transmitted from the pair (cf. Andersson & Rönnberg, 1995). The value was the number of statements (not a proportion).

Adding was measured as the proportion of statements in the transmitted report that were not in the original report. For example, "then he hit another car," was not a statement in the story at all, and would thus be considered added (assuming the participant mentioned hitting one car already). Statements that were changed in the terminal report (and thus not accurate) were not considered adding. A changed statement is something like, "all his friends told him not to drive," which is an obvious morph of the statement "because others at the party encouraged me [to drive]" (Table 1). Changed statements were not examined.

The statements in the transmitted report remaining from the original report were then broken into central and peripheral statements (according to the statement's original classification). The number of peripheral statements was subtracted from the number of central statements, and this number was divided by the total number of central and peripheral statements in the transmitted report. This value was the Central-Peripheral Bias (CPB). This value represented the amount of central information remaining compared to the amount of peripheral

Table 1

Original Story Given to Participants by Statement

Statement	Central or Peripheral	Order
Last night I got drunk at a party	C	1
My drinks were spiked with 180-proof alcohol without my knowledge	P	2
I have little experience with alcohol,	P	3
and I did not feel intoxicated.	P	4
When my friend and I left, I drove	C	5
because others at the party encouraged me.	P	6
I want it known that I would never intentionally drink and drive.	P	7
I ended up traveling much faster than the speed limit.	C	8
I ran a stop sign	C	9
and hit another car.	C	10
I was in shock because of the wreck	P	11
and my friend yelled at me to drive away.	P	12
We left the scene of the accident,	C	13
and a police car was chasing us.	C	14
My friend was yelling to go faster.	P	15
I sped up, making it a high-speed police chase.	C	16
I lost control and spun out in somebody’s lawn.	C	17
My friend got out of the car and ran.	P	18
An officer approached me and I punched him in the face.	C	19
I don’t usually get in fights,	P	20
but I wasn’t thinking clearly yet.	P	21
I was then hand-cuffed and taken to jail.	C	22
My parents were very upset when they came to bail me out.	C	23
My college has revoked my athletic scholarship.	C	24
This was the worst night of my life, and I would take it back if I could,	P	25
if only to prevent the harm I caused to the driver of the car I hit.	P	26

information remaining. Values could range from -1 (entirely peripheral statements) to 1 (entirely central statements). Higher values mean more central information, as compared to peripheral information, remained.

Statement Ordering (SO) was calculated by performing a Spearman's *rho* test on the order of statements in the transmitted statements of the first-generation compared to the original statement (cf. Weldon & Bellinger, 1997). That is, each statement in the original stimulus was given a sequential value based on temporal order. The same was done for transmitted reports (consisting of 26 statements). These values were then correlated with a Spearman's *rho* for ordinal data. The Spearman's *rho*'s of SO were then converted into Fisher *z*'s for better comparison.

The final dependent variable was the rating of the original actor by the participants. Ratings were made on 9-point Likert-type scales (scaled from 0 [not at all] to 8 [extremely]) for how good, self-centered (reverse scored), hurtful (reverse scored), sensitive, responsible, aggressive (reverse scored), and sorry was the initial storyteller. A scale reliability analysis showed high reliability for these scales (Cronbach's $\alpha = .629$), so these ratings were added to create a final value, from 0 (negative) to 56 (positive), of how the participants felt about the actor.

Rater reliability. One rater judged all stories for all components. To ensure the single rater was not biased, a test of inter-rater reliability was performed. Two additional raters rated three random (as chosen by Random.org; Haahr, 1998) stories per group (18 stories total). Inter-rater reliability was determined by calculating Cronbach's alpha on the presence of statements between the three raters (binary data); Cronbach's alpha was quite high ($\alpha = .919$, $n = 468$). Since the inter-rater reliability was high, only the primary rater's judgments were analyzed further.

Configuration. *Figure 1* shows the flow of information. First-generation participants listened to the original stimulus, discussed or not, then transmitted to the second-generation. Note that each first-generation participant made one transmission that went to each of two second-generation participants (one discusser and one non-discusser). This controlled the responses of the second-generation by giving each of the groups (discussing and non-discussing) the exact same source materials. This resulted in 6 conditions: first-generation discussion (D), first-generation non-discussion (N), second-generation discussion following first-generation discussion (D-D), second-generation non-discussion following first-generation discussion (D-N), and a similar pair of second-generation conditions that followed first-generation non-discussion (N-D and N-N).

This study examined the effects of number of eyewitness (2 levels), discussion (2 levels), and transmission source (3 levels). This was not a factorial design, because transmission source is confounded with number of eyewitness and because the second-generation is not independent from the first-generation.

Materials

Initial presentation of the drunken driving scenario was a recording read by an actor on a digital audio player. This single recording was a recreation of Baron et al.'s (1997) scenario of a student who drove while drunk (Table 1). In the account there are 13 central statements and 13 peripheral statements. Participants told their versions of the scenario into a digital audio recorder. Finally, each participant independently rated the original actor.

Procedure

Participants, either first- or second-generation, were brought into their own small rooms. Participants were instructed to listen to a story in the following way:

[You are about to hear a story about a college student.] You will then make an audio tape

> that will transmit the information you learned to another person. In other words, we need you to provide a description of this person's actions that we can give to someone else. Please try to tell an entertaining yet accurate story that communicates much information and holds the other person's interest. ... [Finally,] you will answer some questions about [this person]. (Words in brackets were changed from original; Baron et al., 1997, p. 829)

Lyons and Kashima (2006) found that intent to communicate versus memorization severely influenced how participants reproduced the information, and, since intent to communicate mimics real rumors better, those instructions were used.

Participants were then shown how to use the digital audio recorder to listen to the recorded story (either the original or a reproduction from a first-generation participant, depending on condition). A 2-minute distracter task of drawing a map of the RIT campus (cf. Lyons & Kashima, 2003; 2006) was then completed to prevent rehearsal (cf. Andersson & Rönnberg, 1995; Gabbert et al., 2006; Hope et al., 2008; Paterson & Kemp, 2006a; Wright et al., 2000).

Pairs in the discussion conditions were then brought into the same small room to discuss the recording together. They were instructed to work together to recall the story as accurately as possible from beginning to end. Participants in the non-discussion condition were instructed to “rehearse the story by yourself [and] to recall the story as accurately as possible from beginning to end” (cf. Weldon & Bellinger, 1997). Both groups worked for 5 minutes on this task.

Participants were then shown how to record using the digital audio recorder. They recorded their version of events as best they could into the audio recorder after hearing the first instructions again. Finally, participants rated the original actor on the above traits.

Results

Unless otherwise specified, all t-tests were planned independent samples t-tests. Effect size *r* calculated as in Rosenthal and Rosnow (2007). Effect size *r* was used because of the common use of the *r* statistic (such as for correlations).

Adding

As shown in *Figure 3*, the proportion of added statements was reliably near (and typically below) Lyons and Kashima's (2006) 8%. Further, a one-way ANOVA showed no difference across groups, $F(5, 174) = 0.487$, $p = .786$. Given these low values, adding was not explored further.

Analysis of Paired Data

Two participants who discussed were likely to have more homogenous stories, as were descendents of participants who discussed, so their accuracy scores cannot be considered

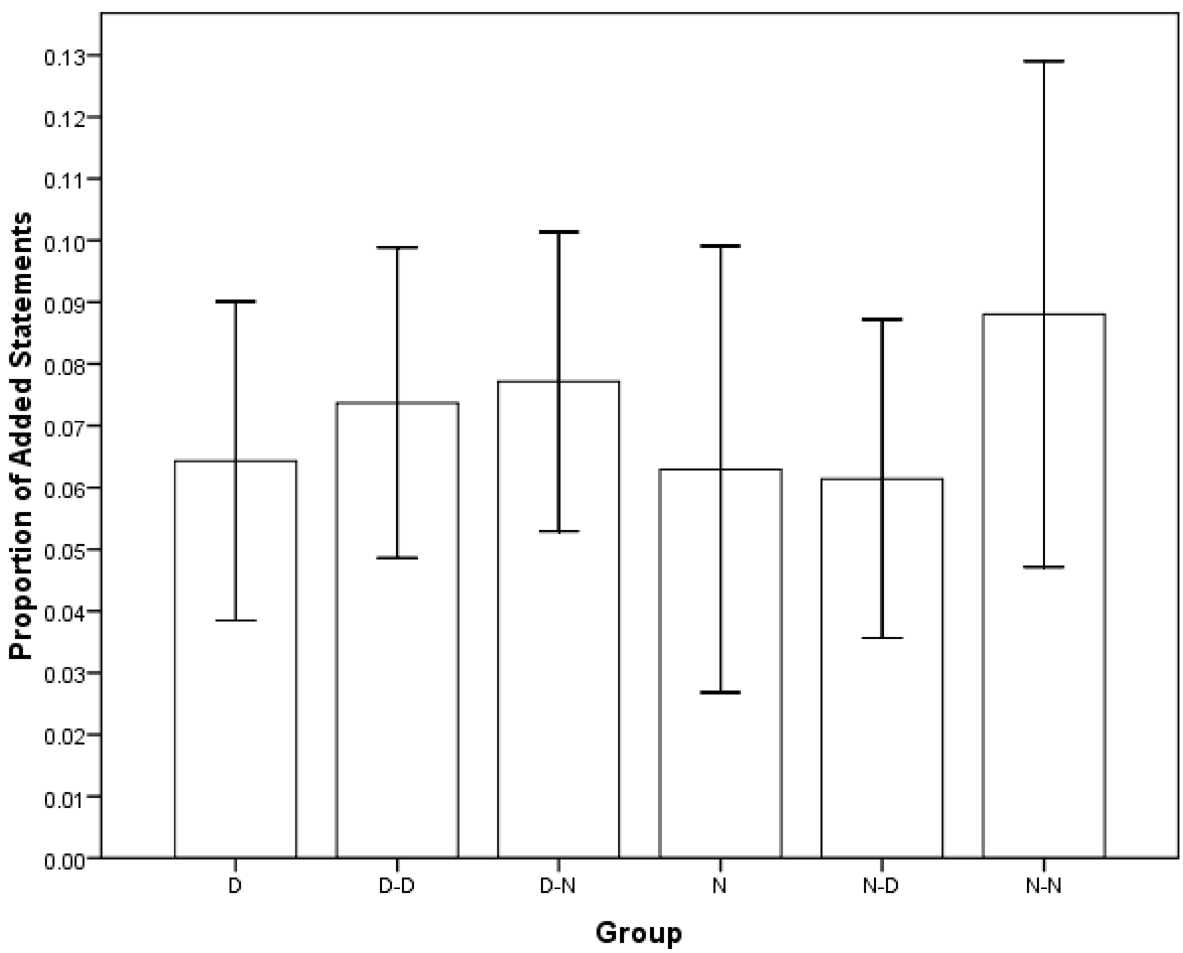

Figure 3. Mean proportion of statements in the transmitted account not in the original account (or changed from a statement in the original account) for each group. Error bars are 95% confidence intervals.

independent. Therefore, pairs who discussed had their absolute accuracy, relative accuracy, CPB, and SO Fisher *z* scores averaged into pair unit scores. CPB and SO Fisher *z* scores were not averaged for the testing of hypothesis 14 because of requirements for cross-generational analysis.

Accuracy

Two planned one-sample t-tests (against 1) of the first-generation and second-generation's relative accuracy checked that each generation always had worse accuracy than their preceding generation (the original story in the case of the first-generation); otherwise there would have been an obvious error in procedure or analysis (subsequent generations cannot gain accuracy beyond what they are told).

First-Generation Information Loss. A planned one-sample t-test (against 1) of the relative accuracy for the first-generation participants ($M = .621$, $SD = .145$) showed the first-generation reliably lost information from the original, $t = 20.249$, $df = 59$, $p < .0005$, effect size $r = .935$. This means, on average, 62.1% of the original story was retained (or, conversely, 37.9% was lost) in the first-generation.

Second-Generation Information Loss. A planned one-sample t-test (against 1) of the relative accuracy for the second-generation participants ($M = .809$, $SD = .187$) showed the second-generation reliably lost information from the first generation, $t = 9.695$, $df = 89$, $p < .0005$, effect size $r = .717$. This means, on average, 80.9% of the received story was retained (or, conversely, 19.1% was lost) in the second-generation. The second-generation retained more of their received story because they had less to remember: The first-generation was given 26 statements each, while the second-generation participants received, on average, about 16, a more manageable amount for their memories.

Hypotheses. Looking at **Error! Reference source not found.**, the accuracy hypotheses can be seen in graphical form. First-generation discussion (D) was predicted to result in the highest transmission accuracy (and actor rating) and the increased accuracy after discussion was predicted to allow second-generation discussers (D-D) to transmit more accurately than non-discussion. Transmission following first-generation non-discussion (N) was predicted to be less accurate than after first-generation discussion (D) and further discussion after non-discussion (N-D) was predicted to reduce accuracy further.

Table 2 shows the means (and *SD*s) of absolute accuracy and relative accuracy for each group. *Figure 4* shows the absolute accuracy results and *Figure 5* shows the relative accuracy results. Generally, the first-generation had better accuracy (higher absolute accuracy and relative

Table 2

Proportion of Statements Remaining in Transmitted Stories Compared to the Original (Absolute) or Received Story (Relative)

	Absolute Accuracy Proportion		Relative Accuracy Proportion	
Group	Mean	*SD*	Mean	*SD*
1st Generation Discussion[a]	.674	.118	.674	.118
2nd Generation Discussion[a]	.541	.129	.815	.158
2nd Generation Non-discussion[b]	.521	.149	.779	.188
1st Generation Non-discussion[b]	.568	.143	.568	.143
2nd Generation Discussion[a]	.500	.115	.925	.184
2nd Generation Non-discussion[b]	.435	.136	.779	.185

Note. The first-generation has identical absolute and relative accuracies since they received the original story.
[a] n = 15 (combined pairs) [b] n = 30

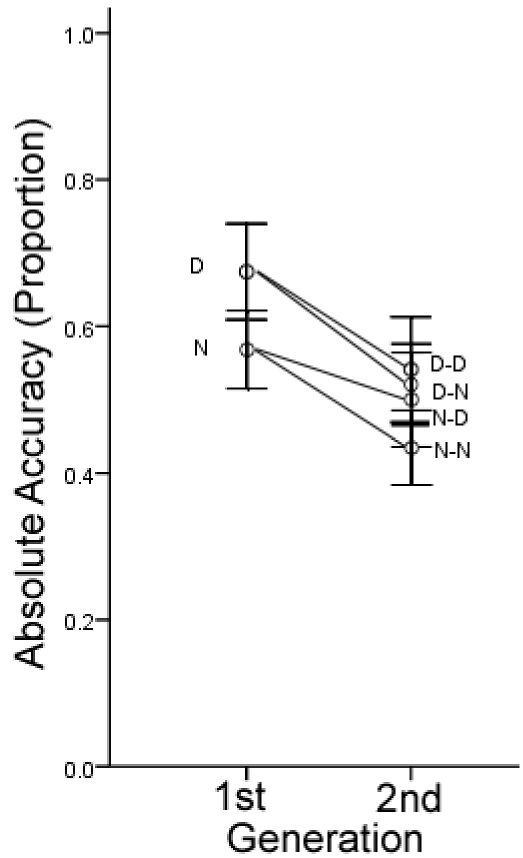

Figure 4. Proportion of statements in the original account remaining in transmitted account (absolute accuracy proportion) for each group. Error bars are 95% confidence intervals.

accuracy) and discussers had better accuracy.

Hypothesis 1. A t-test of the relative accuracy showed discussion in the first-generation (D; M = .674) reliably achieved better accuracy than non-discussion in the first-generation (N; M = .568), t = 2.487, df = 43, p = .008 (one-tailed), effect size r = .355. That is, first-generation discussers retained over 10% more statements compared to non-discussion.

Hypothesis 2. Participants following first-generation discussion (D-D/N; M = .527, SD = .142, n = 45) were more accurate than the participants that followed first-generation non-discussion (N-D/N; M = .456, SD = .132, n = 45), as shown by a t-test of the absolute accuracy, t = 2.461, df = 88, p = .008 (one-tailed), effect size r = .351. That is, participants following first-generation discussion retained about 7.1% more information (compared to the

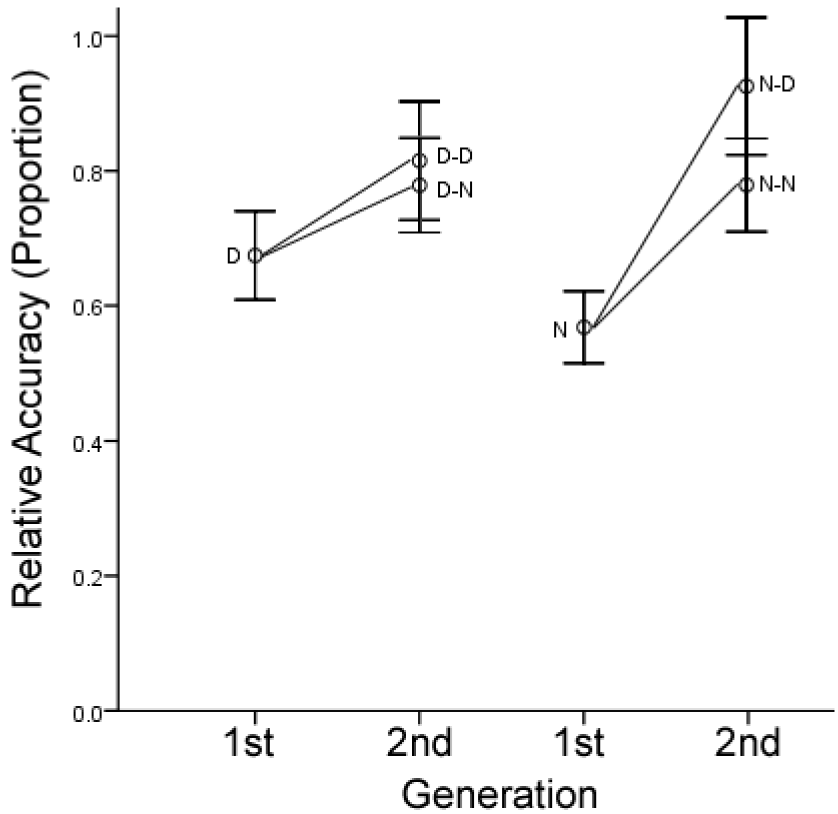

Figure 5. Proportion of statements from the received account remaining in the transmitted account (relative accuracy proportion) for each group. Error bars are 95% confidence intervals.

original) than participants following first-generation non-discussion.

Hypothesis 3. A relative accuracy t-test showed that, of participants following first-generation discussion, those who discussed (D-D; $M = .815$) were not more accurate than those who did not discuss (D-N; $M = .779$), $t = 0.643$, $df = 43$, $p = .262$ (one-tailed), effect size $r = .098$. That is, following first-generation discussion, participants who discussed were not significantly more accurate than those who did not, despite discussers retaining about 3.6% more information (compared to their parents).

Hypothesis 4. Of participants following first-generation non-discussion, those who discussed (N-D; $M = .925$) were not less accurate than those who did not discuss (N-N; $M = .779$), as shown by a t-test of the relative accuracy, $t = -2.509$, $df = 43$, $p = .992$ (one-tailed), effect size $r = .357$. In fact, the effect appears to be reversed. Following first-generation non-discussion, discussers retained about 14.6% more information than non-discussers (compared to their parents).

Cross-Parent Information Transfer

For discussing participants in the second-generation (D-D and N-D), the proportion of statements that came from the cross-parent (compared to the original's number of statements) were examined. The cross-parent was their discussion partner's parent. This exploratory measure examined how much information was transferred during discussion that could not have been gained another way. Differences between groups were tested with a t-test (none were predicted), and the amount of information transfer was compared to the differences between second-generation discussers vs. non-discussers.

Discussers following first-generation discussion (D-D; M = .036, SD = .042) had less cross-parent information transfer than discussers following first-generation non-discussion (N-D; M = .051, SD = .055). A t-test did not find a difference between the two groups, t = 1.227, df = 58, p = .225, effect size r = .159. That is, the 1.5% difference in cross-parent information transfer between groups (in favor of discussers following first-generation non-discussion) was not significant.

The difference of proportion averages between groups following first-generation discussion (.021 in favor of discussers) was lower than the amount of cross-parent information transfer for the discussers, indicating that the slight advantage of discussers (following first-generation discussion, i.e., D-D) can be entirely accounted for by cross-parent information transfer. The difference of proportion between groups following first-generation non-discussion (.065 in favor of discussers), however, was greater than the amount of cross-parent information transfer for the discussers, indicating that the advantage of discussers (following first-generation non-discussion, i.e., N-D) cannot be entirely explained by cross-parent information transfer.

Ratings

The actor rating scale items were reliable, as determined by Cronbach's alpha, α = .629. Therefore, the scores for each item were summed to create the total score for the scale. Table 3 shows the means (and *SD*s) of the ratings of the actor for each group. Generally, the first-generation and discussers had higher ratings of the actor.

Hypotheses. The actor rating hypotheses were exactly the same as the accuracy hypotheses (**Error! Reference source not found.**). First-generation discussion (D) was predicted to result in the highest transmission accuracy (and actor rating) and the increased accuracy after discussion was predicted to allow second-generation discussers (D-D) to transmit more accurately than non-discussion. Transmission following first-generation non-discussion (N) was predicted to be less accurate than after first-generation discussion (D) and further

Table 3

Rating of the Actor by Group

		Overall Rating[a]	
Group	*n*	Mean	*SD*
1st Generation Discussion[b]	15	31.40	6.591
2nd Generation Discussion[b]	15	24.63	4.962
2nd Generation Non-discussion[c]	30	22.40	7.877
1st Generation Non-discussion[c]	30	27.37	7.717
2nd Generation Discussion[b]	15	24.87	5.752
2nd Generation Non-discussion[c]	30	22.00	8.964

[a]Rating scale ranged from 0 to 56 (higher means actor was viewed as more "good") [b]$n = 15$ (combined pairs) [c]$n = 30$

discussion after non-discussion (N-D) was predicted to exacerbate the problem.

Hypothesis 1. A t-test of the actor rating showed participants who discussed in the first-generation (D; M = 31.4) reliably rated the actor higher than participants who did not discuss in the first-generation (N; M = 27.37), t = 1.731, df = 43, p = .045 (one-tailed), effect size r = .255. That is, in the first-generation, discussers rated the actor more positively than non-discussers.

Hypothesis 2. Participants in the second-generation following first-generation discussion (D-D/N; M = 23.14, SD = 7.061, n = 45) did not rate the actor significantly higher than the participants in the second-generation following first-generation non-discussion (N-D/N; M = 22.96, SD = 8.084, n = 45), as shown by a t-test of the actor rating, t = 0.118, df = 88, p = .453 (one-tailed), effect size r = .013. With a difference of only 0.18 points, there was no perceivable difference between participants in the second-generation following discussion versus non-discussion.

Hypothesis 3. An actor rating t-test showed that, of participants in the second-generation following first-generation discussion, the participants who discussed (D-D; M = 24.63) rated the actor higher than the participants who did not discuss (D-N; M = 22.40). Nevertheless, the amount was non-significant, t = 1.000, df = 43, p = .161 (one-tailed), effect size r = .151. The difference in actor rating of 2.23 points was not great enough to distinguish the ratings of discussers and non-discussers following first-generation discussion.

Hypothesis 4. Of participants in the second-generation following first-generation non-discussion, participants who discussed (N-D; M = 24.87) did not rate the actor lower than participants who did not discuss (N-N; M = 22.00). In fact, the effect appears to be reversed (see means in Table 3), as shown by a t-test of the actor rating, t = -1.125, df = 43, p = .867 (one-tailed), effect size r = .169, with discussers rating the actor more positively 2.87 points on the scale higher than non-discussers (both following first-generation non-discussion).

Pooled Accuracy

T-tests also examined these three hypotheses. Discussion groups were predicted to have more accurate transmitters than non-discussion groups as long as discussion occurred in the first-generation.

Table 4 shows the means (and *SD*s) of the Pooled Accuracy (PA) for each group. Generally, in the first-generation, discussion resulted in higher PA than non-discussion. In the second-generation, discussion resulted in lower PA than non-discussion.

Hypothesis 5. A t-test of the PA showed first-generation discussers (D; $M = 20.00$) had non-significantly higher PA than first-generation non-discussers (N; $M = 18.80$), $t = 1.241$, $df = 28$, $p = .112$ (one-tailed), effect size $r = .228$. Despite having, on average, 1.2 more unique statements in a pair, first-generation discussers did not have statistically more statements than first-generation non-discussers.

Table 4

Pooled Accuracy of Pairs by Group

Group[a]	Mean	*SD*
1st Generation Discussion	20.00	2.673
2nd Generation Discussion	16.40	3.247
2nd Generation Non-discussion	17.27	3.105
1st Generation Non-discussion	18.80	2.624
2nd Generation Discussion	15.80	3.144
2nd Generation Non-discussion	15.87	3.270

[a] *n* = 15 for each group. Each *n* represents 1 pair.

Hypothesis 6. Second-generation discussers following first-generation discussion (D-D; *M* = 16.40) did not have significantly higher PA than second-generation non-discussers following first-generation non-discussion (N-N; *M* = 15.87), as shown by a t-test of the PA, $t = 0.448$, $df = 28$, $p = .328$ (one-tailed), effect size $r = .084$. The 0.53 difference in statements was not enough to be statistically different.

Hypothesis 7. A PA t-test showed, of participants following first-generation discussion, second-generation discussers (D-D; *M* = 16.40) did not have significantly higher PA than second-generation non-discussers (D-N; *M* = 17.27), $t = -0.747$, $df = 28$, $p = .770$ (one-tailed), effect size $r = .140$. The 0.87 statement difference in favor of second-generation discussers was slightly (but non-significantly) reversed.

Central-Peripheral Bias

T-tests examined hypotheses 8 through 11 as well. Notice these hypotheses were the inverse of the four accuracy hypotheses. Hypothesis 12 was examined with a Pearson correlation between the Central-Peripheral Bias (CPB) and the actor rating of each participant.

Table 5 shows the CPB means (and *SD*s) for each group. Most of the groups have equivalent bias, but first-generation discussers (D) have less bias and second-generation non-discussers following non-discussion (N-N) have more bias.

Hypothesis 8. A t-test of the CPB showed participants who discussed during the first-generation (D; *M* = .241) had significantly lower CPB than first-generation non-discussers (N; *M* = .356), $t = 2.515$, $df = 40.23$, $p = .008$ (one-tailed), effect size $r = .369$ (adjusted to account for unequal variances, Levene’s test for equality of variances, $F(1,43) = 6.201$, $p = .017$). That is, participants in the first-generation who discussed retained more peripheral information than participants who did not discuss.

Table 5

Central-Peripheral Bias of Transmitted Stories by Group

Group	Mean	*SD*
1st Generation Discussion[a]	.241	.080
2nd Generation Discussion[a]	.382	.111
2nd Generation Non-discussion[b]	.386	.206
1st Generation Non-discussion[b]	.356	.224
2nd Generation Discussion[a]	.390	.120
2nd Generation Non-discussion[b]	.459	.264

Note. Central-Peripheral Bias = (Number of Central Statements – Number of Peripheral Statements) / (Number of Central Statements + Number of Peripheral Statements)
[a]*n* = 15 (combined pairs) [b]*n* = 30

Hypothesis 9. Participants who followed first-generation discussers (D-D/N; *M* = .384, *SD* = .179, *n* = 45) had lower CPB than participants who followed first-generation non-discussers (N-D/N; *M* = .436, *SD* = .228, *n* = 45), but the difference was not significant, as shown by a t-test of the CPB, *t* = 1.199, *df* = 88, *p* = .117 (one-tailed), effect size *r* = .127. The .052 difference in favor of participants following first-generation discussion (lower CPB means comparatively more peripheral retention) was not statistically significant.

Hypothesis 10. A CPB t-test showed, of participants following first-generation discussion, those who discussed (D-D; *M* = .382) had slightly lower CPB than those who did not discuss (D-N; *M* = .386), but the difference was not significant, *t* = 0.071, *df* = 43, *p* = .472 (one-tailed), effect size *r* = .011. The very slight difference in CPB between discussers and non-discussers (both following first-generation discussion) of .004 was not significant.

Hypothesis 11. Participants following first-generation non-discussion, those who discussed (N-D; $M = .390$) did not have higher CPB than those who did not discuss (N-N; $M = .459$), as shown by a t-test of the CPB, $t = -1.201$, $df = 42.799$, $p = .882$ (one-tailed), effect size $r = .181$. In fact, the effect appears to be reversed (see Table 5 for means), with the .069 difference in favor of discussers.

Hypothesis 12. CPB ($M = .379$, $SD = .206$, $n = 135$) was negatively correlated with actor rating ($M = 24.94$, $SD = 7.978$, $n = 135$), $r = -.277$, $p = .001$. This indicates, as central statements increase (compared to peripheral statements), actor rating goes down (*Figure 6*).

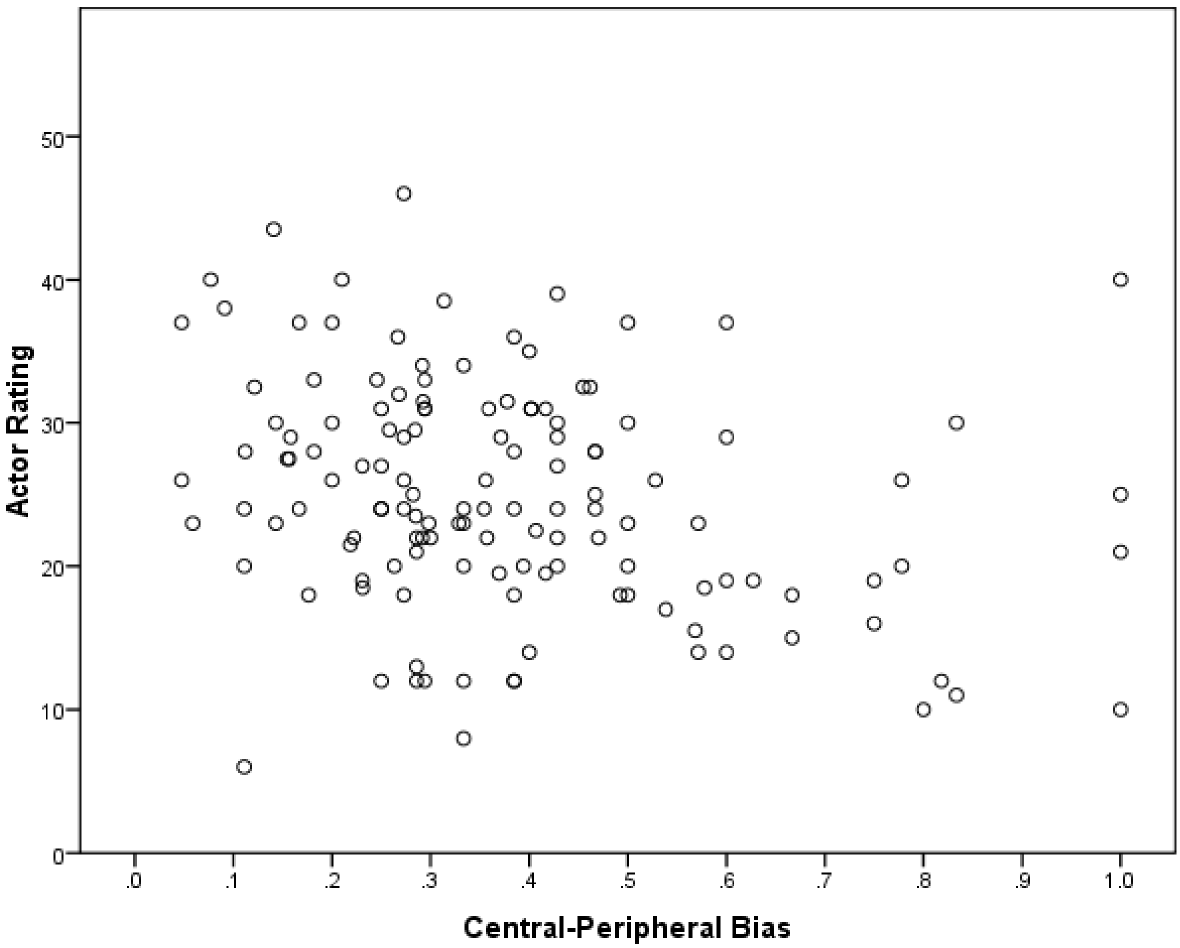

Figure 6. Central-Peripheral Bias (CPB) vs. actor rating for each participant (non-discussers) or pair (discussers). Higher CPB means more central bias relative to the total number of statements the participant relayed. Higher rating means the participant(s) thought the actor was more "good."

Statement Ordering

The Statement Ordering (SO) hypothesis was tested with a Spearman *rho* analysis of the first-generation's statement order compared to the statement order of the original stimulus. The Spearman's *rho*'s were converted into Fisher *z*'s for comparison. A t-test tested the differences between discussers and non-discussers.

Hypothesis 13. The Fisher *z* for the first-generation groups was quite high, 3.043 for discussers (SD = 0.489, n = 15, corresponding r = .995) and 2.854 for non-discussers (SD = 0.570, n = 25, corresponding r = .993). Some participants did not have Fisher *z* scores because their r = 1, causing a divide by 0 error when calculating Fisher *z*. For discussers (3 participants), null *z* values were ignored when averaging across pairs, i.e., the non-null participant's value was considered the average. For non-discussers (5 participants), null *z* values were ignored in this analysis. A t-test showed no difference between groups, t = 1.070, df = 38, p = .291, effect size r = .171. That is, both groups (first-generation discussion and non-discussion) had similar degrees of ordering agreement with the original story.

Statement Ordering and Central-Peripheral Bias

The Fisher *z* SO values were regressed against the corresponding CPB values of the second-generation to test hypothesis 14. Paired averages were not used because of the direct comparisons required between non-paired groups (e.g., N) to paired groups (e.g., N-D) required for regression. Further, first-generation data was duplicated so it could be compared to all of the second generation in one test. This was necessary because each first-generation participant past his or her story on to two participants in the second-generation (one discusser and one non-discusser).

Hypothesis 14. The SO Fisher *z* (M = 2.944, SD = 0.569, n = 104) and CPB (M = .404, SD = .198, n = 120) correlated significantly (r = -.322, p = .001, n = 104). Table 6 shows the

Table 6

Summary of Simultaneous Regression Analysis for First-Generation Statement Order Predicting Central-Peripheral Bias of Second-Generation (N = 104)

Variable	*B*	*SE B*	β
(Constant)	.713	.096	
Statement Order (Fisher *z*)	-.110	.032	-.322*

Note. R^2 = .104; F(1,102) = 11.789, p = .001
*p < .01

results of the regression analysis. SO of the first-generation accounted for about .104 of the variance for the CPB of the second-generation.

Discussion

Information Retention

Table 7 summarizes the results for the accuracy, actor rating, and CPB hypotheses. Two consistent findings emerge: discussion in the first-generation (D) was better than non-discussion in the first-generation (N) and, following first-generation non-discussion, discussion (N-D) was better than non-discussion (N-N). A third finding was hinted at, but statistically non-significant for actor rating and CPB: Second-generation participants following first-generation discussion (D-D/N) performed better than second-generation participants following first-generation non-discussion (N-D/N). “Better” meant higher accuracy, higher actor rating, and lower CPB.

Figure 7 shows a possible explanation. Looking at relative accuracy, second-generation participants performed equally well following first-generation discussion. Following first-generation non-discussion, however, discussion (N-D) recovered information that would otherwise have been lost without discussion (N-N). This recovery of information for the

Table 7

Summary of Accuracy, Actor Ratings, and Central-Peripheral Bias Hypotheses

Hypotheses[a]	Accuracy[b]	Ratings[c]	CPB[d]
D better than N	H1: *$D_{rel} > N_{rel}$	H1: *D > N	H8: *D < N
D-(D/N) better than N-(D/N)	H2: *$D\text{-}(D/N)_{abs} > N\text{-}(D/N)_{abs}$	H2: non-significant	H9: non-significant
D-D better than D-N	H3: non-significant	H3: non-significant	H10: non-significant
N-N better than N-D	H4: †$N\text{-}N_{rel} < N\text{-}D_{rel}$	H4: †N-N < N-D	H11: †N-N > N-D

Note. Hypotheses numbered (H1, H2, etc.) based on numbering in the hypothesis section
[a]The hypotheses for these three variables was patterned on what group was "better" than the other, better being defined by that specific variable. [b]Accuracy is either absolute (compared to original) or relative (compared to previous generation). Higher is always better. [c]Higher ratings were better than lower ratings (the actor was more "good"). [d]Lower CPB was better since it means more peripheral information was retained
*predicted †reversed

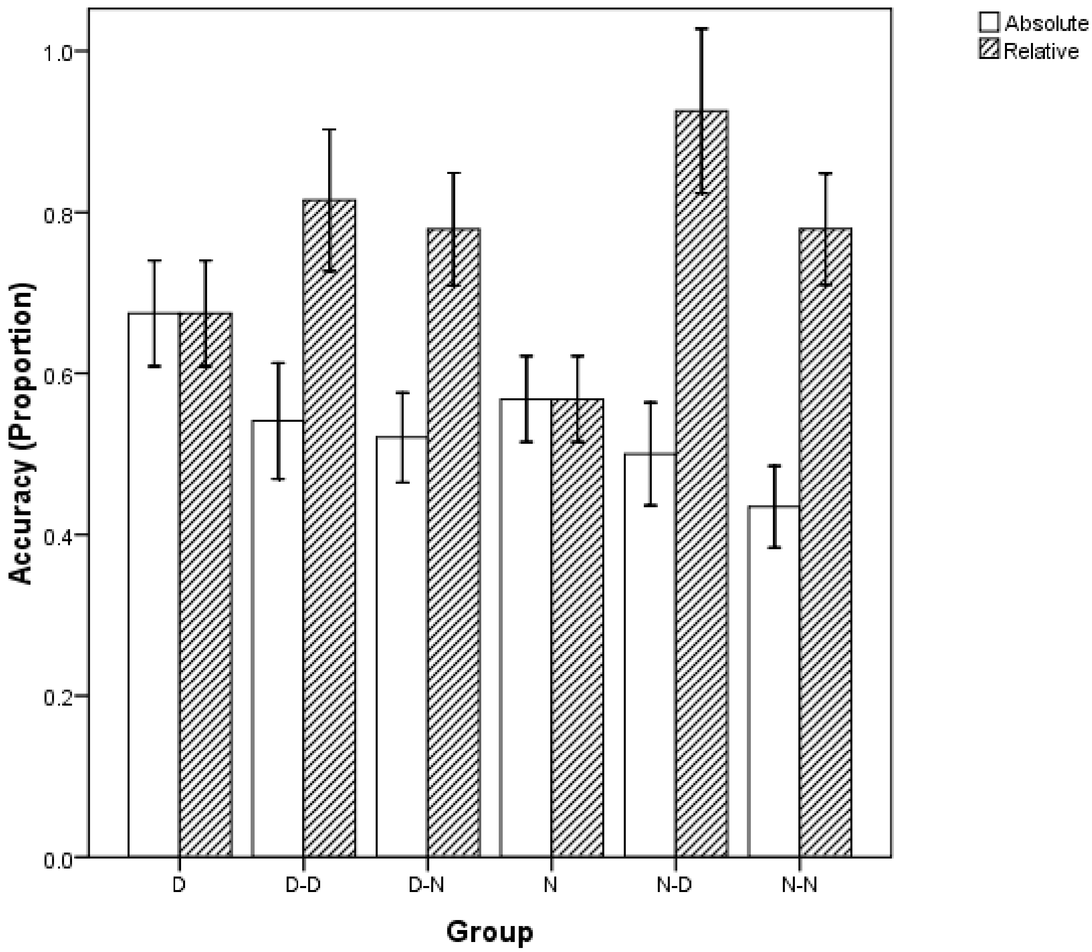

Figure 7. Proportion of statements from the original story in transmitted account (absolute accuracy) and proportion of statements from the received story in the transmitted account (relative accuracy) for each group. Error bars are 95% confidence intervals.

discussers following first-generation non-discussion shows discussion helps recover some of the available information (the pooled statements) from the two stories to offset loss in an individual story.

Pooled Accuracy. None of the PA tests of hypotheses showed differences between the PA of the groups (means in Table 4). This indicates, despite a lack of information transmitted individually, pairs generally transmitted a similar amount of information. This is likely because discussers tend to align their stories and transmit the same information, while non-discussers do not have that opportunity and transmit the information they, individually, found most important and remembered (which changes for each individual). This extends the finding of Stasser and Titus (1985; discussers talk about shared information) to include that discussers transmit shared information.

This means, as long as a person has access to reports from both eyewitnesses (including through another person), it does not matter if those eyewitnesses speak to each other or not. The end result is the same amount of total, non-redundant information.

Central-Peripheral Bias and Actor Ratings. Baron et al. (1997) showed a decrease in rating of the actor could be accounted for by a loss of peripheral statements (compared to central statements). This study replicated that finding (hypothesis 12), showing actor rating was significantly negatively correlated with CPB; as central statements increase (compared to peripheral statements) actor ratings go down.

Ordering

Just like Weldon and Bellinger (1997), SO was quite high ($r > .99$; hypothesis 13). Even with the problem of the ceiling effect, variation of SO in the first-generation was able to account for about 10% of the variability of the CPB in the second-generation (hypothesis 14). It is likely more variation in SO would result in a better correlation with CPB. This is in line with Baron et al.'s (1997) findings in Experiments 2 and 3.

Applications

Rumors. Previous serial chain transmission research was largely replicated. Significant leveling occurred (Table 2). Peripheral statements were leveled more than central statements (cf. Baron et al., 1997). Despite Peterson and Gist's (1951) assertion that adding occurs in more natural environments, there was a minimal amount of adding, consistent with Lyons and Kashima's (2006) findings.

This study indicates that people discussing rumors with other people who have heard the rumor will result in a more homogenous rumor set over time. Conversely, people only receiving a rumor and passing it on later without discussing the rumor with another person who has also heard the rumor will result in a more diverse rumor set over time. Specifically, peripheral information will be retained if there is discussion, but lost if there is not discussion.

Eyewitnesses. The assertion that co-witnesses should discuss was neither supported nor refuted. Looking at the total information that would be available (PA) to a jury (second-generation) from two witnesses, there was not a difference when the co-witnesses (first-generation) discussed or did not discuss (Table 4). This indicates the debate for allowing or prohibiting the discussion of eyewitnesses is something of a non-issue; whether co-witnesses discuss or not, the same total amount of information will be available to the jury. Of course, the participants in this instance were detached and independent from the event, a rare situation in court cases. An interesting study would use people familiar with each other and give them a stake in the "outcome" of their testimony.

Further, Stephenson's (1990) assertion that co-witness testimony omits peripheral information was not supported and, in fact, reversed (Table 5). Nevertheless, when looking at the actor rating (Table 3), juries that do not discuss judged the actor more harshly than juries who discussed. In light of this, juries should always be recommended to discuss a case and forego an initial vote. Otherwise, the jury may condemn a person they do not initially like (i.e., rate lower) instead of fully reviewing the evidence against that person.

Future Research

Discussion as transmission. A significant portion of external validity was ignored for this study: discussion as transmission. For the most part, information is not passed along via one-way transmission. Rather, when someone tells a rumor to someone else, the receiver has the opportunity to discuss and ask questions of the transmitter. In light of this, discussion as transmission (e.g., McAdam, 1962) should be explored along side of discussion before and after transmission.

Parents not from the same discussion pair. As shown in *Figure 1*, discussers in the second-generation following first-generation discussers (D-D) both had parents from the same discussion pair. Since discussing pairs have more homogenous stories, their pooled accuracy shows an amount of information similar to having non-discussing parents would be available to the second-generation discussers. By having parents from different discussion pairs, second-generation discussers would have access to stories with both greater overall accuracy and greater pooled accuracy.

Multiple parents. In the case of a jury trial, the N-D condition is not exactly analogous. In fact, the jurors all listen to all the witnesses, not just a single witness. To more accurately replicate this situation, the design could be modified to compare non-discussing participants with multiple parents each to discussing participants with only one parent each.

Longer chains. While Baron et al. (1997; experiment 1) did not find any differences in information loss when adding a third-generation, it is possible that the incorporation of discussion between generations has a compounding effect as distance from the source increases. For instance, it would be interesting to see if the children of second-generation discussers following first-generation non-discussion (both discussers and non-discussers) performed as well as the children of the second-generation following first-generation discussion (see Table 2) since those three groups currently have no significant difference in accuracy.

Discussion with a non-witness. The accuracy gains of second-generation discussers following non-discussion (N-D) over second-generation non-discussers (N-N) could not be completely explained by cross-parent information transfer. This means that, somehow, discussion helped people tell a more complete story without necessarily giving them more information. A study that has a participant discuss the story with a non-witness (someone not exposed to the original) could shed some light on whether the act of discussing, regardless of information transfer, helps someone to recover information they already know and tell a more complete story.

Manipulate Statement Order. Baron et al. (1997) found that experimentally disorganized accounts caused more extreme ratings of the actor. They also found evidence that extreme ratings of the actor were caused by a lack of peripheral information retention of the listener. Nevertheless, they did not directly connect SO and peripheral information retention. This study attempted to do so, but found that most transmitted stories tended to be ordered too close to the original. Despite the ceiling effect problem, SO accounted for 10.4% of the variability of later CPB (Table 6). It is likely this effect would increase with greater variability of SO.

Shaping. Although the information was available, this study did not examine statements that were merely changed, made no predictions regarding them, and counted them as omitted. Nevertheless, the average changed statement proportion across all participants was almost twice that of added statements ($M = .147$, $SD = .103$; Table 8 shows group breakdown). Further, a post-hoc one-way ANOVA found differences between the groups, $F(5, 174) = 2.844$, $p = .017$. In light of these preliminary findings, further research could examine how discussion affects shaping.

Viewing an event. In this study, participants were relayed a series of events from a source audio recording with a set number and type of statements. Showing a video of the events happening (which would be more externally valid for eyewitness research) would likely show a wider variety of responses from the first-generation and possibly spread apart accuracies and give more reliable results.

Table 8

Proportion of Transmitted Statements Substantially Changed from their Corresponding Statement in the Original Account by Group

Group[a]	Mean	*SD*
1st Generation Discussion	.101	.086
2nd Generation Discussion	.153	.083
2nd Generation Non-discussion	.178	.105
1st Generation Non-discussion	.119	.094
2nd Generation Discussion	.177	.103
2nd Generation Non-discussion	.156	.129

[a]*n* = 30 for each group.

Conclusion

Discussion increased accuracy of individual reports (except in the second-generation following first-generation discussion). Nevertheless, this advantage was dampened by looking at the full amount of information transmitted by a pair (their pooled accuracy) which showed no difference between discussers and non-discussers. This reduces validity of arguments both for and against eyewitnesses discussing with each other; it does not seem to affect the amount of information available to, or used by, the jury. Ordering of subsequent transmissions was always highly similar to the original and accounted for a significant portion of the Central-Peripheral Bias. Because of the ceiling effect of the highly consistent ordering, results are likely to be lower with more variable ordering.

References

Allport, G. W. & Postman, L. J. (1945). The basic psychology of rumor. *Transactions of the New York Academy of Sciences, 8*, 61-81.

Allport, G. W. & Postman, L. (1946). An analysis of rumor. *Public Opinion Quarterly, 10,* 501-517.

Andersson, J. & Rönnberg, J. (1995). Recall suffers from collaboration: Joint recall effects of friendship and task complexity. *Applied Cognitive Psychology*, *9*(3), 199-211.

Baron, R. S., David, J. P., Brunsman, B. M., & Inman, M. (1997). Why listeners hear less than they are told: Attentional load and the Teller-Listener Extremity Effect. *Journal of Personality and Social Psychology*, *72*(4), 826-838.

Basden, B. H., Basden, D. R., & Henry, S. (2000). Costs and benefits of collaborative remembering. *Applied Cognitive Psychology, 14*, 497-507.

DiFonzo, N. & Bordia, P. (2007). *Rumor psychology: Social and organizational approaches*. Washington, DC: American Psychological Association.

Gabbert, F., Memon, A., & Allan, K. (2003). Memory conformity: Can eyewitnesses influence each other's memories for an event? *Applied Cognitive Psychology, 17*, 533-543.

Gabbert, F., Memon, A., Allan, K., & Wright, D. B. (2004). Say it to my face: Examining the effects of socially encountered misinformation. *Legal and Criminological Psychology, 9*, 215-227.

Gabbert, F., Memon, A., & Wright, D. B. (2006). Memory conformity: Disentangling the steps toward influence during a discussion. *Psychonomic Bulletin & Review*, *13*(3), 480-485.

Gilbert, D. T., Pelham, B. W., & Krull, D. S. (1988). On cognitive busyness: When person perceivers meet persons perceived. *Journal of Personality and Social Psychology*, *54*(5), 733-740.

Gilovich, T. (1987). Secondhand information and social judgment. *Journal of Experimental Social Psychology*, *23*(1), 59-74.

Haahr, M. (1998). Random.org [Computer software]. Retrieved November 14, 2009, from http://www.random.org/integers/

Henningsen, D. D. & Henningsen, M. L. M. (2004). The effect of individual difference variables on information sharing in decision-making groups. *Human Communication Research*, *30*(4), 540-555.

Hope, L., Ost, J., Gabbert, F., Healey, S., & Lenton, E. (2008). 'With a little help from my friends...': The role of co-witness relationship in susceptibility to misinformation. *Acta Psychologica*, *127*(2), 476-484.

Inman, M. L., Reichl, A. J., & Baron, R. S. (1993). Do we tell less than we know or hear less than we are told? Exploring the teller-listener extremity effect. *Journal of Experimental Social Psychology*, *29*(6), 528-550.

Kanematsu H., Mori K., & Mori H. (2003). Memory distortion in eyewitness pairs who observed nonconforming events and discussed them. *Journal of Faculty of Education, Shinshu University, 109*, 75-84. (Original work published 1996, in Japanese.)

Leavitt, H. J. & Mueller, R. A. H. (1951). Some effects of feedback on communication. *Human Relations*, *4*, 401-410.

Luus, C. A. E. & Wells, G. L. (1994). The malleability of eyewitness confidence: Co-witness and perseverance effects. *Journal of Applied Psychology*, *79*(5), 714-723.

Lyons, A. & Kashima, Y. (2003). How are stereotypes maintained through communication? The influence of stereotype sharedness. *Journal of Personality and Social Psychology, 85*(6), 989-1005.

Lyons, A. & Kashima, Y. (2006). Maintaining stereotypes in communication: Investigating memory biases and coherence-seeking in storytelling, *Asian Journal of Social Psychology*, *9*, 59-71.

McAdam, J. R. (1962). *The effect of verbal interaction on the serial reproduction of rumor*. Unpublished doctoral dissertation, Indiana University.

Paterson, H. M., & Kemp, R. I. (2005). Co-witness discussion: A survey of police officers' attitudes, knowledge, and behaviour. *Psychiatry, Psychology and Law, 12*(2), 424-434.

Paterson, H. M., & Kemp, R. I. (2006a). Comparing methods of encountering post-event information: The power of co-witness suggestion. *Applied Cognitive Psychology*, *20*(8), 1083-1099.

Paterson, H. M., & Kemp, R. I. (2006b). Co-witnesses talk: A survey of eyewitness discussion. *Psychology, Crime & Law*, *12*(2), 181-191.

Peterson, W. A. & Gist, N. P. (1951). Rumor and public opinion. *American Journal of Sociology, 57*, 159-167.

Rosenthal, R. & Rosnow, R. L. (2007). *Essentials of behavioral research: Methods and data analysis*. New York: McGraw-Hill.

Skagerberg, E. M. (2007). Co-witness feedback in line-ups. *Applied Cognitive Psychology*, *21*(4), 489-497.

Stasser, G. & Titus, W. (1985). Pooling of unshared information in group decision making:

Biased information sampling during discussion. *Journal of Personality and Social Psychology*, *48*(6), 1467-1478.

Stephenson, G. M. (1990). Should collaborative testimony be permitted in courts of law? *The Criminal Law Review,* May, 302-314.

Warnick, D. H. & Sanders, G. S. (1980). The effects of group discussion on eyewitness accuracy. *Journal of Applied Social Psychology, 10*(3), 249-259.

Weldon, M. S., & Bellinger, K. D. (1997). Collective memory: Collaborative and individual processes in remembering. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*(5), 1160-1175.

Wright, D. B., Memon, A., Skagerberg, E. M., & Gabbert, F. (2009). When eyewitnesses talk. *Current Directions in Psychological Science, 18*(3), 174-178.

Wright, D. B., Self, G., & Justice, C. (2000). Memory conformity: Exploring misinformation effects when presented by another person. *British Journal of Psychology, 91*(2), 189-202.

Yuker, H. E. (1955). Group atmosphere and memory. *The Journal of Abnormal and Social Psychology*, *51*(1), 17-23.