The effects of question type and eyewitness temperament on accuracy and quantity of recall for a simulated misdemeanor crime

LAUREN R. SHAPIRO -- *Department of Psychology and Special Education, Emporia State University, Emporia, Kansas 66801 (lshapiro@emporia.edu)*

The present study examined the effect of question format on accuracy and quantity of testimony. Forty college students' memory for a videotaped theft was assessed through open-ended or multiple-choice questionnaires. High accuracy was found for central information elicited with an open-ended questionnaire and for peripheral information elicited with a multiple-choice questionnaire. Quantity was higher for central than for peripheral features and was higher with multiple-choice than with open-ended questionnaires. Lipton's (1977) notion of an inverse relationship between accuracy and quantity due to cognitive set was modified in order to incorporate the role question format had on different types of information. Several witness temperament characteristics were associated with central and peripheral recall using multiple-choice questionnaires.

Keywords: eyewitness testimony, individual differences, adult memory, simulated crime.

In the last three decades, researchers have investigated how the format of questions used to elicit eyewitness testimony affects both accuracy and quantity of information. The ability of the eyewitness to provide police officers and others in the judicial system with a clear understanding of what transpired during a crime is imperative for the prosecution of suspected perpetrators. When interviewing witnesses, a mixture of open-ended questions (i.e., requiring extensive responses) and closed-ended questions (i.e., requiring a simple yes/no response or a selection of two or more choices) is used to elicit a description of the suspect, information about the victim, and a description of the crime. Research findings typically indicate that overall recall from open-ended, unbiased questions (i.e., in which no answer is suggested), although less complete, is more accurate than overall recall from closed-ended questions, such as correct leading (i.e., in which a correct answer is suggested) and incorrect leading questions (i.e., in which an incorrect answer is suggested) (Dodd & Bradshaw, 1980; Loftus, Miller, & Burns, 1978; Smith & Ellsworth, 1987). Using leading questions inadvertently increases inaccuracy as at least one of the options provided will be wrong (e.g., "Did he wear a sweatshirt or jacket?") and the suggestion itself may interfere with the recall process.

The negative relationship between accuracy and quantity of overall recall is likely due to the fact that a witnessed event represents the type of complex-stimulus situation in which more information is learned than can be reported before availability of the information declines (Sperling, 1960). According to Lipton (1977), closed-ended questions impact on witnesses' cognitive set by limiting the range of responses to one of the given choices in multiple-choice questions or by restraining the answer to yes/no responses in correct/incorrect leading questions. For this question format, cognitive set is "narrowest" as witnesses are required to respond, even if they

do not know the answer, and should result in high quantity, but low accuracy. In contrast, open-ended questions widen witnesses' cognitive set as most questions lead to some response, although witnesses usually provide information only when they are sure of it. Consequently, this question format should result in low quantity, but high accuracy.

Although a number of investigators have studied the effect of questions on recall, their use of open-ended questions prior to closed-ended questions to elicit recall confounds type of question with accuracy and quantity. To disentangle these factors, different groups of witnesses must receive each type of question. Lipton (1977) showed college students a filmed murder-robbery and then administered either an open-ended questionnaire or a multiple-choice questionnaire. Consistent with his contentions, accuracy was higher for those given the open-ended rather than the multiple-choice questionnaire (i.e., 83% vs. 56%), whereas quantity was higher for those given the multiple-choice rather than the open-ended questionnaire (i.e., 75% vs. 32%). The current study attempted to replicate and extend Lipton's findings by examining accuracy and quantity for central and peripheral recall of a simulated misdemeanor theft. Shapiro, Blackford, and Chen (2005) admonished researchers not to rely on overall recall to determine whether witnesses remember a crime, as various types of information about crimes are not recalled equally well. Specifically, central information is more accurately recalled than peripheral information, particularly when open-ended questions are used (Cassel & Bjorklund, 1995; Clifford & Scott, 1978; Shapiro, Blackford & Chen, 2005). However, to elicit recall for peripheral information, specific cues in the form of unbiased and leading questions may be needed (Cassel, Roebers & Bjorklund, 1996). Thus, one modification to Lipton's contentions is that high accuracy would be expected for central information with open-ended questions and for peripheral information with multiple-choice

questions. A second modification would be that the high quantity expected for multiple-choice questions would favor central over peripheral information due to the use of specific cues to elicit salient features.

In addition to type of questions, individual differences may influence accuracy and quantity of testimony. Initial research examined the relationship between recall and personality characteristics, such as imaging, introversion/extroversion, and need for approval (Marks, 1972; Schill, 1966). However, more recent endeavors have explored the linkage between memory and temperament or "expression of behavior" (Thomas & Chess, 1977) as a means of understanding why witnesses vary in the accuracy and quantity of information reported about a crime. Ornstein, Shapiro, Clubb, Follmer, and Baker-Ward (1997) proposed that certain temperament characteristics affect eyewitnesses' perception and attention to an event as it unfolds (e.g., activity level, emotional intensity, persistence), whereas other dimensions (e.g., adaptability, approach/withdrawal, distractibility) impact on their adjustment to the interview context and hence the extent of their reports.

Results from suggestibility studies in the adult eyewitness literature support this contention. For example, Gudjonsson (1988) found that suggestibility was high in shy/avoidant and unassertive adults. Shapiro, Blackford and Chen (2005) reported that shy, highly active, or emotionally intense adults who were given incorrect leading suggestions demonstrated high levels of suggestibility for peripherally related crime features, whereas distractible, emotionally intense, or withdrawn adults produced high rates of suggestibility for the victim's appearance. High rates of suggestibility were also found in the suspect's appearance with shy or distractible adults and in bicycle features with non-persistent adults. In contrast, no published eyewitness studies and only one unpublished study (Palmer, Brandt, Chen, & Shapiro, 1998) have examined how temperament may affect recall for central and peripheral information elicited with open-ended questions. Palmer et al. found that easy-going witnesses who have irregular personal regimens demonstrated low recall levels for central features; whereas, difficult witnesses who are slow-to-adapt to new situations demonstrated low recall levels for peripheral details. In summary, temperament does seem to affect adults' encoding and retrieval of events and to mediate their responses to open-ended and incorrect leading questions.

The present research examined how the type of questions used to elicit recall for central and peripheral features of a crime affected accuracy and quantity, as well as the relationship between temperament and recall. College students were shown a videotape of a simulated crime and then given either an open-ended or multiple-choice questionnaire. Hypothesis 1a predicted that accuracy for central features would be higher for the open-ended group than for the multiple-choice group, whereas, Hypothesis 1b predicted that accuracy for peripheral features would be higher for the multiple-choice group than for the open-ended group. Hypothesis 2a predicted that quantity would be higher for the multiple-choice group than for the open-ended group, whereas, Hypothesis 2b predicted that quantity for central features would be higher than for peripheral features in both groups. Hypothesis 3 predicted that negative temperament traits (e.g., shy, distractible) would be associated with low accuracy and quantity.

METHOD
Participants

The participants were 40 undergraduate students, 21 of whom were men (*M* age = 20; 4 years, Mode age = 19; 8 years, Range = 18 to 23 years), from a small midwestern university. Students were recruited through sign-up sheets posted on the psychology experiment bulletin board and received research credit in exchange for their participation. Consistent with the makeup of the university, the participants were predominantly from middle-class, Caucasian families.

Materials

Stimulus — A VHS videotape of a trip to the zoo with an embedded theft scene was developed and used specifically for this project. The videotape was 12 minutes long and featured female adolescent twins who visited the zoo. There was a two-minute sequence in the beginning of the film in which the twins are witnesses to a bike theft. Despite several attempts to borrow a preadolescent boy's bike, an adolescent girl is repeatedly denied permission to use it. She leaves the scene, sneaks back, and then steals the bike. The appearance and behaviors of the adolescent perpetrator were congruent with stereotypically male rather than female characteristics (e.g., she had short hair, wore a black shirt, and punched the boy on the arm).

Memory questionnaires — Memory for the bike theft was assessed using either an *open-ended* or *multiple-choice* questionnaire. The assessment focused on three aspects of the bike-theft event: a) the bicycle characteristics, b) criminal actions, and c) the actors' physical characteristics and clothing. At the beginning of each questionnaire, there was a short paragraph explaining that police officers collect information from witnesses about crime.

The open-ended questionnaire consisted of two sections. In the first section, respondents were given direct, nonbiased probes to elicit information in particular categories (i.e., name, sex, facial characteristics, hair style and color, clothing,

height, weight, build, and age) for each actor (i.e., victim, perpetrator, other) in the event. In the second section, respondents were instructed to think about the bike theft and encouraged to provide as much detail as possible when providing testimony about the theft.

The multiple-choice questionnaire was organized using four categories (i.e., bicycle, actions, physical characteristics, and clothing). Within each category, there was a set of specific questions (e.g., "What was the color of the bike?") followed by three alternatives that included *correct leading* (e.g., "The bike was black."), *incorrect leading* (e.g., "The bike was red."), and *not sure* (e.g., "I don't remember.") choices. The order of correct and incorrect alternatives was counterbalanced, with not-sure answers always last. That is, one version of the questionnaire was organized with correct leading choices followed by incorrect leading choices and the other version was structured with incorrect leading choices followed by correct leading choices.

Temperament questionnaire — The New York Longitudinal Scales Adult Temperament Questionnaire (ATQ; Chess & Thomas, 1998) was used. This 54-item scale is divided into 9 subscales shown in Table 1 (see sample questions). For each question, respondents provided an estimation of how they behave in a given situation ranging using a Likert-type scale from 1 (almost never) to 7 (almost always). The internal consistency of the ATQ was measured using coefficient alpha reliabilities ranging from 0.69 to 0.83, with a median of 0.76.

Procedure

Each participant came to the laboratory and sat in a quiet room to watch the video under the pretense that the experimenter needed to finish setting up the experiment. After watching the film, the experimenter administered the ATQ and then gave the participant either the open-ended or multiple-choice questionnaire (determined by random assignment).

Scoring

Recall — Four types of proof are required in a court case to demonstrate theft of the bicycle occurred, including that the defendant took the bicycle, that it belonged to another person, that the defendant did not have permission to use the bicycle, and that the defendant performed the theft intentionally (Cassell, 1991). The 27 pre-identified features of the theft shown in Table 2 were organized into five categories that are central to proving a misdemeanor theft or are tangential to the crime. The central features consisted of central crime (CC), central appearance of the suspect (CA), and the stolen bicycle (BK) and are used to help establish guilt of the accused. The peripheral features included peripheral crime (PC) and peripheral appearance (PA) and serve to support the credibility of the witness.

Two types of scores—accuracy and quantity—were calculated. Central accuracy was computed as the total number of central

Table 1. Definition of temperament characteristics and sample questions.

1 Hardly ever	2 Infrequently	3 Once in a while	4 Sometimes	5 Often	6 Very often	7 Almost always	
Activity Level (the amount of physical motion or movement during the day). "I don't enjoy physically active sports. I would rather sit and watch such a game."							1 2 3 4 5 6 7
Rhythmicity (the regularity of bodily functions such as eating and sleeping). "I usually can't predict when I want to eat lunch or dinner, in contrast to the people who get hungry at the same time each day."							1 2 3 4 5 6 7
Adaptability (the ease with which a person adjust to changes in routine). "When I travel if the bus or train is very crowded and uncomfortable I get use to it after a few minutes or so."							1 2 3 4 5 6 7
Approach/ withdrawal (response to new things – curiosity vs. caution in dealing with the unfamiliar). "I like to go to a party or gathering where I meet new people."							1 2 3 4 5 6 7
Emotional intensity (the amount of energy expressed in response to situations, whether positive or negative). "If I'm criticized or treated unfairly, I show very little annoyance."							1 2 3 4 5 6 7
Mood (the quality of emotions expressed; either positive or negative in tone). "If someone does a stupid thing, I don't get annoyed, but can see the funny side of it."							1 2 3 4 5 6 7
Distractibility (the ease with which ongoing behavior is interrupted by unrelated sights and sounds). "No matter how many people are talking around me, I have no trouble concentrating on the people to who I am talking."							1 2 3 4 5 6 7
Persistence (the degree to which the person continues activities without interruption). "I don't like to concentrate on a job or project a long time. I prefer to take frequent breaks, even if I an enjoying the activity."							1 2 3 4 5 6 7
Sensory threshold (sensitivity to changes or differences in light, sound, taste, and texture). "When I travel by train or bus I don't mind if it is very warm or cold, even if others are complaining."							1 2 3 4 5 6 7

Table 2. Twenty-seven features in videotaped theft organized by categories.

Category	Central features
Central crime	1. Victim was owner of bicycle
	2. Suspect denied permission to use bicycle
	3. Suspect and victim physically struggle over bicycle
	4. Suspect steals bicycle
Bicycle	5. Color of frame
	6. Model
Central appearance (suspect)	7. Name
	8. Hair color and other defining characteristics
	9. Hair style
	10. Height: taller than victim
	11. Age: older than victim
	12. Clothing

Category	Peripheral features
Peripheral crime	1. Victim sitting on bench when suspect approached
	2. Suspect touched bicycle, started to wheel it away
	3. Victim got bicycle and moved it back to bench
	4. Suspect punches boy's arm
	5. Suspect made threatening gesture and remarks
	6. Suspect calls boy, "a stupid jerk"
	7. Victim angry when bicycle was stolen
	8. Victim's father came and comforted her
Peripheral appearance	9. Victim's name
	10. Victim's hair color, hairstyle, other defining characteristics
	11. Father's hair color, hairstyle, other defining characteristics
	12. Victim's clothes
	13. Victim's shoes
	14. Perpetrator's shoes
	15. Perpetrator's watch

features mentioned (i.e., CC, CA, BK) divided by the total number of accurate features mentioned. Peripheral accuracy was computed as the total number of peripheral features mentioned (i.e., PC, PA) divided by the total number of accurate features mentioned. Category accuracy was also calculated by dividing the number of features recalled for a category by the total of features recalled. For example, a participant may accurately recall 15 features of which 3 were

CC features, 3 were CA features, 1 was BK feature, 5 were PC features, and 3 were PA features. Central accuracy would be 7/15 or .47 and Peripheral accuracy would be 8/15 or 0.53. PC accuracy would be 5/15 or 0.33; CC Accuracy, CA accuracy, and PA accuracy would each be 3/15 or 0.20; and BK accuracy would be 1/15 or 0.07.

Central (peripheral) quantity was computed as the total number of central (peripheral) features mentioned divided by the total possible central (peripheral) features (i.e., 12 for central and 15 for peripheral). Category quantity was computed as the total number of features for a category divided by the total possible features. Using the example from above, Central quantity would be 7/12 or 0.58 and Peripheral quantity would be 8/15 or 0.53. CC and CA quantity would each be 3/12 or 0.25; BK would be 1/12 or .08; PC quantity would be 5/15 or 0.33; and PA quantity would be 3/15 or 0.20.

Temperament — A program was created in *Microsoft Excel* to categorize the Likert ratings on the ATQ according to the nine subscales and convert them to standardized scores. High scores on the ATQ are interpreted as negative aspects of the characteristics, specifically inactive, habitually rigid, slow-to-adapt, environmentally insensitive, shy/withdrawing, distractible, emotionally intense, non-persistent, and serious.

RESULTS
Recall

How does the question format affect accuracy and quantity of the type of information recalled? In the first analysis, central accuracy and peripheral accuracy were subjected to a 2 X 2 (Condition X Type of Feature) mixed model Repeated Analysis of Variance (ANOVA). In the second analysis, category accuracy was subjected to a 2 X 5 (Condition X Category) mixed model Repeated ANOVA. In the third analysis, central quantity and peripheral quantity were subjected to a 2 X 2 (Condition X Type of Feature) mixed model Repeated ANOVA. In the fourth analysis, category quantity was subjected to a 2 X 5 (Condition X Category) mixed model Repeated ANOVA. Condition (open-ended and multiple-choice) served as the between-subjects factor in all analyses, whereas type of feature (central vs. peripheral) or category (CC, BK, CA, PC, PA) served as the within-subjects factor. Geisser-Greenhouse corrections were applied to all effects that involved repeated measures and the effects reported were significant even with this correction. Tukey's HSD post-hoc tests or t-tests were used to examine all significant effects ($p < 0.05$).

Central and peripheral accuracy — A main effect of type of feature was interpreted within a significant Condition X Type of Feature interaction, $F(1,36) = 17.72$, $p < 0.001$, $\eta^2 = 0.33$. As shown on the left side of Figure 1, central accuracy
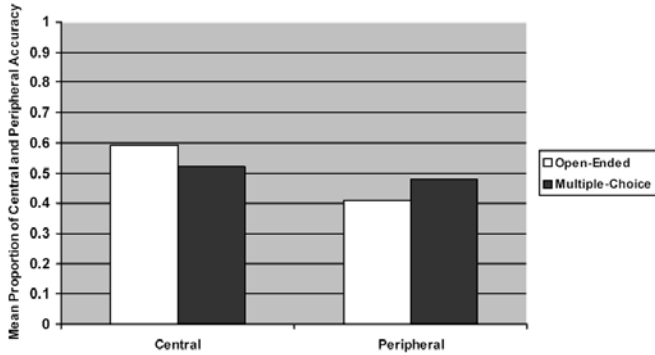
Figure 1. Mean proportions of central and peripheral accuracy by condition.

Figure 3. Mean proportions of central and peripheral quantity by condition.

was higher for witnesses in the open-ended group than for those in the multiple-choice group. In contrast, the right side shows peripheral accuracy was higher for witnesses in the multiple-choice group than for those in the open-ended group. A feature effect demonstrated that central accuracy was higher than peripheral accuracy for witnesses in the open-ended group, but not for those in the multiple-choice group.

Category accuracy — A main effect for category was interpreted within the significant Condition X Category interaction, $F(4,152) = 8.17$, $p < .001$, $\eta^2 = 0.18$. As shown in Figure 2, CC and CA accuracy was higher for witnesses in the open-ended group than for those in the multiple-choice group. In contrast, BK and PA accuracy was higher for witnesses in the multiple-choice group than for those in the open-ended group.

Central and peripheral quantity — Main effects of condition and of feature were interpreted within a significant Condition X Feature interaction, $F(1,38) = 7.35$, $p < 0.02$, $\eta^2 = 0.16$. As shown in Figure 3, central quantity (shown on the left side) and peripheral quantity (shown on the right side) were higher for witnesses in the multiple-choice group than for those in the open-ended group. Both groups also provided higher central quantity than peripheral quantity.

Category quantity — Main effects of condition and of category were interpreted within a significant Condition X Category interaction, $F(4,152) = 5.64$, $p < 0.001$, $\eta^2 = 0.13$. As shown in Figure 4, CC, BK, PC, and PA quantity was higher for witnesses in the multiple-choice group than for those in the open-ended group.

Relationship between temperament characteristics and recall

To what extent are temperament characteristics related to accuracy and quantity of the type of information recalled? A series of Pearson product-moment correlation analyses using temperament characteristics and the accuracy/quantity measures were performed for the open-ended and multiple-choice groups separately.

Central and peripheral accuracy — Witnesses in the multiple-choice group who are slow-to-adapt to new situations, $r(20) = -0.49$, $p < 0.03$, were associated with low
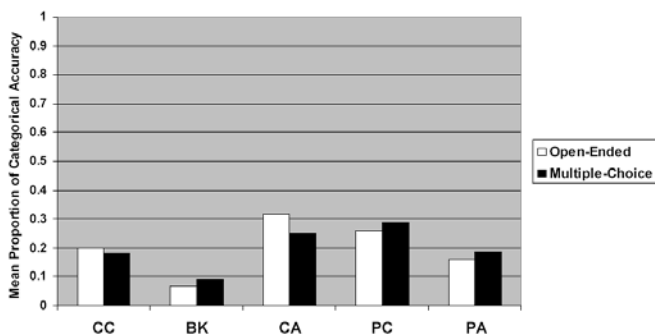

Figure 2. Mean proportions of category accuracy by condition. CC - central crime, BK - bicycle, CA - central appearance, PC - peripheral crime, PA - peripheral appearance.
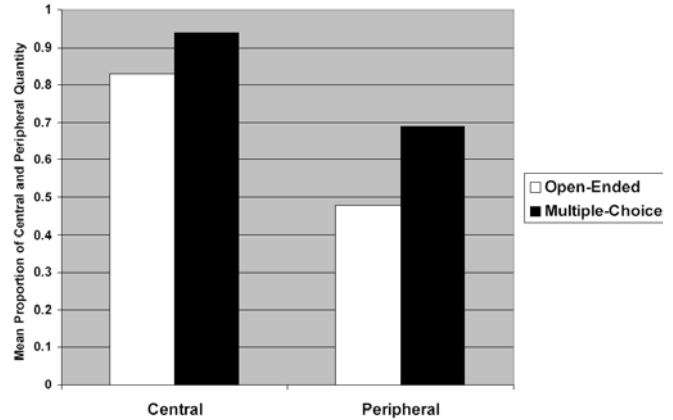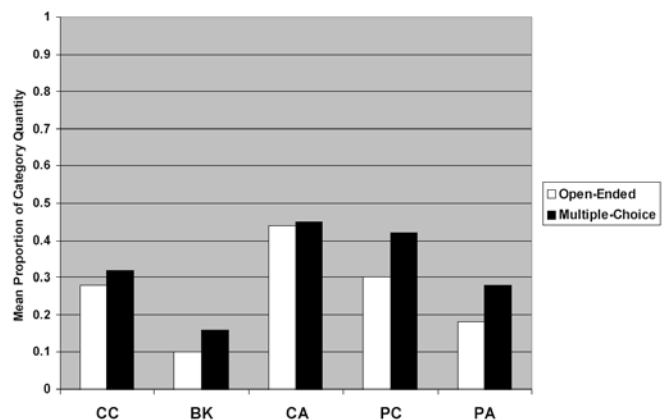

Figure 4. Mean proportions of category quantity by condition. CC - central crime, BK - bicycle, CA - central appearance, PC - peripheral crime, PA - peripheral appearance.

peripheral accuracy. Witnesses in the multiple-choice group who have a serious world-view were associated with high central accuracy, $r(20) = 0.63$, $p < 0.01$, but low peripheral accuracy, $r(20) = -0.60$, $p < 0.01$. None of the correlations for witnesses in the open-ended group were significant.

Category accuracy — Witnesses in the multiple-choice group who are shy were associated with high accuracy for central appearance of the suspect, $r(20) = 0.56$, $p < 0.02$; whereas, those who are emotionally intense were associated with high accuracy for bicycle, $r(20) = 0.64$, $p < 0.01$. Witnesses in the forced group who have a serious world-view were associated with high accuracy for bicycle, $r(20) = 0.52$, $p < 0.02$, and for peripheral crime, $r(20) = 0.53$, $p < 0.02$, but low accuracy for peripheral appearance, $r(20) = -0.72$, $p < 0.001$. One puzzling correlation was significant for the open-ended group. Witnesses who are highly distractible were associated with high accuracy for peripheral appearance, $r(20) = 0.45$, $p < 0.05$. No other correlations were significant.

Central and peripheral quantity — Witnesses in the multiple-choice group who have a serious world-view were associated with low peripheral quantity, $r(20) = -0.73$, $p < 0.001$. No other correlations were significant.

Category quantity — Witnesses in the multiple-choice group who are shy were associated with high quantity for central appearance of the suspect, $r(20) = 0.53$, $p < 0.02$, whereas, witnesses who have a serious world-view were associated with low quantity for peripheral appearance, $r(20) = -0.74$, $p < 0.001$. No other correlations were significant.

DISCUSSION

One question addressed by the current study was, "*How does the question format affect accuracy and quantity of the type of information recalled?*" The findings confirmed Hypothesis 1a in that central recall, particularly for the crime and appearance of the suspect, was more accurate for the open-ended group than for the multiple-choice group. There was also evidence supporting Hypothesis 1b in that peripheral recall, especially for appearance of the victim and her father, was more accurate for the multiple-choice group than for the open-ended group. The results provided support for Hypothesis 2a in that quantity for both types of features and for four of the five categories was higher for the multiple-choice group than for the open-ended group. Evidence also supported Hypothesis 2b given that quantity for central features was higher than for peripheral features in both groups.

A second question addressed by the present research was, "*To what extent are temperament characteristics related to accuracy and quantity of the type of information recalled?*"

Some support was found for Hypothesis 3 in that certain negative temperament traits (i.e., slow-to-adapt and serious world-view) were associated with low accuracy and quantity for peripheral features when multiple-choice questions were used. Palmer et al. (1998) also reported that slow-to-adapt witnesses demonstrated low recall of peripheral details. However, witnesses with other negative traits (i.e., shyness, emotionally intense) who were asked multiple-choice questions about central features were associated with high accuracy for suspect appearance and high quantity for both suspect appearance and bicycle. Although these traits typically limit attention, perception, and retrieval, it is possible that information about salient, but not secondary aspects of the crime, may have been recalled well because specific prompts cued witnesses' memory.

The findings have theoretical implications. This study replicated the inverse relationship between accuracy and quantity proposed by Lipton (1977). That is, the use of open-ended questions resulted in high accuracy with low quantity; whereas, the use of multiple-choice questions resulted in low accuracy with high quantity. This investigation also extended Lipton's findings by addressing the differential role that question format had on accuracy and quantity of central and peripheral information. Open-ended questions yielded high accuracy for central crime features; whereas, multiple-choice questions resulted in high accuracy for peripheral crime features. These results suggest that wide cognitive sets facilitate accurate responses when information is salient to the event, but not when information is secondary to the event. This idea is also consistent with "cognitive effort" in that it is easier to recognize rather than recall details and minor actions in a crime. The finding that multiple-choice questions resulted in high quantity for both central and peripheral crime features re-affirms the notion that narrow cognitive sets encourage eyewitnesses to provide responses, even at the risk of lowering the accuracy of their answers.

There are also applied implications of these findings. First, this study demonstrated the importance of examining central and peripheral recall rather than overall recall when considering accuracy and quantity. Police and others in the legal field should be aware that open-ended questions are best suited for eliciting accurate information about the crime needed to establish guilt of the suspect, but may not yield high quantity of peripheral information needed to build witness credibility. Second, this investigation suggested that certain negative eyewitness temperament traits may contribute to low accuracy and quantity of peripheral information about the crime and criminal when multiple-choice questions are asked. Finally, police may consider developing and using a standardized, open-ended questionnaire for particular crimes, such as theft, as a preliminary source of data collection at a crime scene. The benefit of using this type of questionnaire

would be that police could gather accurate information expeditiously while they perform other tasks, such as interviewing other witnesses and canvassing the area. In this way, eyewitnesses' memory for the crime and criminal is less likely to fade or to become tainted. The police could then increase the quantity of information obtained by using the questionnaire during a follow-up interview to probe for additional features.

REFERENCES

Cassel, W.S. (1991). *Child eyewitness testimony: The search for truth and justice in the American way.* Unpublished dissertation, Florida Atlantic University.

Cassel, W.S. and Bjorklund, D.F. (1995). Developmental patterns of eyewitness memory and suggestibility. *Law and Human Behavior, 19,* 507-32.

Cassel, W.S., Roebers, C.E., and Bjorklund, D.F. (1996). Developmental patterns of eyewitness responses to repeated and increasingly suggestive questions. *Journal of Experimental Child Psychology, 61*, 116-133.

Chess, S. and Thomas, A. (1998). *The New York Longitudinal Scales Adult Temperament Questionnaire Test Manual.* Scottsdale, Arizona: Behavioral-Developmental Initiatives.

Clifford, B.R. and Scott, J. (1978). Individual and situational factors in eyewitness memory. *Journal of Applied Psychology, 63,* 352-359.

Dodd, D.H. and Bradshaw, J.M. (1980). Leading questions and memory: Pragmatic constraints. *Journal of Verbal Learning and Verbal Behavior, 19,* 695-704.

Gudjonsson, G.J. (1988). Interrogative suggestibility: Its relationship with assertiveness, social-evaluative anxiety, state anxiety, and method of coping. *British Journal of Clinical Psychology, 27,* 159-166.

Lipton, J.P. (1977). On the psychology of eyewitness testimony. *Journal of Applied Psychology, 66*, 79-89.

Loftus, E.F., Miller, D.G. and Burns, H.J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory, 4,* 19-31.

Marks, D. (1972). Individual differences in the vividness of visual imagery and their Functions. In P.W. Sheehan (Eds.) *The Function and Nature of Imagery.* London: Academic Press.

Ornstein, P.A., Shapiro, L.R., Clubb, P.A., Follmer, A. and Baker-Ward, L. (1997). The influence of prior knowledge on children's memory for salient medical experiences. In N.L. Stein, P.A. Ornstein, B. Tversky, & C. Brainerd (Eds.), *Memory for everyday and emotional events* (pp. 83-111). Mahwah, NJ: Lawrence Erlbaum Associates.

Palmer, C., Brandt, C., Chen, C.-F. and Shapiro, L.R. (1998). *The relationship of temperament and memory in children.* Paper presented at the Great Plains Conference, Lincoln, Nebraska.

Schill, T. (1966). Effects of approval motivation and varying conditions of verbal reinforcement on incidental memory for faces. *Psychological Reports, 19,* 55-60.

Sperling, R. (1960). The information available in brief visual presentation. *Psychological Monographs, 74* (11, Whole No. 498).

Shapiro, L.R., Blackford, C. and Chen, C.-F. (2005). Eyewitness memory for a simulated misdemeanor crime: The role of age and temperament in suggestibility. *Applied Cognitive Psychology, 19,* 267-289.

Smith, V.L. and Ellsworth, P.C. (1987). The social psychology of eyewitness accuracy: Misleading questions and communicator expertise. *Journal of Applied Psychology, 72,* 294-300.

Thomas, A. and Chess, S.C. (1977). *Temperament and development.* New York: Brunner/Mazel.