Sex Differences in the Contributions of Appearance-Related Messages to Body Esteem and Perceptions of Thinness
Carrie L. Giant, Rhonda M. Passino, and Lesa Rae Vartanian
Indiana University–Purdue University Fort Wayne

Effects of High and Low Saturation of Red Hue on Long-Term Memory Performance
Ani G. Nikolova and Jill L. Quilici
California State University, Northridge

Effects of DNA and Eyewitness Evidence on Juror Decisions
Naomi J. Freeman and Diana L. D. Punzo
Earlham College

Part-Set Cuing is Due to Strong, Not Weak, List Cues
Monica L. Boice and Gary J. Gargano
Saint Joseph’s University

Human Sex Differences in Aggression Within an Evolutionary Model
Barbara J. Hagenah, Christopher Heaps, Eugene Gilden, and Michael Roberts
Linfield College

Published quarterly by Psi Chi, The National Honor Society in Psychology
Psi Chi is the National Honor Society in Psychology, founded in 1929 for the purposes of encouraging, stimulating, and maintaining excellence in scholarship, and advancing the science of psychology. Membership is open to graduate and undergraduate men and women who are making the study of psychology one of their major interests and who meet the minimum qualifications. Psi Chi is a member of the Association of College Honor Societies (ACHS) and is an affiliate of the American Psychological Association (APA) and the American Psychological Society (APS). Psi Chi's sister honor society is Psi Beta, the national honor society in psychology for community and junior colleges.

Psi Chi functions as a federation of chapters located at more than 990 senior colleges and universities. The Psi Chi National Office is located in Chattanooga, Tennessee. A National Council, composed of psychologists who are Psi Chi members and who are elected by the chapters, guides the affairs of the organization and sets policy with the approval of the chapters.

 Psi Chi serves two major goals—one immediate and visibly rewarding to the individual member, the other slower and more difficult to accomplish, but offering greater rewards in the long run. The first of these is the Society’s obligation to provide academic recognition to its inductees by the mere fact of membership. The second goal is the obligation of each of the Society’s local chapters to nurture the spark of that accomplishment by offering a climate congenial to its creative development. For example, the chapters make active attempts to nourish and stimulate professional growth through programs designed to augment and enhance the regular curriculum and to provide practical experience and fellowship through affiliation with the chapter. In addition, the national organization provides programs to help achieve these goals, including national and regional conventions held annually in conjunction with the psychological associations, research award and grant competitions, certificate recognition programs, national and regional chapter awards, and national service projects.

For more information about Psi Chi, contact the Psi Chi National Office, P.O. Box 709, Chattanooga, TN 37401-0709; telephone (423) 756-2044; www.psichi.org.

Journal Purpose Statement
The twofold purpose of the Psi Chi Journal of Undergraduate Research is to foster and reward the scholarly efforts of undergraduate psychology students as well as to provide them with a valuable learning experience. The articles published in this journal represent primarily the work of the undergraduate student(s). Faculty supervisors, who deserve recognition, are identified by an asterisk next to their name or on a separate byline.

Since the articles in this journal are primarily the work of undergraduate students, the reader should bear in mind that: (1) the studies are possibly less complex in design, scope, or sampling than professional publications and (2) the studies are not limited to significant findings. The basis for accepting papers for publication is the agreement among three professional reviewers that the project, hypothesis, and design are well researched and conceived for someone with an undergraduate level of competence and experience.

Instructions for Contributors
The Psi Chi Journal of Undergraduate Research encourages undergraduate students to submit manuscripts for consideration. Submissions are accepted for review on an ongoing basis. Although manuscripts are limited to empirical research, they may cover any topical area in the psychological sciences.

1. The primary author of a submitted manuscript must be an undergraduate student who is a member of Psi Chi. Manuscripts from graduate students will be accepted only if the work was completed as an undergraduate student. Additional authors other than the primary author may include non-Psi Chi students as well as the faculty mentor or supervisor. Membership verification information for the primary author must be included.

2. Only original manuscripts (not published or accepted for publication elsewhere) will be accepted.

3. All manuscripts must be prepared according to the Publication Manual of the American Psychological Association (5th ed.).

4. What to submit:
   a. Four copies of the complete manuscript. Near-letter-quality print is required on all copies. Should you desire a masked (blind) review, make sure that identifying names, affiliations, etc. appear only on the title page and nowhere else on the manuscript; i.e., manuscripts should be reasonably free of clues to the identity of the authors. Footnotes that identify the author(s) should appear on a separate page. You must request masked review.
   b. An e-mail address so that receipt of your manuscript can be acknowledged.
   c. A self-addressed envelope with sufficient postage for the return of your manuscripts when an editorial decision has been reached.

A sponsoring statement from the faculty supervisor who attests: (1) that the research adhered to APA ethical standards; (2) that the supervisor has read and critiqued the manuscript on content, method, APA style, grammar, and overall presentation; and (3) that the planning, execution, and writing of the manuscript represent primarily the work of the undergraduate student.

Submit all manuscripts to:
Dr. Stephen F. Davis, Managing Editor
Psi Chi Journal of Undergraduate Research
649 Hideaway Lane East
Lindale, TX 75771-5301

The Psi Chi Journal of Undergraduate Research (ISSN 1089-4136) is published quarterly in one volume per year by Psi Chi, Inc., The National Honor Society in Psychology, P.O. Box 709, Chattanooga, TN 37401-0709.

Subscriptions are available on a calendar-year basis only (Spring–Winter). U.S. rates are as follows (four issues): Individual $20; Institution $40. For international rates or other information contact: Daniel P. Bockert, Director of Publishing, Psi Chi National Office, P.O. Box 709, Chattanooga, TN 37401-0709; telephone (423) 756-2044; fax (toll-free) 1-877-774-2443; e-mail journal@psichi.org. Printed in the USA. Periodicals postage paid at Chattanooga, TN, and additional mailing offices.

Statements of fact or opinion are the responsibility of the authors alone and do not imply an opinion on the part of the officers or members of Psi Chi. Copyright 2001 by Psi Chi, The National Honor Society in Psychology. Postmaster: Send address changes to Psi Chi Journal of Undergraduate Research, P.O. Box 709, Chattanooga, TN 37401-0709.
Sex Differences in the Contributions of Appearance-Related Messages to Body Esteem and Perceptions of Thinness

Past research has indicated that satisfaction with one’s body shape is negatively correlated with mass media influence and appearance-related feedback from family and peers. The present research investigated sex differences in the contributions of those sociocultural factors to body esteem and perceptions of thinness. One hundred sixty-six women and 113 men completed questionnaires measuring body esteem, perceptions of current body shape compared to perceptions of desired and perceived ideal body shapes, extent of susceptibility to mass media model influence, frequency of teasing from family and peers, and frequency of body-related discussions with peers. Correlations among these variables indicated sex differences, and reinforced the importance of sociocultural influences on perceptions of the body for both men and women.

Carrie L. Giant
Rhonda M. Passino
Lesa Rae Vartanian*
Indiana University–Purdue University Fort Wayne

The occurrence of low body esteem seems to have increased over the past 20 years (Rieves & Cash, 1996). Low body esteem can lead to excessive dieting and exercise, which in turn can lead to chronic weight problems, growth disturbances, and eating disorders (Davis, Kennedy, Ravelski, & Dionne, 1994; Pugliese, Lifshitz, Grad, Fort, & Marks-Katz, 1983). In fact, as the prevalence of eating disorders is on the rise (Nagel & Jones, 1992), the “ideal” body portrayed in the media has been getting thinner (Wiseman, Gray, Mosimann, & Ahrens, 1992). Because of these patterns, researchers have been actively investigating factors that may contribute to low body esteem.

Most of the studies regarding body esteem and body image only include women as participants. One reason for the wide exclusion of men is that women usually have lower body esteem than do men (e.g., Grogan, Williams, & Conner, 1996; Miller et al., 2000) and perceive their own body shape as more dissimilar to their desired body shape than do men (McKinley, 1998; Silberstein, Striegel-Moore, Timko, & Rodin, 1988). Additionally, women outnumber men when it comes to eating disorders by a ratio of 10 to 1 (Crisp, 1980). For women in the United States, thinness has become almost synonymous with beauty (Thompson, 1990). Furthermore, women are more popular candidates for body esteem research than are men because researchers believe that women experience greater societal pressures to be a particular size and shape, which is believed to affect their body esteem negatively (Delaney, O’Keefe, & Skene, 1997; Ogletree, Williams, Raffeld, Mason, & Fricke, 1990; Rothblum, 1990; Silverstein, Perdue, Peterson, & Kelly, 1986).

Sociocultural factors are thought to be the strongest influences on the development of low body esteem (Thompson, 1992). Researchers have been investigating the influence of the mass media as a

Author note. Carrie L. Giant and Rhonda M. Passino, undergraduate students, Department of Psychology; Lesa Rae Vartanian, Assistant Professor, Department of Psychology.

Portions of this findings were presented at the 14th Biennial Conference on Human Development (Memphis, Tennessee, April 14–16, 2000), and the annual meeting of the Midwestern Psychological Association (Chicago, Illinois, May 4–6, 2000).

Correspondence concerning this article should be addressed to Lesa Rae Vartanian, PhD, Department of Psychology, Indiana University–Purdue University Fort Wayne, Neff Hall 380D, 2101 Coliseum Boulevard East, Fort Wayne, IN 46805. Electronic mail may be sent to VartaniL@ipfw.edu, and faxes may be sent to (219) 481-6972.
sociocultural factor that contributes to low body esteem. Societal messages portrayed in the mass media help to define physical attractiveness and unattractiveness for the general public (Cash & Pruzinsky, 1990). However, these societal messages often portray an ideal that is unachievable (Fallon, 1990). Research has shown that individuals compare their own bodies to mass media models (Grogan et al., 1996). Sixty percent of girls regularly read at least one magazine that they consider to be an important source for information about ideal shape, diet, fitness, and beauty (Levine, Smolak, & Hayden, 1994). In addition, 22% of girls who regularly read magazines report high interest in imitating the fashion models (Levine et al., 1994). Researchers posit that the mass media restricts the range of physical beauty, promotes slenderness as the ideal for beauty and fitness, promotes slenderness as the path to social, sexual, and occupational success, and portrays fat as negative (Levine & Smolak, 1996). In a recent study, women viewed photographs of “ideal” (i.e., slender, attractive) female models to determine if viewing the photographs had an effect on their self-evaluations (Henderson-King & Henderson-King, 1997). Results indicated that after viewing the models in the photographs, thinner women evaluated their sexual attractiveness more positively, whereas heavier women evaluated their attractiveness more negatively, particularly when their body shape was more dissimilar to the “ideal” body shape. Additional research indicated that women who viewed fashion magazines containing images of the thin ideal body immediately felt less satisfied with their own bodies and desired to be more thin, than did women who viewed newspapers (Turner, Hamilton, Jacobs, Angood, & Dwyer, 1997). Similar research has shown that women who view magazines that contain more images of the thin ideal body are more dissatisfied with their current bodies (Stice, Schupak-Neuberg, Shaw, & Stein, 1994).

Other factors that contribute to low body esteem among women are teasing about weight or body shape and other appearance-related feedback from family and friends (e.g., Levine et al., 1994). Negative appearance-related comments from family and friends are thought to be internalized and thus to relate to negative perceptions of one’s body. Appearance-related criticism as recalled during childhood is significantly related to increased dysfunctional body image in adulthood for women (Rieves & Cash, 1996). Women most often identified “peers in general” as the critics (62%), followed by friends (47%) and brothers (41%). In a similar study, researchers investigated three sociocultural influences—parents, peers, and magazines—as possible factors associated with girls’ body satisfaction (Levine et al., 1994). Results showed that direct parental pressure from mother and father to be thin correlated positively with weight concerns and dieting behaviors. Additionally, criticism from family members about weight or shape, body-related discussions with peers, and susceptibility to magazine model influence significantly contributed to body dissatisfaction and dieting behaviors. Results indicated that 42% of the girls reported talking with their friends about weight, shape, and dieting at least sometimes. In fact, some researchers suggest that peers may have greater influence on body esteem than mass media models, perhaps because verbal feedback from peers has a more direct impact than viewing mass media models (e.g., Cash, Cash, & Butters, 1988). Thus, exposure to appearance-related feedback from family members and peers and susceptibility to mass media model influence appear to play significant roles in the development of girls’ and women’s body esteem.

Although researchers have had empirical support for focusing on women, new findings suggest that men should also be included in studies on body esteem and body image. The prevalence of anabolic steroid use suggests that men are also dissatisfied with their bodies. One out of 15 male high school seniors report taking anabolic steroids, and 27% of those young men report taking steroids to improve their physical appearance (Fackelmann, 1988). Additionally, the 1999 “Monitoring the Future” study by the National Institute of Drug Abuse (NIDA) indicated that anabolic steroid use among young men has increased from 1.8% in 1991 to 2.7% in 1999 (NIDA, 2000). Grogan et al. (1996) claimed that men should also be included in studies on body esteem because Western attitudes about the male body are changing, and men are becoming increasingly concerned with body image. Whereas thinness appears to be vogue for women, men appear to be concerned with being too thin (i.e., “scrawny”). For instance, one study indicated that whereas 35% of men desired to be thinner, 43% desired to be heavier (Silberstein et al., 1988). Although recent trends indicate that men increasingly have body image concerns, some researchers posit that men’s body esteem would not be affected by portrayals of idealized male bodies because men are less likely to make appearance comparisons (Heinberg & Thompson, 1992).

In order to develop and strengthen the emerging evidence that men also experience body image concerns, some recent studies have included men in their samples. For instance, Grogan et al. (1996) examined the effects of viewing same-sex photographic models on women’s and men’s body esteem. In this
study, two groups viewed one of two types of photographs: attractive models or landscapes. The female models were thin and well proportioned, whereas the male models were well toned and muscular. The participants also completed a body esteem scale before and after viewing the photographs. Body esteem scores before viewing the photographs revealed that women, on average, had lower body esteem than men. However, after viewing the attractive models, all participants demonstrated a significant decrease in body esteem. In other words, this study demonstrated that both women’s and men’s body esteem is affected similarly by viewing attractive models. Subsequently, as Grogan et al. also point out, these findings do not support Heinberg & Thompson’s (1992) notion that men’s body esteem would not be negatively affected by exposure to mass media models. Grogan et al. offer one study to the body esteem literature on men, and it only focuses on the influence of mass media models. The subsequent step is to further examine men’s body image concerns by investigating the contributions of multiple sociocultural factors to men’s body esteem.

The purpose of the present study was to extend previous research on body esteem by further investigating sex differences in body esteem and their relation to susceptibility to mass media models and appearance-related feedback from family and peers. Past research has shown that when considering body esteem alone women tend to have lower body esteem than men (e.g., Grogan et al., 1996). Thus, a prediction of the current study is that women’s body esteem will be lower on average than men’s. In addition, past research has shown that women’s body esteem is negatively affected when women are exposed to mass media models or to body-related criticism from family and peers (e.g., Levine et al., 1994; Stice et al., 1994; Turner et al., 1997), and that men’s body esteem is negatively affected when men are exposed to mass media models (Grogan et al., 1996). Recent findings have also indicated that men are increasingly concerned about their bodies and that men’s and women’s body esteem is similarly affected by exposure to mass media models (Grogan et al., 1996). Thus, another prediction of this study is that a high susceptibility to mass media models, as well as a high rate of appearance-related feedback from family and peers, will be similarly associated with lower body esteem for both men and women. Additionally, a high influence of sociocultural factors is predicted to be associated with both men and women similarly perceiving their own body shape as more dissimilar to their desired and perceived societal ideal body shapes.

Method

Participants

Participants for this study were 279 students enrolled in introductory psychology courses at a mid-sized midwestern commuter campus. The sample consisted of 166 women and 113 men with a mean age of 22.6 years (range: 17 to 48 years, SD = 6.85). The sample was predominantly Caucasian (88.1%), with 5.4% African Americans, 2.9% Hispanic Americans, 2.2% multiracial, 1.0% Asian Americans, and 0.3% Native Americans. The researchers recruited participants through posted sign-up sheets that indicated the topic under investigation, the procedure to be used, and possible risks of completing the project. The researchers treated the participants in accordance with APA guidelines for ethical treatment of human participants (APA, 1992).

Measures

For the purposes of this study, the researchers created a Sociocultural Factors Questionnaire. Although most items were adapted from subscale items created by Levine et al. (1994), the researchers added new items in order to obtain specific information regarding physical fitness, same-sex versus opposite-sex peer teasing and dieting, and childhood versus current appearance-related feedback. The original subscales by Levine et al. were Peer Diet (3 items), Peer Teasing (1 item), Family Teasing (5 items), Magazine Information (7 items), and Model (8 items). The final measure contained 42 items, which assessed (a) the extent to which appearance-related mass media messages influence participants’ notions about body image, (b) the extent to which participants receive appearance-related compliments or criticism from family and peers, and (c) the extent to which participants’ peers express body image concerns. Because new items were added to preexisting subscales, the researchers completed a new factor analysis. This exploratory factor analysis supported the existence of three distinct factors among 31 items that demonstrated acceptable internal consistency: Mass Media Susceptibility (α = .85), Family and Peer Teasing (α = .84), and Peers with Body Concerns (α = .73). Appendix A displays those items within each factor. Depending on the nature of the item, the participant chooses how many individuals he or she knows who fit the question asked (from 1 = none to 5 = all), how often the item occurs (from 1 = never to 5 = all the time), or how much he or she agrees with the item (from 1 = disagree to 5 = agree). Higher scores on each subscale reflect a greater susceptibility to mass media model influence, more appearance-related teasing from family and peers, and a higher frequency of
exposure to discussions with peers about body concerns, respectively.

The researchers administered the Body Esteem Scale for Adolescents and Adults (Mendelson, White, & Mendelson, 1997) to obtain individuals’ perceptions and feelings about their bodies. The questionnaire contains 20 items, which Mendelson et al. (1997) found to have strong test–retest reliability and convergent validity. Examples of those items are “I’m happy about the way I look” and “I think about ways to change my weight.” Participants choose how often they feel the way the item describes (from 1 = never to 5 = all the time). Three of the original 20 items were dropped in the present study to maximize internal consistency. One average score was utilized from the remaining 17 items. The scale’s internal consistency, as measured by the alpha coefficient, was .95. Higher scores reflect more positive feelings about one’s body, or in other words, higher body esteem.

For the purposes of this study, the researchers created a Body Shape Questionnaire to assess pictorially individuals’ body shape perceptions in terms of thinness and roundness (see Appendix B). The researchers selected the male and female figure drawings from a larger continuum of figure drawings as presented by Stunkard, Sorensen, and Schulsinger (1983). Participants selected from a series of five figure drawings differing in thinness: (a) their perceived current body shape, (b) their desired body shape, and (c) the body shape they believe the media, family, and friends expect them to have (i.e., perceived ideal body shape). The actual number of the figure chosen for each item was retained as the score for that item, and thus could range from 1 to 5. A higher score reflects that the participant selected a thinner figure, whereas a lower score reflects that the participant selected a rounder figure. In order to assess the extent to which participants’ desired and perceived ideal body shapes differed from their current body shape, two difference scores were created (c.f. McKinley, 1998; Silberstein et al., 1988). These scores were calculated by subtracting the participants’ own perceived current body shape figure from their desired body shape figure and their perceived current body shape figure from their perceived ideal body shape figure. The absolute values of the resulting numbers were then obtained so that discrepancies in one direction of desiring to be thinner did not cancel out discrepancies in the other direction of desiring to be rounder (c.f. Silberstein et al., 1988). The absolute values reflect the degree, or amount, of body change desired. Therefore, larger absolute values reflect greater desired body change, whereas smaller absolute values reflect less desired body change. The following analyses only utilize the absolute values of the two discrepancy scores.

Procedure
Participants completed the packet of questionnaires in groups of 10 to 30 individuals. The researcher informed the participants that the study concerned how individuals feel and think about their bodies, how individuals feel about and perceive mass media models, and how often family and friends provided appearance-related comments. The researcher distributed blank envelopes containing the above three questionnaires, presented in random order. When the packets were completed, the researcher instructed the participants to put all materials back into the envelope, close the envelope, and turn it in. Thus, participants’ responses were anonymous. Sessions lasted approximately 45 min.

Results
Sex Differences in Mean Body Esteem and Body Shape Discrepancy Scores
A t test for independent samples was performed on men’s and women’s body esteem scores in order to determine the existence of sex differences. The test indicated a significant difference between the body esteem means for men (M = 3.99, SD = .64) versus women (M = 2.91, SD = .85), t(273) = 4.99, p < .001. In addition, t tests for independent samples were completed for both body shape discrepancy scores. The test for the absolute value of the desired-current discrepancy indicated no significant difference between the means of men (M = .57, SD = .58) versus women (M = .67, SD = .63), t(279) = –1.38, p > .05. Likewise, for the absolute value of the perceived ideal-current discrepancy, the test also indicated no significant difference between the means of men (M = .77, SD = .63) versus women (M = .84, SD = .77), t(279) = –.78, p > .05. Therefore, men and women report requiring a similar amount of body change to match their desired and perceived ideal body shapes.

The Contributions of Sociocultural Factors to Body Esteem and Perceptions of Thinness by Sex
Correlations between body esteem and the sociocultural factors were investigated separately for men and women. Both men’s and women’s body esteem was found to be significantly negatively correlated with susceptibility to mass media model influence, appearance-related teasing, and exposure to discussions with peers about body concerns (see Table 1 for correlations by sex). Although both sets of correlations for men and women were significant and in the same direction, Fisher’s r to z transformations
(Daniel, 1995) were utilized to determine if their magnitudes differed significantly. A significant sex difference was found between body esteem and susceptibility to mass media model influence, \( z = -3.29, p < .05 \). However, no sex differences were indicated between body esteem and appearance-related teasing, \( z = -1.26, p > .05 \), or between body esteem and exposure to discussions with peers about body concerns, \( z = -1.76, p > .05 \). Thus, the current findings indicated that appearance-related teasing and exposure to discussions with peers about body concerns similarly affect men’s and women’s body esteem. However, the findings did indicate that women’s body esteem is related more strongly to susceptibility to mass media model influence than is men’s body esteem.

In order to examine whether the relations between the sociocultural factors considered and the absolute values of the two body shape discrepancy scores differed for men versus women, zero-order correlations among the two body shape discrepancy scores and the three sociocultural factors were conducted separately for men and women (see Table 1 for the correlations by sex). The analysis yielded only two pairs of correlations indicating that men and women might be similarly affected—those correlations between appearance-related teasing from family and peers and the absolute value of the desired-current discrepancy (for men, \( r[111] = .23, p < .05 \); for women, \( r[165] = .42, p < .001 \)) and the absolute value of the perceived ideal-current discrepancy (for men, \( r[109] = .31, p < .05 \); for women, \( r[165] = .37, p < .001 \)). Thus, the higher degree of teasing men and women experience, the greater amount of body shape change necessary to match their desired and perceived ideal body shapes. Although these four correlations for men and women were statistically significant and in the same direction, Fisher’s \( r \to z \) transformations (Daniel, 1995) were again utilized to determine significant differences among the magnitudes. A significant sex difference was indicated for the relation between appearance-related teasing and the absolute value of the desired-current discrepancy, \( z = -2.22, p < .05 \). Thus, the magnitude of the relation for women is stronger than for men. However, no sex difference was indicated for the relation between appearance-related teasing and the absolute value of the perceived ideal-current discrepancy, \( z = -.22, p < .05 \). Thus, the magnitude of the relation for women is stronger than for men. However, no sex difference was indicated for the relation between appearance-related teasing and the absolute value of the perceived ideal-current discrepancy (for men, \( r[111] = .23, p < .05 \); for women, \( r[165] = .42, p < .001 \)). Thus, the higher degree of teasing men and women experience, the greater amount of body shape change necessary to match their desired and perceived ideal body shapes. However, higher levels of the same sociocultural influences were not related in a similar way for men.

### Discussion

The present study attempted to replicate and extend previous research findings by investigating the impact of multiple sociocultural factors simulta-
necessarily on men’s and women’s body esteem and on perceptions of current body thinness versus perceptions of desired and perceived ideal body thinness. First, we predicted that women’s body esteem would be lower on average than men’s body esteem. This hypothesis was affirmed as the current analysis indicated that women’s body esteem was, in fact, lower on average than men’s. These present results concerning sex and body esteem replicated past research (Grogan et al., 1996; Miller et al., 2000). Although men are becoming increasingly concerned with their body image, as shown by increases in anabolic steroid use (NIDA, 2000) and by recent research portraying negative effects on men’s body esteem due to viewing attractive models (Grogan et al., 1996), men continue to score higher, on average, on measures of body satisfaction. However, although a sex difference existed for overall body esteem, no sex differences were found among the absolute values of the body shape discrepancy scores. Because men desire a similar amount of body change with respect to perceptions of thinness as do women in order to match their desired and perceived ideal body shapes, the notion that men have increasing body image concerns was reinforced. Men desire to change their bodies just as much as women, but men do not report as low a body esteem as do women. This result can probably be explained by the notion that societal pressures on women to be a particular size and shape are still more pronounced than those pressures on men (Rothblum, 1990), and that women feel more pressure to conform to those societal ideals (Ogletree et al., 1990). Therefore, when a woman believes her body shape does not conform to her desired or perceived ideal body shapes, she is likely to feel more dissatisfied with her body than would a man who has the same beliefs. Furthermore, there may be a wider range of acceptable body shapes for men, which may enable them to feel more satisfied with their bodies even when their current body shape does not match their desired or perceived ideal body shapes.

Second, the researchers hypothesized that a high susceptibility to mass media model influence and a high rate of appearance-related feedback from family and peers would be similarly associated with lower body esteem for both men and women. The current results affirmed this hypothesis, as lower body esteem was associated with higher levels of susceptibility to mass media models, exposure to discussions with peers about body concerns, and appearance-related teasing from family and peers for both men and women. These results are consistent with past research that has indicated that women report lower body esteem when they also report higher levels of susceptibility or exposure to sociocultural influences (Henderson-King & Henderson-King, 1997; Levine et al., 1994; Rieves & Cash, 1996; Stice et al., 1994; Turner et al., 1997) and that men report lower body esteem after exposure to attractive models (Grogan et al., 1996). Thus, there is strong evidence supporting the proposition that sociocultural factors play a significant role in individuals’ perceptions and feelings about their bodies. It appears that individuals are internalizing the criticisms and other negative appearance-related feedback from close family members and friends, as well as mass media messages about the ideal body shape. Although all three sociocultural factors negatively affected both men’s and women’s body esteem, women’s body esteem was more strongly related to susceptibility to mass media model influence. This result is consistent with the notion that women face more societal pressures to be thin (Ogletree et al., 1990). However, this finding is inconsistent with past findings indicating that men’s and women’s body esteem is affected similarly by portrayals of ideal body shapes (Grogan et al., 1996), as the present findings indicated that women were affected more profoundly than men. However, men’s and women’s body esteem was affected similarly by appearance-related teasing from family and friends and by exposure to discussions with peers about body concerns. Although low body esteem may not appear on the surface to be a problematic issue for men, their body esteem is significantly affected by these sociocultural influences that idealize a thin body. Although Heinberg & Thompson (1992) suggested that men’s body esteem would not be negatively affected by exposure to body-related sociocultural influences, the present results indicated that men’s body esteem was inversely related to all three sociocultural factors. With respect to exposure to peers discussing body concerns and appearance-related teasing, the notion that men’s body esteem is affected similarly to women’s body esteem by sociocultural factors was replicated in the present study. These body image concerns of men may play a significant role in the recent increase of anabolic steroid use among men (Fackelmann, 1988; NIDA, 2000). Men also appear to have body image concerns and should be included in future studies on body esteem and body image.

Lastly, to further extend the findings of past research, this study also investigated sociocultural influences of discrepancies between men’s and women’s perceptions of current body thinness versus perceptions of desired body thinness and the body shape they believe most people want them to have in terms of thinness (i.e., perceived ideal body shape). We predicted that a high susceptibility to mass media influence
model influence and a high rate of appearance-related feedback from family and peers would be associated with both men and women perceiving their own body thinness as more dissimilar to their desired and perceived ideal body shapes. However, the data only partially confirmed this hypothesis; just one set of correlations indicated that men and women were similarly affected. Higher rates of appearance-related teasing from family members and friends was the only sociocultural factor associated with greater discrepancies between both men’s and women’s current body shapes and their desired and perceived ideal body shapes. This finding is consistent with the Cash et al. (1983) notion that appearance-related information from peers has a greater effect on body dissatisfaction than do mass media models. Apparently, the appearance-related criticisms from family members and peers have a very direct and salient impact on one’s internalized perceptions about the body. The remaining correlations among the sociocultural factors and the body shape discrepancy scores indicated the following sex differences: Women’s desired-current discrepancies were more profoundly affected by high rates of appearance-related teasing than were men’s desired-current discrepancies. Furthermore, whereas women’s desired-current and perceived ideal-current discrepancies were significantly related to high rates of exposure to discussions with peers about body concerns and susceptibility to mass media model influence, men’s body shape discrepancies were not. These current findings indicate that women desire changes in body thinness when confronted with exposure to mass media models and body-concerned peers, whereas men do not.

These results suggest that low body esteem and body image concerns remain problematic for women to a greater extent than for men. However, recent research (e.g., Grogan et al., 1996) and the present results also indicate that men have some body image concerns when considered in relation to certain sociocultural influences (i.e., exposure to pictures of “ideal” male bodies, appearance-related teasing). The present findings reinforce the prevalent notion that sociocultural factors strongly affect body esteem and body shape preferences, especially for women. Finally, although women’s body concerns appear to be associated with a greater extent with sociocultural influences, men seem to have increasing concerns about body image and body shape.

The present study is not without limitations. For example, the lack of cultural diversity in the predominately Caucasian sample made it impossible for the researchers to adequately examine ethnic or racial differences. Ethnic or racial differences should be considered when possible, as some research has indicated that ethnic traditions are important elements of beliefs about body shapes (e.g., Collins, 1991; Mumford, Whitehouse, & Platts, 1991; Rucker & Cash, 1992). Furthermore, additional research might consider body image and body esteem along other dimensions of body shape besides thinness (i.e., physical fitness, muscularity) that may be related to sociocultural factors. Despite these limitations, the present findings replicate past research indicating that women’s body esteem is negatively affected by multiple sociocultural influences and also extend past research by demonstrating that men’s body esteem is also affected by multiple sociocultural influences. Whereas some sex differences exist among the associations of body esteem and body image with sociocultural influences, in some respects men and women are affected similarly, particularly when it comes to the influence of interpersonal factors, such as teasing and exposure to peers’ dissatisfaction with their own bodies. Thus, a better understanding of how individuals view and feel about their physical selves will likely result from studies that include men as well as women.

References


APPENDIX A

Subscales and Items from the Sociocultural Factors Questionnaire

Subscale—Teasing/Criticism From Others ($\alpha = .84$)
1. When you were younger, how often did same-sex peers tease you about being too fat or too thin?
2. When you were younger, how often did same-sex peers tease you about not being muscular and/or physically fit enough?
3. Now in your life, how often do same-sex peers tease you about being too fat or too thin?
4. Now in your life, how often do same-sex peers tease you about not being muscular and/or physically fit enough?
5. When you were younger, how often did opposite-sex peers tease you about being too fat or too thin?
6. When you were younger, how often did opposite-sex peers tease you about not being muscular and/or physically fit enough?
7. Now in your life, how often do opposite-sex peers tease you about being too fat or too thin?
8. Now in your life, how often do opposite-sex peers tease you about not being muscular and/or physically fit enough?
9. How often does your brother (or brothers) tease you about your weight or body shape?
10. How often does your sister (or sisters) tease you about your weight or body shape?
11. How often do you get criticized about your weight or body shape by an adult female family member?
12. How often do you get complimented about your weight or body shape by an adult female family member?
13. How often do you get criticized about your weight or body shape by an adult male family member?
14. How often do you get complimented about your weight or body shape by an adult male family member?

Subscale—Mass Media Influence ($\alpha = .85$)
1. I try to look like some of the models I see by purchasing the products they advertise.
2. If I diet or try to build muscle, I look to the models in magazines for inspiration (e.g., put their pictures on a wall in your room).
3. I compare myself to the models I see.
4. I would like to look like the models I see in magazine and TV ads.
5. I feel bad about my own body after seeing attractive models in magazines and on TV.
6. I am motivated to exercise or work out after seeing attractive models in magazines and on TV.
7. Magazines/TV influence my idea of the perfect body.
8. I think the models in magazines/on TV are confident and happy.
9. I think the models are among the best looking people I have ever seen.
10. When I read a magazine or watch TV, I enjoy looking at the models in the ads.
11. How often do you look at or read magazines?

Subscale—Peers Expressing Body Image Concerns ($\alpha = .73$)
1. Of the same-sex peers you know, how many would like to be thinner?
2. Of the same-sex peers you know, how many would like to be more muscular/physically fit?
3. Of the same-sex peers you know, how many are dieting/trying to lose weight?
4. Of the same-sex peers you know, how many talk about weight, weight loss, and/or building muscle?
5. Of the opposite-sex peers you know, how many would like to be more muscular/physically fit?
6. How often have you changed your appearance to look like an attractive, same-sex peer?
APPENDIX B

Body Shape Questionnaire

DIRECTIONS: This questionnaire asks for your opinions about the way your own body looks in terms of thinness/roundness, and how thin/round you think others think people’s bodies should look. Please answer the questions by looking at the drawings of male and female bodies below, and by circling the number that goes with the body that best shows your opinion (even if not exactly).

Questions 1–3 are for WOMEN ONLY:

1. Which of the female bodies above do you think you look like right now?
   1 2 3 4 5

2. Which of the female bodies above would you like to look like?
   1 2 3 4 5

3. Which of the female bodies above do you think everyone (friends, TV, magazines) wants you to look like?
   1 2 3 4 5

Questions 4–6 are for MEN ONLY:

1. Which of the male bodies above do you think you look like right now?
   1 2 3 4 5

2. Which of the male bodies above would you like to look like?
   1 2 3 4 5

3. Which of the male bodies above do you think everyone (friends, TV, magazines) wants you to look like?
   1 2 3 4 5
Effects of High and Low Saturation of Red Hue on Long-Term Memory Performance

Ani G. Nikolova
Jill L. Quilici*
California State University, Northridge

We purposely incorporate color in many elements of our visual world, and yet we have not fully examined its potential to trigger physiological and psychological reactions in individuals. Researchers have found that the different properties of color can affect the direction of our emotional reactions in terms of feelings (Jacobs & Suess, 1975; Weller & Livingston, 1988), anxiety (Profusek & Rainey, 1987), and arousal or level of alertness (Batra, Urvashi, & Muhar, 1998; Jacobs & Hustmyer, 1974; Mikellides, 1990; Valdez & Mehrbain, 1994; Wilson, 1966). Studies have also established color reactions as a function of personality and psychopathology (Frank, 1976), reactions to color concepts (Adams & Osgood, 1973), and behavior changes due to color properties (Damhorst & Reed, 1986). However, we know little about the possible effect of color on various cognitive processes such as learning, memory performance, reasoning, and problem solving.

Why Color Saturation?

Experimenters have typically confounded saturation and brightness properties of color in studies investigating the various effects of color, thus leading to methodologically weak and sometimes contradictory results. In most previous studies (Ainsworth, Simpson, & Cassell 1993; Jacobs & Suess, 1975; Profusek & Rainey, 1987; Wilson, 1966), researchers have studied color by manipulating color hue (i.e., wavelength) solely and have failed to adequately con-
trol for its other two properties, brightness (i.e., black-to-white quality) and saturation or chrome (i.e., purity or vividness). Unlike these studies, Valdez and Mehrabian (1994), in their research on the effects of color on emotions, found significant emotional reactions (arousal, pleasure, and dominance feelings) to both brightness and saturation variations of color. In this study, they addressed various methodological and conceptual issues by using a wide range of color samples and by providing experimental control in investigating the separate effects of color hue, saturation, and brightness. They used 76 color stimuli of each hue level, varying as much as possible in brightness and saturation, and used 250 participants to rate the samples. Valdez and Mehrabian used Mehrabian’s (1978) verbal report scales for pleasure–displeasure, arousal–nonarousal, and dominance–submissiveness measures to assess emotional reactions to the different properties of color. The arousal–nonarousal scale contained eight word pairs (e.g., “frustrated–sad”, “troubled–dull”) that differed with respect to arousal but were equal on dominance and pleasure measures. The results of their study indicated that less bright and more saturated colors elicit significantly higher arousal reactions in individuals than more bright and less saturated colors, regardless of color hue. Surprisingly, results relating hue and arousal were nonsignificant.

Mikellides (1990) also examined the effects of color properties on arousal and obtained similar results. In contrast to Valdez and Mehrabian (1994), Mikellides measured arousal by means of EEG and EKG recordings and manipulated color stimuli by using rooms painted in red and blue variations of saturation. Mikellides also found that chromatic strength (saturation) was the property of color that most affected our perceptions of how exciting or calming a color was. Again, there was no significant relation found between color hue and physiological arousal.

Profusek and Rainey (1987) examined a similar theme by looking at the effects of rooms painted in red or pink on human anxiety level. They found that rooms painted in pink were significantly less psychologically arousing than rooms painted in red. However, more recently, Ainsworth et al. (1993), who also examined the effects of colored rooms on states of anxiety and on work performance in an office environment, came to the conclusion that color, in particular red and blue-green, does not have an effect on the variables studied. These results are not very surprising because, even though neither of these experiments controlled for brightness and saturation dimensions of color, pink hue generally tends to be of higher brightness and lower saturation than the red and blue-green hues. Evidently, we could explain the contradiction of these studies by an existing correlation between levels of arousal and color brightness and/or saturation but not between arousal and hue variations.

If Color and Arousal are Significantly Correlated, Can We Hypothesize That Color and Mental Performance are Correlated as Well?

Psychologists who have studied the amount of mental effort required to perform a task have found that the amount of mental capacity available varies with the level of arousal (Kahneman, 1973; Yerkes & Dodson, 1908). In particular, they came to the conclusion that more mental capacity is available when arousal is moderately high. At the same time, performance suffers when arousal level is either too high or too low for the specific task. Consequently, we can assume that if certain variations of color brightness or saturation have an effect on the state of arousal in individuals, then color should have an effect on the amount of mental capacity available for performing various cognitive tasks. Thus, if we look at the relation between memory performance and color saturation, we would expect that memory outcome will fluctuate in response to saturation variations. However, because the relation between hue and arousal is weak (Mikellides, 1990; Valdez & Mehrabian, 1994), we would not anticipate a significant effect of hue alterations on memory performance.

Is Memory Performance Affected by Color Properties or by the Memory of Color Context?

We know little about the exact mechanism of our physiological reactions to color and its impact on memory performance. Despite that lack of knowledge, many researchers who have conducted studies on the effect of environmental changes on recall (Dulsky, 1935; Weiss & Margolius, 1954) and recognition (Murnane & Phelps, 1993) have used background color as an environmental context. Thus, by varying the color saturation and brightness of the cards on which paired associates were placed, Dulsky (1935) found that participants recalled less information when he changed the color background presented in the learning situation for the recall test as compared to when the color background during recall remained the same as the background during study. Dulsky interpreted the results as due to the color context factor alone (i.e., color background during learning vs. color background during recall), omitting the possibility of any color property effects.

Weiss and Margolius (1954) found similar results in their study. They defined same context as colored
background during both study and recall, and altered context as either colored background during study and gray color during recall or gray background at study and colored context at test. According to their results, recall was better when the colored backgrounds remained the same compared to when they were changed. They also found that memory performance was significantly better when they had employed color contexts, compared to the condition in which participants viewed a gray background during study or recall. Although we could attribute the significance of these results to the stimulating or arousing effect of the colored background compared to the effect of neutral (gray) background, Weiss and Margolius concluded that forgetting increased due to the change of environmental context.

More recently, Murnane and Phelps (1993) found that memory recognition also decreases with changes in color context. Thus, they concluded that memory performance declines when we test participants in an environmental context (i.e., color context) that is different from the study context, not only when we measure memory by recall, but also when we test memory by recognition.

However, in none of these three studies have researchers specified the exact color characteristics in terms of brightness and saturation. More than that, researchers have not considered the possible stimulating or inhibiting physiological effects of certain color characteristics on recognition or recall. Apparently, they have failed to consider the possible effects of color properties on mental performance when changing the environmental contexts and, thus, have most likely confounded the results of their studies. Whether performance on memory tasks is affected by changes in the environmental context or by the properties of color itself is unclear.

Summary of Research Goals

The main goal of this study was to examine whether manipulation of color properties, manipulation of color context, or both factors have an effect on memory recall. The experiment built on the hypothesis that if manipulation of color saturation can increase or decrease arousal reactions (Mikellides, 1990; Valdez & Mehrabian, 1994) and if levels of arousal and memory performance are significantly correlated (Yerkes & Dodson, 1908), then we expected that the alteration of saturation of the environmental color during study or test would affect the outcome in memory recall.

Based on previous research, we made the following hypotheses:

1. There would be significant main effects for color at study on memory performance as well as color at test on memory performance. In particular, we expected that participants viewing the highly arousing color stimulus (red hue of high saturation) at the time of material study would perform significantly better on the recall task than participants viewing the low-arousing (red hue of low saturation) or neutral (medium gray) color stimulus during study, regardless of the color at test. We expected the same pattern of results for the main effect of the color at test: we hypothesized that students viewing the high-saturated color at the time of recall would score significantly higher on the memory test than students looking at the low-saturated or neutral color during recall, regardless of the properties of color during study. Simultaneously, because the neutral (gray) color has no saturation and no hue (Munsell Color System, available from the Macbeth division of Kollmorgen Corporation), we anticipated that the low-saturated color, by decreasing the level of arousal (Valdez & Mehrabian, 1994), would also lower memory performance. Thus, we hypothesized that the recall scores of students viewing the low-saturated color at study or test would be significantly lower than the recall scores of students viewing the neutral color. Because of the lack of saturation value of the neutral color, we assumed that its properties would not affect the initial level of arousal, and thus when looking at the main effects of color we considered participants presented with this stimulus during study or test to represent control groups.

2. There would be no significant contextual interaction between color at study and color at test. Thus, we hypothesized that there would be no significant differences found between memory performance of participants presented with the same color stimuli during study and test and performance of participants presented with different color stimuli during study and test.

Color hue was not a variable in this experiment because, as we mentioned above, Valdez and Mehrabian (1994) have clearly demonstrated that hue itself does not have a significant effect on arousal. Because we built our hypotheses on the assumption that memory performance and states of arousal are significantly correlated, if hue does not have an effect on arousal, then the relation between hue and memory should be weak as well. We did not consider the results of some of the studies mentioned above (Murnane & Phelps, 1993; Weiss & Margolius, 1954), which implied that color hue accounted for significant changes in memory, because these researchers failed to control for brightness and saturation of the color stimuli.
We chose randomly the red hue from the variation of basic colors (hues) of the visual spectrum.

Method

Participants

We tested a randomly drawn sample of 180 students. All participants were female undergraduates from California State University, Northridge, who partially fulfilled an introductory psychology course requirement by voluntarily participating in the Psychology Department Participant Pool. We included only female participants in the study because women show a significantly stronger pattern of reactions to the different properties of color than men do (Valdez & Mehrabian, 1994).

Design

We randomly assigned participants to one of the nine treatment combinations, according to a $3 \times 3$ randomized-groups design. Each of the two independent variables—color presented at the time of study and color presented at the time of test—had three levels: high saturation of red hue, low saturation of red hue, and neutral (medium-gray) hue. The dependent variable was memory recall.

Materials and Setting

For each color variation, there were two color-sample boards of the size $30 \text{ cm} \times 45 \text{ cm}$. We took all color stimuli from the Munsell System for color notation (available from the Macbeth division of Kollmorgen Corporation), and we chose them to be of equal brightness. We used the following specifications of color stimuli: Munsell red with saturation (chrome) of 14 and brightness (value) of 4 (high-saturated red), Munsell red with saturation of 4 and brightness of 4 (low-saturated red), and Munsell gray with brightness of 4 (N4) and no saturation (neutral color). We placed the two sample boards for each treatment condition on the wall and on the surface of a desk, each one at the background of a medium-gray poster. The laboratory room contained no windows and was illuminated with fluorescent tubes, which approximated daylight condition.

We used 30 common English words, which were between four and six letters in length and contained either one or two syllables, to create our study list for the dependent variable. The first author recorded the words from the study list with an equal 2-s interval between each word by pronouncing all words with the same volume and tension of voice. Also, we used pseudoisochromatic plates (Hardy, Rand, & Rittler, 1957) for the purpose of detecting defective color vision in the participants. The color vision test consisted of six screening plates. Each plate consisted of a maximum of two colored symbols or no symbols, located at different positions of the plate. The sequence of the six plates contained the following combinations of shapes: (a) a circle and an X, (b) a triangle and an X, (c) a circle, (d) no symbols, (e) a circle and an X, and (f) a circle and a triangle.

Procedure

We assigned participants to one of the conditions based on their order of appearance and tested them individually. We seated each student at a desk, approximately 75 cm from the color board positioned on the wall. We kept the environment quiet throughout the experiment and held the room temperature constant at $21^\circ \text{ C}$.

Prior to the experiment, we tested all participants for defective color vision. The first author presented each participant with the six screening plates in sequence and asked the participants to indicate the number, shape, and location of all colored symbols on the plate. First, the experimenter asked the participant how many colored symbols she saw. Second, the experimenter instructed the participant to indicate the shapes of the colored symbols by allowing the participant to use her own names (e.g., a circle could be labeled as “zero,” “circle,” etc.). Third, the experimenter asked the participant to point to the location of the symbol. An error consisted of failure to see all symbols, or to name incorrectly any symbol, or to point to an incorrect location of any symbol. We identified no color-blind students during the test; all students gave correct responses to all questions referring to the six plates.

The experiment consisted of three parts.

1. Listening task (2.5 min). Upon arrival, we asked each student to look at the color in front of her for a period of 1 min. The purpose of this procedure was to get the person’s eye photoreceptors adapted to the color stimulus of the surrounding, which is an important condition for the validity of the experiment. Next, we asked the student to listen to the word list, played by a tape recorder (1.5 min), and to try to remember as many of the words as possible. The words were presented auditorily so that students could keep their eyes focused on the color-sample boards.

2. A short math task (50 s). During the second part of the experiment, we asked each participant to start counting aloud, backward, from 100 to 0, each time subtracting three numbers (e.g., 100, 97, 94). The purpose of this procedure was to prevent students from rehearsing the words. Students did not look at the color-sample boards while carrying out this task. While students were counting, we changed color-sample boards when it was necessary.
3. Memory recall (2.5 min). We used a recall test to assess the effects of color manipulations and change of color context on long-term memory performance. We informed the participants that the time allowed for completing the task would be 2.5 min. We also instructed each student to pronounce aloud all of the words remembered, while looking at the color in front of her.

Results

We used a $3 \times 3$ randomized-groups analysis of variance to examine the effects of color at study (high, low, neutral) and color at test (high, low, neutral) on recall. The number of correct words recalled by each individual served as data for the corresponding condition. Incorrect responses and words recalled more than one time were excluded from the data.

Does color saturation at study have an effect on memory recall? As predicted, the main effect of color at study was significant, $F(2, 171) = 4.46$, $p < .05$. Figure 1 illustrates the means and standard errors for this effect. Furthermore, analysis of main comparisons for color at study (Factor A) revealed significant differences between the recall scores of students in the high-saturated group and students in the low-saturated group, $F_{A\text{comp}}(1, 171) = 7.33$, $p < .01$, and between the recall scores of the high-saturated and neutral groups, $F_{A\text{comp}}(1, 171) = 6.03$, $p < .03$. Thus, participants presented with the high-saturated red at the time of material study scored significantly higher on the recall task than participants presented with the low-saturated red and participants viewing the neutral stimulus. However, contrary to what we hypothesized, the analysis of the main comparison between memory performance of students in the low-saturated group and students in the neutral group was not significant, $F_{A\text{comp}}(1, 171) = 0.063$, $p > .05$. Thus, the recall scores of students looking at the low-saturated color during study were not significantly different from the recall scores of students looking at the neutral stimulus.

Does color saturation at test have an effect on memory recall? Likewise, the main effect of color at test was also significant, $F(2, 171) = 14.01$, $p < .01$. Analysis of main comparisons for color at test (Factor B) revealed significant differences between the memory performance of participants in the high-saturated group and the performance of participants in the low-saturated group, $F_{B\text{comp}}(1, 171) = 27.86$, $p < .01$, as well as between the recall scores of the high-saturated and neutral groups, $F_{B\text{comp}}(1, 171) = 5.66$, $p < .03$. Thus, similar to the pattern of results for Factor A, students tested in the presence of the high-saturated stimulus scored significantly higher on memory recall than students tested in the presence of the low-saturated stimulus and students presented with the neutral color during recall. As predicted, memory performance of students viewing the gray color at test was significantly better than the memory performance of students looking at the low-saturated red, $F_{B\text{comp}}(1, 171) = 8.41$, $p < .01$. (Refer to Figure 2 to see the means and standard errors for the main effect of color at test.)
Does a change in color context from time of study to time of test reduce memory recall? The present results revealed no significant interaction effect between the two factors on memory recall, $F(4, 171) = 2.37$, $p > .05$. As hypothesized, memory recall of students viewing the same color stimuli during study and during test did not differ significantly from the recall of students viewing a different color stimulus during study than they viewed at test. Table 1 illustrates the mean recall scores and standard errors as a function of color at study and color at test.

### Discussion

The present study provides evidence for the hypothesis that manipulating color saturation of the environment during study or test, while holding other color properties (hue and brightness) constant, can affect human memory performance. As hypothesized, the current results indicate that material learned in the presence of a high-saturated color is recalled significantly more than when the same material is learned in the presence of a low-saturated or a neutral (gray) color. The same results apply when these colors are presented during the memory test, regardless of the environmental color at study. At the same time, the findings demonstrate that scores on memory recall are significantly lower when a low-saturated color stimulus is presented during the test compared to when a neutral stimulus is presented. These results are consistent with Valdez and Mehrabian’s findings (1994), showing that high-saturated colors elicit significantly higher arousal in individuals than low-saturated colors, regardless of hue and brightness values.

Furthermore, the finding that there was a significant relation between the low-saturated red and the neutral gray color during testing (recall was higher for participants viewing the gray color), but not between the same colors during study, is quite interesting. Because too much or too little arousal can hinder performance on tasks requiring mental effort (Yerkes & Dodson, 1908), we can infer that the low-saturated color of testing, by inducing low arousal, decreases memory performance. Yet, according to the present results, the same does not apply to color at study. Whether these findings could be explained by the fact that study and recall are two distinct information-processing stages generated by different physiological mechanisms, or by a possible interactive effect of type of information processing and level of saturation of a given hue, should be further investigated.

The results also provide evidence for the hypothesis stating that change of color context does not have a significant effect on memory recall. Prior studies have established that changing the environmental context between learning and retrieval plays an important role in many explanations of forgetting (Godden & Baddeley, 1975; Smith, Glenberg, & Bjork, 1978). Apparently, this theory does not apply to color saturation context, even though many researchers have used color in studies on context manipulation and memory. Based on our results, researchers must raise questions about the cause of context-dependent memory effects found by manipulating the color of environmental context. It is possible that changes in color saturation rather than changes in environmental context caused these findings. Color saturation at study and color saturation at test have independent (not interactive) effects on memory performance. To illustrate, we found that memory performance of participants viewing the high-saturated color during both study and test was not significantly different from the memory performance of participants looking at the low-saturated color during study but with a high-saturated stimulus at test. However, when we looked at the separate effects of color at study and color at test we encountered significant results. Thus, we found that students viewing the high-saturated color at study scored significantly higher on recall than students presented with a low-saturated color during study, regardless of the color properties at test. The same pattern of the results was observed for color at test—participants presented with the high-saturated color during recall performed significantly better in memory than participants tested in a low-saturated environment.

Evidently, manipulating the properties of color seems to play a crucial role in the outcome on memory recall tests. Clearly, researchers should consider the stimulating and inhibiting effects of color characteristics in studies on memory and manipulation of color contexts. Furthermore, because our knowledge of the effects of color brightness and hue

---

**TABLE 1**

<table>
<thead>
<tr>
<th>Color at test</th>
<th>Color at study</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Red (high)</td>
</tr>
<tr>
<td>Red (high)</td>
<td>14.20 (0.56)</td>
</tr>
<tr>
<td>Red (low)</td>
<td>12.05 (0.50)</td>
</tr>
<tr>
<td>Gray</td>
<td>13.35 (0.88)</td>
</tr>
</tbody>
</table>
on memory is limited, researchers should use care when using colors to examine the effects of environmental changes and memory performance.

Apparently, color has the potential to trigger physiological reactions in humans. The mechanism leading to these reactions, though, is beyond the scope of this paper. One possible explanation of the results might be that color’s potential to trigger arousal is responsible for the changes in our mental capacity, or the amount of mental effort available for processing information. In other words, the direction of our arousal reactions to the different properties of color shapes the outcome of our memory performance. Because there is a strong relation between mental capacity and level of alertness (Kahneman, 1973; Yerkes & Dodson, 1908), we infer that the potential of color to affect our mental activation may account for its effect on memory recall. However, further research should examine whether there is a direct relation between color, arousal, and memory.

A potential criticism of this study must be addressed—namely, the lack of male participants, which limits the generalizability of our findings. Because women tend to be significantly more sensitive, in terms of arousal and other emotional reactions, to saturation and brightness variations of color than men (Valdez & Mehrabian, 1994), we decided not to include male participants in order to avoid further complexity of the experimental design. Future research should investigate the effects of color properties on memory of male participants, as well as any possible sex differences in memory performance due to different patterns of emotional reactions to color properties of men and women. Because men tend to show significantly weaker arousal reactions to color properties than women, if there is a significant correlation between levels of color saturation (brightness) and levels of physiological arousal, we would expect that color saturation (brightness) would have a significantly stronger effect on the memory of women than on the memory of men.

Another criticism of the study could be that its findings are primarily limited to the auditory presentation of study material and does not address visual presentation of information, displayed separately or embedded in the color stimulus. If we had presented participants with study materials in a written format on a separate piece of paper, participants would have visually focused on the study list and not as much on the color stimulus of the environment. The implementing of such a procedure would have interfered with the primary goals of our study. Yet, a critic might argue that if we had presented the list of words as part of the color stimulus, the outcome of the recall results could have been different. However, recent research suggests that color and form are represented separately in memory, are assessed independently, and can be integrated if the task requires it, but that such integration is a deliberate strategic response, not the natural consequence of encoding (Hanna & Remington, 1996). Therefore, if we had presented the test material visually as part of the color stimulus instead of auditorily, we would have increased the possible flaw due to individual differences in the knowledge and use of memory strategies. Although auditory presentation does not guarantee us that participants did not use any memory strategies in their attempt to perform well on the memory task, because we randomly assigned the participants to the conditions based on the order of their appearance, we have minimized the possibly confounding effect stemming from differences between the participants due to their degree of knowledge and use of memory strategies.

Further studies should also examine the possible effect of color on other cognitive processes, such as reasoning, problem solving, and performance on various tasks requiring mental effort. If high color saturation increases arousal, which, in turn, increases capacity for mental effort, then we would expect that high-saturated colors would significantly increase performance on problem solving, reasoning, and other similar cognitive processes compared to the low-saturated colors. Concurrently, researchers should investigate the effect of different percentage variations of color brightness and saturation on memory, as well as on other mental processes. As we mentioned earlier, similar to the effect of color saturation on arousal, different variations of color brightness significantly affect the levels of emotional arousal in humans. According to Valdez and Mehrabian (1994), less bright colors elicit significantly higher arousal reactions in individuals than more bright colors. Based on their findings and on the current results, we would expect that less bright colors would increase memory recall and more bright colors would decrease recall. Thus, we would expect a different pattern of results relating color brightness and memory compared to the results of the current study on color saturation and memory; while high-saturated colors tend to improve memory performance, as supported by the present results, we would expect that very bright colors would hinder memory performance.

Because we have purposely incorporated color in many elements of the human world, researchers should pursue a theoretical explanation of the shared reactions to color properties. Such knowledge would be of great application in many environmental settings of the academic and industrial world where
learning and memory performance is crucial and valuable. Choosing wisely the colors of the school setting, such as the color of the school classrooms, libraries, and other learning facilities, might improve significantly students’ academic performance. Currently, the walls of most classrooms and libraries are painted in white color, which tends to be of high brightness and low saturation. Our findings support the idea that students who learn or recall information in the presence of a low-saturated color setting score significantly lower on a memory test than students who learn or recall information in the presence of a high-saturated color setting. Therefore, in the school environment, it is important that we implement high-saturated colors and avoid the use of low-saturated color stimuli in order to optimize the conditions for students’ academic success. Concurrently, an appropriate use of color in the work environment could enhance the performance of office workers carrying out mentally demanding tasks by optimizing their performance in memory. However, because the present findings are limited to the effects of saturation levels of red hue with medium brightness, and because we examined its effects during a brief period of study (1.5 min) and test (2.5 min) time, future studies should include other levels of saturation and brightness variations of different hues and investigate the relation between color properties and memory for a longer interval of time. If the color of the environment has the potential to affect our productivity by inducing higher levels of alertness and, consequently, higher attention and better performance on mental tasks, then further exploration and utilization of the findings in this area is worthwhile.

References
Effects of DNA and Eyewitness Evidence on Juror Decisions

The present experiment investigated whether DNA evidence, eyewitness evidence, or a combination of both was more persuasive to mock jurors. The study also explored whether varying the credibility of the testimony affected persuasiveness. The researchers hypothesized that the combination of eyewitness and DNA evidence would be the most persuasive to jurors. Eyewitness evidence would be more persuasive than DNA evidence, and credible testimony would be more persuasive than discredited testimony. One hundred-fifty college students read an excerpt from a court transcript describing a first-degree murder trial. Participants rendered a verdict and answered questions concerning confidence, understanding, reliability, and persuasiveness of testimony. The analyses partly supported the hypotheses. DNA evidence was more persuasive, reliable, influential, and less likely to be viewed as wrong than eyewitness evidence, regardless of whether it was discredited. DNA evidence elicited more guilty verdicts, and jurors were more confident in their verdict. We discuss these results in light of the 2-routes-to-persuasion theory.

Naomi J. Freeman
Diana L. D. Punzo*
Earlham College

The persuasiveness of different types of evidence in criminal cases has been an ongoing issue in the field of forensic psychology. For years, researchers have studied the influence of accuracy, confidence, and reliability of eyewitness testimony on juror decisions (Brigham & Bothwell, 1983; Cutler, Penrod, & Martens, 1987; Loftus, 1980; Wrightsman, 1991). Eyewitness testimony and its effect on jurors are topics of interest not only to the psychological world, but also to the legal community because “the evidence produced by an eyewitness often makes the difference between an unsolved crime and a conviction” (Wrightsman, 1991, p. 157). Often, however, eyewitnesses are inaccurate in their identification of perpetrators (Wrightsman, 1991). Mistakes in identification can occur from the time when the crime is committed up to the time when the witness testifies at the trial. In addition, new experiences, leading questions, or police lineups can alter memories. Despite the unreliability of eyewitness identification, research has found that jurors are heavily persuaded by this form of evidence (Loftus, 1980; Wrightsman, 1991).

Many theories have attempted to explain why eyewitness testimony is such a powerful form of evidence. Elizabeth Loftus (1980) has suggested that people do not have a general understanding of the working of human memory. People tend to believe that memory works similar to a videotape recorder. However, Loftus has suggested that, in a way, people actually construct their memories. Loftus also argues that in everyday life truly precise memories are very rarely demanded of people. The common errors in recollection that occur daily go unnoticed because they are not particularly important. Because people do not understand the exact workings of human memory and tend to trust their own memories, they also tend to trust the memories of others. Therefore, jurors are more likely to trust the memory of an eyewitness, especially a confident eyewitness, than other forms of evidence. Other witnesses, such as experts, can provide only a small piece of information; an eyewitness, however, can describe events that occurred

Author note. Naomi J. Freeman, Earlham College, and Diana L. D. Punzo, PhD, Earlham College Department of Psychology. The author would like to thank Katharine S. Milar, PhD, Earlham College, for her assistance in preparing the manuscript for publication, and Lynda Hall, PhD, Ohio Wesleyan University, for her assistance with the statistical analysis. The author would also like to thank Martin Lewis, JD, for his assistance in the development of the court transcripts used in the study.

Correspondence concerning this article should be addressed to Naomi J. Freeman at freemna@earlham.edu.
in detail to the jurors, thus providing a great deal of information for the mental construction of events in jurors’ minds. Loftus suggested that it might be easier for jurors to work with a smooth recollection of the events, even if they are inaccurate, than to have to piece together fragments of the event on their own.

To test her theory that people will be more persuaded by eyewitness than other forms of evidence, Loftus (1980) asked participants to read a case in which a bad check had been written to purchase a television. Participants participated in one of four evidence conditions: (a) an eyewitness (the store clerk) who positively identified the defendant as the man who wrote the check, (b) a polygraph expert who testified that a polygraph test stated the defendant was lying when he claimed he did not write the check, (c) a fingerprint expert who examined prints left on the counter of the store by the person who wrote the check and the fingerprints matched the defendant’s prints, and (d) a handwriting expert who examined the check and claimed that it matched the defendant’s handwriting. Results indicated that convictions were highest in the eyewitness condition (78%), followed by fingerprints (70%) and polygraph test (53%), and lowest in the handwriting condition (34%). These results confirm that eyewitness evidence is one of the most persuasive kinds of evidence that can be presented.

Research on whether juries believe discredited eyewitnesses has produced equivocal results. Loftus (1980) has reported that jurors place a great deal of weight on eyewitness testimony, even if that testimony is suspect or discredited. McCloskey and Egeth (1983) conducted a meta-analysis of eyewitness research and determined that “jurors” often disregard eyewitness testimony that has been convincingly discredited. Kennedy and Haygood (1992) have argued that jurors are aware of the unreliability of eyewitnesses. The findings on the persuasiveness of a discredited eyewitness are, thus, inconclusive.

The persuasiveness of other types of evidence, such as deoxyribonucleic acid (DNA) evidence, has been less extensively investigated compared to eyewitness evidence. DNA profiling is a technique that allows an individual to be identified by comparing a sample of DNA material from the crime scene, such as hair, blood, or semen, to the DNA of a suspect (Joyce, 1990). A forensic examiner often uses statistical language to present DNA evidence. For example, a DNA forensic examiner is likely to state that the blood found at a crime scene matches the defendant’s blood and that there is only a 2% chance that a randomly selected person from the population would have that DNA pattern or that 1 out of every 170 million to 1.2 billion people would have that DNA pattern.

Although DNA evidence is extremely reliable, it is often ignored by jurors (Goodman, 1992; Smith, Penrod, Otto, & Park, 1996; Wells, 1992). Goodman (1992) studied the extent to which prospective jurors comprehend and use probabilistic evidence when rendering a verdict. Participants read a court transcript describing a homicide. Goodman varied the frequency with which DNA from blood found at the crime scene matched the defendant. Results indicated that jurors’ estimates of guilt were lowest when the likelihood of a random match between the blood found at the crime scene and the blood type of the defendant was greatest (Goodman, 1992). She also found that the jurors recognized that the scientific evidence was significantly valuable, but they “failed to make fine distinctions between probability evidence that were mildly incriminating, moderately incriminating, and strongly incriminating” (Goodman, 1992, p. 371).

When Goodman (1992) compared the weight mock jurors placed on the statistical evidence to a Bayesian analysis, which predicted how much weight jurors should have given that evidence, she found that mock jurors underused the scientific evidence in reaching a verdict. Additional research by Wright, MacEachern, Stoffer, and MacDonald (1996) supported Goodman’s findings and suggested that people do not understand statistical evidence and, therefore, are reluctant to use it when rendering a verdict.

In order to discover if people are reluctant to make guilty decisions when the evidence is based only on statistical evidence, Wells (1992) conducted a series of five experiments. Wells concluded that even students who have a general understanding of statistical information disregard statistical evidence when it is presented.

In their review of jurors’ use of scientific evidence, Cooper, Bennett, and Sukel (1996) argued that most jurors are not able to understand or comprehend legal concepts or scientific data. People who would likely understand this evidence, such as college students, lawyers, and doctors, are often exempt from serving on juries. Theories of persuasion ( Petty & Cacioppo, 1981) suggest that when evidence is complex or difficult to understand, jurors will take the peripheral route to persuasion and base their decision on factors other than the strength of the argument, such as the expert’s credibility. Cooper et al. (1996) tested this hypothesis and discovered that when mock jurors evaluated extremely complex
scientific evidence, they based their verdict on other cues besides the evidence, such as credibility of the witness. The more complex the evidence, the more jurors relied on the credibility of the witness to make a decision in the case. Taken together, these results suggest that when jurors view complex scientific testimony, such as DNA evidence, factors other than the content of the testimony will influence believability.

The purpose of the present experiment was to examine the relative importance of eyewitness evidence and DNA evidence to mock jurors. Participants read an excerpt of a court transcript describing a first-degree murder trial. The transcript contained either eyewitness evidence, DNA evidence, or both types of evidence. We also manipulated the credibility of the testimony (either credible or discredited). After the participants read the transcript, they rendered a verdict in the case and answered several questions concerning understanding of testimony, confidence in verdict, credibility of witness, and persuasiveness of testimony.

Based on the results of previous research (Goodman, 1992; Loftus, 1980; Wells, 1992) and the theory of persuasion espoused by Petty and Cacioppo (1981), we hypothesized that the combination of the eyewitness and DNA evidence would be the most persuasive to jurors. Eyewitness evidence would be more persuasive than DNA evidence, and credible testimony would be more persuasive than discredited testimony.

**Method**

**Participants**

One hundred fifty college students (107 women, 43 men) participated in one of six different experimental conditions. Participants, ages 18 to 23, attended either Ohio Wesleyan University or Earlham College. Most participants received course extra credit in return for their participation.

**Materials and Procedure**

The researcher told participants only that the experiment was about juror’s perception of criminal cases. Participants completed the experiment in groups ranging from 5 to 40 participants. Participants received excerpts from a court transcript and were instructed to read the transcript carefully and to imagine they were real jurors serving on the case. An attorney assisted with the writing of the court transcripts, which were written specially for this study. There were two independent variables: (a) the type of evidence that was being presented, either eyewitness evidence, DNA evidence, or both types of evidence; and (b) the credibility of the testimony, either credible testimony or one that had been discredited through a cross-examination. In the combined condition, in which participants evaluated both types of evidence, the order was counterbalanced; half the participants read the DNA evidence first, whereas the other half read the eyewitness evidence first. Thus, there were six experimental conditions to which participants in each experimental session were randomly assigned. Each experimental condition received a different color court transcript. For example, participants in the eyewitness credible condition received a purple transcript, whereas participants in the combined discredited condition read a blue transcript. Participants read a direct examination as well as a cross-examination in all conditions.

In the eyewitness credible condition, the prosecution presented evidence that an eyewitness saw the defendant running from the scene of the crime holding a knife. In addition, this eyewitness picked the defendant out from a lineup, and during the trial, identified the defendant, David K. Johnson, as the perpetrator. In the cross-examination, little damage was done to the witness’s credibility. The eyewitness testified that he did not need glasses at night, that there was plenty of light in the area, and that he got a good look at the defendant.

In the eyewitness discredited condition, participants read the same prosecution evidence used in the eyewitness credible condition, but in addition, they received a defense attorney’s cross-examination of the eyewitness in which he challenged the credibility of the eyewitness. The defense attorney pointed out that the eyewitness had very poor eyesight and was not wearing his glasses the night of the murder. The defense also presented evidence that stated it was dark the night of the murder and there was poor lighting in the area, which would have made identification of the perpetrator extremely difficult.

In the DNA credible condition, the prosecution presented a DNA forensic examiner who explained what DNA is and the testing process. The DNA expert then stated that he tested two items for this case: blood retrieved from the sidewalk outside the victim’s house and samples from under the fingernails of the victim. The DNA expert testified that the profile from the two DNA items matched that of the defendant, David K. Johnson. The results from the DNA testing were presented in frequencies. Jurors read that these patterns occur in 1 out of every 170 million to 1.2 billion people. In the cross-examination, the forensic examiner testified to the number of hours put into the case and that the lab had no previous errors on the state proficiency tests.

In the DNA discredited condition, participants viewed the direct examination used in the DNA credi-
ible condition, but were then presented with a cross-examination that questioned the reliability of DNA testing. Previous lab errors were presented as well as possibilities that the DNA evidence was contaminated.

In the combined credible condition, participants read a combined court transcript of the eyewitness credible condition and the DNA credible condition. Participants in the discredited combined condition viewed the same materials used in the eyewitness discredited condition combined with the materials used in the DNA discredited condition.

After they had read the designated court transcript, participants gave the researcher their transcript before receiving the questionnaire to ensure that they did not refer back to the transcripts while answering questions. After the transcript was collected, participants completed a questionnaire. Participants rendered a verdict, either guilty or not guilty, and then indicated their confidence in that verdict. Other questions concerned understanding of evidence, credibility of witness, accuracy and reliability of the testimony, and persuasiveness, importance, and influence of testimony. Participants answered questions on a Likert-type scale ranging from 1 to 9 in which 1 was less and 9 was more of the quality being asked. The participants in the eyewitness and DNA conditions answered 18 questions, whereas participants in the combined condition answered 32 questions. Each participant completed the questionnaire individually without knowledge of other participants’ responses. After participants completed all parts of the measures, they were debriefed and thanked for their time.

Results

Preliminary Analyses

In the combined condition, composite scores for each dependent variable were obtained by taking the average of the eyewitness and DNA scores. Each condition’s court transcript was on different color paper; therefore, in order to determine whether the transcript color was a confounding variable, 48 additional participants were administered the eyewitness credible transcript, with 6 participants receiving each of the eight colors. A one-way analysis of variance (ANOVA) compared the effect of color of the court transcript on verdicts, confidence in verdict, and degree of guilt. No significant effects were discovered, indicating that color of transcript had no effect on these dependent variables.

A preliminary analysis also compared the responses of men and women on several of the dependent variables. No significant differences were revealed; therefore, this variable was not examined further.

A two-way ANOVA was conducted as a manipulation check to determine if type of evidence and credibility of testimony affected witness credibility. This analysis ensured that, in fact, there was a significant difference between the credible and discredited conditions. A significant interaction was revealed, $F(2, 144) = 3.127$, $p < .05$, indicating that the credibility of testimony interacted with type of evidence to influence credibility of witness. The manipulation check indicated that the credibility manipulation worked only in the eyewitness condition. Neither the DNA nor the combined condition revealed a significant difference in mean witness credibility between the credible and discredited conditions. Although there was a main effect for credibility of testimony, $F(1, 144) = 23.515$, $p < .01$, this main effect was due entirely to the eyewitness condition. It appears that one cannot discredit a DNA witness. It should also be mentioned that the main effect for type of evidence was significant, $F(2, 144) = 38.556$, $p < .01$. Mock jurors rated the DNA evidence as more credible than the combined or the eyewitness evidence. Participants viewed the DNA evidence as more credible regardless of whether the testimony was discredited or not.

All data were analyzed using two-way ANOVAs to determine if type of evidence and credibility of testimony had an effect on the dependent variables. Post hoc analyses were conducted with least significant difference (LSD) tests.

Guilty Verdicts and Confidence Ratings

Table 1 shows the mean scores and standard deviations for degree of guilt and confidence in verdict for each condition. A significant main effect for type of evidence was found for confidence in verdict and degree of guilt of defendant, $F(2, 144) = 7.028$, $p < .01$, and $F(2, 138) = 2.83$, $p < .01$, respectively. Participants in the eyewitness conditions rated the defendant as less guilty ($p < .05$) and were less confident ($p < .05$) in their verdict than participants in either the DNA only or combined conditions, which did not significantly differ.

Tables 2 and 3 show the results for the variable of guilty verdicts. We conducted chi-square tests of independence to determine if type of evidence and credibility of testimony had an effect on verdict. A significant effect for type of evidence was discovered, $\chi^2(5, N = 150) = 29.76$, $p < .01$, indicating that the DNA condition produced more guilty verdicts than was expected by chance, whereas the eyewitness condition yielded fewer guilty verdicts than was expected by chance. Results confirmed a significant effect for credibility of testimony, $\chi^2(1, N = 150) = 3.86$, $p < .05$. Participants in the credible conditions were more
Effects of DNA and Eyewitness

TABLE 1

Means and Standard Deviations of Ratings of Guilt and Confidence as a Function of Type of Evidence and Credibility of Testimony

<table>
<thead>
<tr>
<th>Dependent variables</th>
<th>Eyewitness</th>
<th>DNA</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Credible</td>
<td>Discredited</td>
<td>Credible</td>
</tr>
<tr>
<td>Degree of guilt</td>
<td>5.43 1.61</td>
<td>4.97 2.12</td>
<td>7.09 1.50</td>
</tr>
<tr>
<td>Confidence in verdict</td>
<td>5.08 2.14</td>
<td>5.29 1.92</td>
<td>6.96 1.31</td>
</tr>
</tbody>
</table>

Note. Higher scores indicate more of the attribute.

likely to vote guilty than participants in the discredited conditions.

Persuasiveness and Importance of Evidence

Table 4 shows the mean scores and standard deviations for persuasiveness of evidence and importance of evidence. Results revealed a significant main effect for credibility of testimony on the variable of persuasiveness, $F(1, 144) = 5.582, p < .05$, suggesting that the more credible the testimony, the more persuasive it was to prospective jurors. A significant main effect for type of evidence was also discovered, $F(2, 144) = 6.137, p < .01$. Post hoc tests showed that eyewitness evidence was less persuasive than the combined and DNA conditions ($p < .05$); however, there was no significant difference between the DNA and the combined conditions.

The two measures, importance of evidence to the verdict and influence of testimony, were combined by adding the two scores to make a composite score for the variable of “importance.” A significant main effect for type of evidence was revealed for the variable “importance,” $F(2, 144) = 10.318, p < .01$. Post hoc analyses showed that the DNA evidence was

TABLE 2

Verdict as a Function of Type of Evidence

<table>
<thead>
<tr>
<th>Type of evidence</th>
<th>Verdict</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Eyewitness</td>
</tr>
<tr>
<td>Guilty</td>
<td>19</td>
</tr>
<tr>
<td>Not guilty</td>
<td>30</td>
</tr>
<tr>
<td>Total</td>
<td>49</td>
</tr>
</tbody>
</table>

TABLE 3

Verdict as a Function of Credibility of Testimony

<table>
<thead>
<tr>
<th>Credibility of testimony</th>
<th>Verdict</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Credible</td>
</tr>
<tr>
<td>Guilty</td>
<td>57</td>
</tr>
<tr>
<td>Not guilty</td>
<td>20</td>
</tr>
<tr>
<td>Total</td>
<td>77</td>
</tr>
</tbody>
</table>

TABLE 4

Means and Standard Deviations of Ratings of Persuasiveness and Importance of Evidence as a Function of Type of Evidence and Credibility of Testimony

<table>
<thead>
<tr>
<th>Dependent variables</th>
<th>Eyewitness</th>
<th>DNA</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Credible</td>
<td>Discredited</td>
<td>Credible</td>
</tr>
<tr>
<td>Persuasive</td>
<td>5.88 2.03</td>
<td>5.46 2.13</td>
<td>6.73 1.28</td>
</tr>
<tr>
<td>Importance</td>
<td>13.04 4.23</td>
<td>12.50 4.93</td>
<td>15.62 2.74</td>
</tr>
</tbody>
</table>

Note. Higher scores indicate more of the attribute.
significantly more important \((p < .05)\) to the verdict than the eyewitness evidence or the combined evidence. No significant difference was found between the eyewitness and combined conditions.

**Understanding of Evidence**

Table 5 shows mean scores and standard deviations for the variables understanding of evidence, technicality of evidence, and difficulty of evidence. A significant main effect for type of evidence was found for the variables understanding of evidence, \(F(2, 144) = 11.321, p < .01\), technicality of evidence, \(F(2, 144) = 15.277, p < .01\), and difficulty of evidence, \(F(2, 144), p < .01\). Post hoc LSD tests revealed that eyewitness testimony yielded significantly \((p < .05)\) more understanding of evidence than both the combined and DNA evidence; however, DNA evidence did not differ from the combined condition. In addition, the DNA evidence was viewed as significantly \((p < .05)\) more technical than both the eyewitness evidence and the combined evidence. The fact that no significant difference was found between the eyewitness and combined evidence implies that mock jurors did not believe the combined evidence to be any more technical than the eyewitness evidence. Consistent with the previous findings, participants found the DNA evidence more difficult to understand than the combined evidence \((p < .05)\), and eyewitness evidence, which was the least difficult to understand.

**Accuracy of Testimony and Witness**

Table 6 shows the means and standard deviations for the variables likelihood of errors and reliability of evidence. A significant main effect for type of evidence on likelihood of errors was found, \(F(1, 144) = 13.421, p < .01\). Participants believed that there was a greater chance that the eyewitness was wrong in his identification than that the lab had made errors during the DNA testing.

We created the variable, reliability of evidence, by adding together the scores for the ratings of reliability of testimony and believability of witness. ANOVAs yielded significant main effects for credibility of testimony, \(F(1, 144) = 14.383, p < .01\), and type of evidence, \(F(2, 144) = 25.837, p < .01\), on the variable of reliability of evidence. Participants viewed the

---

**TABLE 5**

Means and Standard Deviations of Ratings of Understanding of Evidence as a Function of Type of Evidence and Credibility of Testimony

<table>
<thead>
<tr>
<th>Dependent variables</th>
<th>Eyewitness</th>
<th>DNA</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Credible</td>
<td>Discredited</td>
<td>Credible</td>
</tr>
<tr>
<td>Understanding</td>
<td>M 8.08  SD 1.68</td>
<td>M 7.83  SD 1.79</td>
<td>M 6.73  SD 1.25</td>
</tr>
<tr>
<td>Technicality</td>
<td>M 5.48  SD 2.53</td>
<td>M 5.25  SD 2.12</td>
<td>M 7.42  SD 1.24</td>
</tr>
<tr>
<td>Difficulty</td>
<td>M 2.04  SD 1.01</td>
<td>M 1.88  SD .78</td>
<td>M 4.99  SD 2.23</td>
</tr>
</tbody>
</table>

*Note.* Higher scores indicate more of the attribute.

**TABLE 6**

Means and Standard Deviations of Ratings of Accuracy of Testimony and Witness as a Function of Type of Evidence and Credibility of Testimony

<table>
<thead>
<tr>
<th>Dependent variables</th>
<th>Eyewitness</th>
<th>DNA</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Credible</td>
<td>Discredited</td>
<td>Credible</td>
</tr>
<tr>
<td>Errors</td>
<td>M 5.44  SD 2.16</td>
<td>M 5.58  SD 2.13</td>
<td>M 3.62  SD 1.81</td>
</tr>
<tr>
<td>Reliability of evidence</td>
<td>M 11.52  SD 3.89</td>
<td>M 9.10  SD 3.56</td>
<td>M 15.58  SD 1.92</td>
</tr>
</tbody>
</table>

*Note.* Higher scores indicate more of the attribute.
credible testimony as more reliable than the discredited testimony. Participants thought the DNA evidence was more reliable than the combined evidence and the eyewitness evidence. The eyewitness was the least reliable form of evidence \( (p < .05) \). It appears that adding the eyewitness evidence to the DNA evidence to produce the combined condition reduces the judgments of reliability of the DNA evidence.

Aspects of the Witness

Table 7 shows the means and standard deviations for the variables witness intelligence and confidence. The dependent variable intelligence of witness revealed a significant main effect for type of evidence, \( F(2, 138) = 8.444, p < .01 \). Post hoc analyses showed that the witness was seen as significantly \( (p < .05) \) more intelligent in the DNA conditions than in the combined or eyewitness conditions. Jurors viewed the witness in the combined conditions as significantly \( (p < .05) \) more intelligent than the witness in the eyewitness conditions.

The variable confidence of the witness yielded a significant main effect for credibility of testimony, \( F(1, 138) = 10.262, p < .01 \). This main effect means that mock jurors believed the witness was more confident in the credible testimony condition than in the discredited testimony condition.

Discussion

The purpose of the present experiment was to investigate whether certain types of evidence—DNA, eyewitness, or a combination of both—are more persuasive to mock jurors. The experiment also explored whether varying the credibility of the testimony affected persuasiveness. Results indicated that DNA evidence was more persuasive to mock jurors regardless of whether or not it was discredited. DNA evidence received the most guilty verdicts and participants were more confident in their verdict in the DNA conditions than in the combined or eyewitness conditions. The DNA evidence alone was more convincing to jurors than when it was added to the eyewitness evidence. Overall, it appeared that adding the eyewitness evidence onto the DNA evidence to create the combined conditions actually reduced the persuasiveness of the DNA evidence.

The results supported previous research (Kennedy & Haygood, 1992; McCloskey & Egeth, 1983), which indicated that discrediting a witness weakens the influence of that witness in the juror’s mind. The results, however, contradicted the findings of Loftus (1980), who suggested that jurors overestimate the accuracy of eyewitness evidence, even when it has been discredited. The results also contradicted previous DNA evidence research, which suggested that jurors ignore DNA and scientific evidence because it is too complex (Goodman, 1992; Smith et al., 1996; Wells, 1992).

The results of the present research seem to suggest that either jurors are aware of the inconsistencies and unreliability that surround eyewitness evidence and, therefore, do not place a great deal of weight on it when rendering a verdict or that jurors have more faith in DNA evidence than eyewitness evidence. Mock jurors believed the DNA expert witness more than they believed the eyewitness. In addition, participants viewed the eyewitness as less reliable, less accurate, less confident, less intelligent, and less credible than the DNA expert witness. Jurors also perceived the eyewitness evidence to be less influential and important to their decision and to be less persuasive than the DNA evidence or the combined evidence. The findings also suggest that the eyewitness was more likely to be wrong in his/her identification than the lab was to have errors in the DNA testing process. Mock jurors thought to a greater degree that the defendant was guilty in both the credible and discredited DNA conditions followed by
both the credible and discredited combined conditions, than both eyewitness conditions.

The results of this study could have profound implications not only for the psychological world, but for the legal community as well. It may be that jurors will believe an expert DNA witness regardless of whether the testimony is suspect. If this is the case, then a lawyer only needs to present the DNA evidence and it would be believed regardless of its validity. Furthermore, the more credible and intelligent the witness appears to the jurors, the more likely the evidence is to be believed.

Contrary to previous findings, eyewitness testimony does not appear to be persuasive to jurors. Jurors seem to be aware of the inconsistency and unreliability of eyewitness recollection. The results of the current research suggest that eyewitness evidence is not a powerful type of evidence, which contradicts previous research findings (Loftus, 1980). In fact, because jurors are able to take the central route to persuasion when examining eyewitness evidence, the evidence may be scrutinized more than DNA evidence, which allows the juror to take the peripheral route. When jurors take the central route, they examine and analyze the evidence, therefore, jurors would notice the unreliability and inconsistencies that surround eyewitness identification. Because of the technicality of the evidence in the DNA conditions, the mock jurors, as in the study conducted by Cooper et al. (1996), were forced to take the peripheral route to persuasion and, therefore, perceived the DNA expert witness to have greater expertise and credibility than the eyewitness.

There were several limitations of this study. Due to time constraints and lack of funding, written court transcripts were used. The use of a live courtroom setting or the use of a video recording of a trial so that jurors could actually see and hear the testimony would improve this study. The responses of the jurors to reading the evidence might be different than the responses from hearing and seeing the evidence being presented by a real person. It may be that participants are better able to process technical information when reading it than when hearing it from a witness.

In addition, the education level of the participants used in this experiment may have influenced the results. Previous research indicates that the average juror has an eighth-grade education (Cooper et al., 1996). Because we used only college students, who have a much higher education level than the average juror, the results may not generalize to the larger population. In addition, because most college students have a background in science and may have studied DNA, they may have more confidence in the reliability of DNA testing than the typical juror. Another limitation with the sample was that all participants were between the ages of 18 and 23. This age group is the least likely to actually serve on a real jury. Although one study (Kassin & Barndollar, 1992) found that students and nonstudents (adults) do not significantly differ in their sensitivity to factors that affect eyewitness evidence, the present study should be conducted again using both college students and adults to verify that the findings are robust across age groups.

Another limitation of the present study was that participants did not engage in the jury deliberation process. Participants rendered a verdict and completed the questionnaire individually without knowledge of other participants’ responses. Results may have varied if the participants had discussed their verdicts with other mock jurors.

To increase knowledge as to how jurors perceive various types of evidence, several future studies could be conducted. One possible study could address the issue of how DNA evidence is presented. The use of visual aids could affect persuasiveness of the evidence. It would be interesting to investigate how jurors’ understanding of DNA evidence is affected by the way the DNA is presented. In addition, understanding could be influenced by jurors’ previous knowledge of mathematics and science. Jurors who have a vast understanding of mathematics and science may be more prone to trust DNA and believe it is a reliable form of evidence than those jurors who are not as experienced with math and science. Further research conducted in the psychology laboratory can provide additional insight concerning the variables that influence juror decision making.

References


Part-Set Cuing is Due to Strong, Not Weak, List Cues

Research in memory on the part-set cuing paradigm has shown that the presentation of study list items as cues at retrieval impairs recall for the remaining items of the study list (e.g., Rundus, 1973). Subsequent research by Peynircioglu (1987) has indicated that the part-set cuing effect may also occur in nonmemory tasks. In Peynircioglu’s experiment, participants generated words from a larger source word. At retrieval, participants were provided with 0 or 8 exemplars as cues to help generate other words. The results indicated that fewer words were generated in the cued condition than in the noncued condition, suggesting that the mechanism producing the effect was similar to that found in traditional memory tasks (e.g., blocking, strategy disruption). The procedure in the present study was similar to Peynircioglu’s except participants generated exemplars from common conceptual categories (e.g., fruits). The number and popularity of cue words was manipulated. The results indicated that a part-set cuing effect occurred only when popular cues were provided. This finding suggests that the part-set cuing effect is due to the strength of the cue words and thus supports predictions made by blocking theory, which suggests that study list items with higher retrieval strength block access to study list items with weaker retrieval strength.

Monica L. Boice
Gary J. Gargano*
Saint Joseph’s University

In studies of categorical recall, recall for items related to a common conceptual category, findings indicate that the presentation of category labels enhances recall (Tulving & Pearlstone, 1966). However, other research has shown the presentation of study list members has the opposite effect (Brown, 1968). For example, Brown (1968) asked participants to recall as many of the 50 states as possible. Brown found when participants were provided with 25 states, as cues to help them recall the remaining 25, performance on the recall test decreased. Similar research by Rundus (1973) and Roediger (1974), in which participants recalled studied items in the presence of cues, has shown that there is an inverse relation between the number of cues presented and the number of category members recalled. This paradoxical finding in memory research is referred to as the part-set cuing effect.

Several theories have been proposed to explain the part-set cuing effect, including strategy disruption, blocking, and cue overload. Brown (1968) attributed the part-set cuing effect to strategy disruption. According to strategy disruption, the specific retrieval strategy of each participant is disrupted when cues are presented at retrieval. Brown believed each participant has a unique retrieval strategy, and interfering with that strategy, by providing cues, inhibits recall. For example, when participants recalled as many states as they could, they may have done so systematically by using alphabetical order or by using proximity from east to west. The cues disrupt the retrieval strategy because they are not presented in the systematic strategy used by the participant.

Rundus (1973) attributed the part-set cuing effect to blocking theory. According to blocking theory, items compete for access at retrieval. Items that have high retrieval strength inhibit, or block, access for items with weaker retrieval strength. The retrieval strength for an item becomes stronger and more accessible when the item is studied. The accessibility for an item also increases each time the item is retrieved (Bjork & Bjork, 1988). Thus, when some of the study list items are presented as cues, their retrieval strength increases. According to blocking theory, items with high retrieval strength will block

Author note. Monica L. Boice, Saint Joseph’s University; Gary J. Gargano, Saint Joseph’s University.

This research was presented at the annual meeting of the Eastern Psychological Association in March 2000.

Correspondence concerning this article should be addressed to Dr. Gary J. Gargano, Department of Psychology, Saint Joseph’s University, 5600 City Avenue, Philadelphia, PA 19131, or via electronic mail at gargano@sju.edu.

Copyright 2001 by Psi Chi, The National Honor Society in Psychology (Vol. 6, No. 3, 118–122 / ISSN 1089-4136). *Faculty Supervisor
access to items with weaker retrieval strength, making it more difficult to recall noncued items.

Rundus (1973) used category members as cues to determine if conceptually related items would be better cues than category names. Rundus found a decrease in the number of exemplars recalled as the number of exemplars presented as cues increased. Rundus attributed the effect to blocking, because cue words intruded during recall, blocking access to other less accessible words.

According to cue-overload theory (Mueller & Watkins, 1977), the number of retrieval cues has an inhibitory effect on recall. Each occurrence of an item, whether it is presented on the study list or as a cue, increases the number of instances related to the category. As the number of instances increases, the cue becomes overloaded and is less effective at aiding recall. Thus, when participants are presented with study list items as cues, recall will be lower than when zero instances are provided as cues. Because both intralist and extralist cues overload the cue, both will decrease recall.

Mueller and Watkins (1977) manipulated the type of retrieval cue to support cue-overload theory. After studying a list of categorized words, participants were presented with either intralist or extralist cues. Intralist retrieval cues consisted of items that were included in the original study list. Extrinsic retrieval cues consisted of items that were not included in the original study list but were related to the list items. The results supported predictions of cue-overload theory because a part-set cuing effect was found when both intralist and extralist cues were used.

Peynircioglu (1987) has indicated that the part-set cuing effect may not be limited to memory tasks. Peynircioglu provided participants with a nonmemory task of generating words from the letters of a larger source word. Participants were provided with zero or eight words as cues to help generate other words. Results indicated that fewer words were generated in the cued condition compared to the noncued condition. Peynircioglu concluded that a similar mechanism found in traditional memory experiments (e.g., blocking, strategy disruption, cue overload) may be responsible for the effect.

Although there is some evidence for each theory that offers explanations for part-set cuing, the mechanism producing the effect is still not fully understood. The purpose of the present study is twofold. Research has indicated some support for each theoretical explanation of the part-set cuing effect, but few studies have contrasted the predictions of these theories. The present study tested predictions made by blocking, strategy disruption, and cue-overload theories. Additionally, the present study further investigated whether part-set cuing can be generalized to nonmemory tasks.

In the present study, participants generated category members when provided with a category name and, in some cases, category members as cues. We manipulated the number and popularity of cues provided on the recall test. Participants were provided with either zero or eight popular, unpopular, or random category exemplars as cues. The dependent measure was the number of noncue items generated from each category. A part-set cuing effect will be found if recall is lower in the cued conditions than in the uncued conditions.

Blocking, strategy disruption, and cue-overload theories make contrasting predictions based on different theoretical explanations of the part-set cuing effect. According to strategy disruption, cues should be disruptive regardless of their popularity, because each instance of a cue will disrupt the participants’ unique retrieval methods (Brown, 1968). According to blocking theory, popular cues should be more disruptive than unpopular cues, because stronger, more accessible items will block access to less accessible items (Rundus, 1973). Cue-overload theory predicts lower recall as the number, not type, of exemplars presented as cues increases (Watkins, 1975). Thus, the popularity of the cue should not affect recall.

**Method**

**Participants**

One hundred eighty undergraduate students, ranging in age from 19 to 22 years, at Saint Joseph’s University volunteered to participate in the study as an option for fulfilling a course requirement. All participants were debriefed and given credits for their participation.

**Design**

The design was a 3 (popularity of cue) × 2 (number of cues) between-subjects factorial design. One variable was the popularity of exemplars provided (popular, unpopular, random). The other variable was the number of cues provided (zero or eight). The dependent measure was the number of words correctly recalled. Participants were randomly assigned to one of the six groups.

**Materials**

A separate norming study was conducted to determine the most frequently generated responses for each category. Fifteen conceptual categories were selected from the Battig and Montague (1969) category norms (see Appendix A). A similar group of
120 participants were presented with each category name and asked to generate as many exemplars as possible. Based on the results of the norming study, we selected the eight most frequent responses for each category as the popular cues; we used the eight least frequent responses as the unpopular cues (see Appendix B). The eight random cues were items that were randomly chosen from all of the possible responses.

The test form consisted of 15 pages, each page having one of the category names at the top of the page for the zero-cue conditions. In the cued conditions, depending on the condition for popularity, eight cues were listed below the category name. For example, in the popular-cue condition on each page of the test form, the category name was followed by the eight most popular responses selected as cues (see Appendix C). We presented all the categories and cues in the same random order to all the participants.

### Procedure

The participants were randomly placed into one of the six experimental groups receiving either zero or eight popular, unpopular, or random cues. We told the participants that on each page of the test booklet they would see the name of a conceptual category. Their task was to generate as many instances of the category as possible. In the cued conditions, participants were further instructed that they were provided with cues, in order to help them generate other category members. The participants were allowed 1 1/2 min to recall words from each category. Participants were debriefed following the experiment.

### Results

A 3 × 2 factorial analysis of variance revealed a significant interaction between the type of cue and number of cues provided at recall, $F(2, 174) = 42.84$, $MSE = 23217.92$, $p < .05$. To determine if a part-set cuing effect occurred, a simple main effects analysis was conducted for each cue condition (popular, unpopular, and random) by the number of cues provided at recall (0 or 8). A part-set cuing effect was found in the popular-cue condition, $F(1, 174) = 197.94$, $MSE = 107272.82$, $p < .05$, and the random-cue condition, $F(1, 174) = 18.33$, $MSE = 9933.07$, $p < .05$, but not in the unpopular-cue condition ($p > .05$).

A comparison of the mean scores indicated higher recall when no cues were provided ($M = 157.57$) than when cues were provided ($M = 73.00$) in the popular-cue condition. Similar findings were found in the random-cue condition, with higher recall when no cues were provided ($M = 144.30$) than when cues were provided ($M = 118.57$). Means and standard deviations for each condition are displayed in Table 1.

To determine the source of the part-set cuing effect, a second simple main effects analysis was conducted for the number of cues provided at recall (0, 8) for each cue condition (popular, unpopular, random). The results indicated that an equal number of words were recalled for the zero-cue condition ($p > .10$) but not for the eight-cue condition, $F(2, 174) = 68.95$, $MSE = 37365.74$, $p < .05$. A Tukey post hoc test indicated a linear trend with recall being significantly higher ($p < .05$) for the unpopular-cue condition ($M = 142.47$) than for the random-cue condition ($M = 118.57$) and the popular-cue condition ($M = 73.00$). The decreased recall in the popular-cue condition is due to the fact that participants cannot recall the more accessible popular words, because they are provided as cues. These items, in turn, then block access to less accessible items. This finding is supported by a comparison of the frequency of the types of words recalled. Participants recall an average of 78.28 popular words across all conditions, whereas they recall an average of only 3.51 unpopular words across all conditions.

### Discussion

The results of the Cue Type × Number of Cues interaction indicated that a part-set cuing effect was found in the popular- and random-cue conditions.
but not in the unpopular-cue condition. This finding does not support the predictions of strategy disruption and cue overload, because these theories predict the popularity of the cues would have no effect on recall.

Strategy disruption predicted that recall would decrease when cue words were provided, regardless of their popularity, because any type of cue disrupts the participants’ unique retrieval methods. Strategy disruption could not account for the effect because a part-set cuing effect was found in the popular- and random-cue conditions, but not in the unpopular-cue condition.

Likewise, cue-overload theory could not account for the findings. Cue-overload theory predicted that recall would decrease as the number of items associated with the category increased. According to this theory, a part-set cuing effect should be observed in all cue conditions. Cue-overload could not explain the effect because a part-set cuing effect was only found in the popular and random conditions, not in the unpopular condition.

Blocking theory, however, can account for the findings. Blocking theory predicted that popular cues would be more disruptive than unpopular cues, because more accessible items block access to less accessible items. The results indicated that when the more accessible popular words were presented as cues, retrieval for the less accessible words decreased. Also, because some cue words in the random-cue condition were popular, a smaller part-set cuing effect occurred.

The results suggest that when participants were unable to generate responses with high retrieval strength in the popular-cue condition, because they were provided as cues, recall was suppressed. These strong cue items then blocked access to weaker, less accessible items. However, in the unpopular-cue condition, the weaker, less accessible items did not block access to stronger, more accessible items. When comparing the frequency of the type of words recalled across all conditions, participants recalled more popular words than unpopular words. It appears as though participants retrieve the most popular responses first. Then, participants either stop recalling additional words, or the recall of popular cues blocks access to less accessible, or unpopular responses.

The findings of the present study support those of Peynircioglu (1987), that part-set cuing can be found in nonmemory tasks. In addition, a memory mechanism seems to be responsible for the effect. However, because Peynircioglu did not state whether she used strong or weak cues, it is hard to generalize her findings to a memory mechanism as she claims.

References


APPENDIX A

List of All Category Names Used

<table>
<thead>
<tr>
<th>Furniture</th>
<th>Flowers</th>
<th>Vegetables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fabrics</td>
<td>Fruits</td>
<td>Weather Phenomenon</td>
</tr>
<tr>
<td>Birds</td>
<td>Trees</td>
<td>Carpenters’ Tools</td>
</tr>
<tr>
<td>Professions</td>
<td>Sciences</td>
<td>Fish</td>
</tr>
<tr>
<td>Sports</td>
<td>Insects</td>
<td>Musical Instruments</td>
</tr>
</tbody>
</table>

APPENDIX B

Example of Popular, Unpopular, and Random Cues for the Tree Category

<table>
<thead>
<tr>
<th>Popular cues</th>
<th>Unpopular cues</th>
<th>Random cues</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pine</td>
<td>Cottonwood</td>
<td>Beech</td>
</tr>
<tr>
<td>Maple</td>
<td>Balsam</td>
<td>Hemlock</td>
</tr>
<tr>
<td>Redwood</td>
<td>Gingko</td>
<td>Spruce</td>
</tr>
<tr>
<td>Oak</td>
<td>Holly</td>
<td>Locust</td>
</tr>
<tr>
<td>Dogwood</td>
<td>Mahogany</td>
<td>Walnut</td>
</tr>
<tr>
<td>Willow</td>
<td>Mimosa</td>
<td>Olive</td>
</tr>
<tr>
<td>Birch</td>
<td>Mulberry</td>
<td>Sycamore</td>
</tr>
<tr>
<td>Palm</td>
<td>Poplar</td>
<td>Sequoia</td>
</tr>
</tbody>
</table>

APPENDIX C

Example of Test Form Used in the Popular-Cue Condition

*Instructions:*

Please write down as many members of the category listed below.

TREES
Oak
Pine
Maple
Willow
Redwood
Birch
Dogwood
Palm
Human Sex Differences in Aggression
Within an Evolutionary Model

Barbara J. Hagenah
Christopher Heaps*
Eugene Gilden
Michael Roberts
Linfield College

This investigation used evolutionary theory and social learning theory to predict sex differences in human aggression. One hundred fifteen students estimated their frequency of aggression in the following contexts: (a) compete to attract members of the opposite sex, (b) co-opt resources from others, and (c) negotiate status hierarchies. We assessed aggression type in 3 categories: direct physical, direct verbal, and indirect. We determined sex role by using the Bem Sex-Role Inventory (Bem, 1978). The social model predicts that men will engage more frequently than women in all types of aggression. The evolutionary model predicts that men and women will engage in similar frequencies of total aggression but differ in the type they use; our data supported this prediction. These findings were more consistent with the evolutionary model, indicating that this theory provides a sound basis for predicting and understanding sex differences in aggression.

A

GRESSION IS A UNIVERSAL HUMAN BEHAVIOR that psychologists define as “any action undertaken with the apparent intent of causing physical or psychological harm” (Burbank, 1987, p. 72); it is generally used to gain resources or other advantages. The primary model used to explain aggression is social learning theory. According to this theory, individuals are taught to be aggressive through the influence of surrounding social and environmental conditions (Feldman, 2001). In addition, sex differences in aggression are a result of the different social roles that each sex acquires, which are regulated by social norms (Eagly, 1987). Eagly (1987) explains that the male gender role typically “encourages varying forms of aggression” (p. 72) whereas the female gender role has “little emphasis on aggressiveness” (p. 72). A social learning model argues that these preferences dictate aggressive patterns in individuals.

Evolutionary theory provides another model for understanding aggression. It uses biological, sociological, and historical factors to explain differences in aggression between the sexes and across cultures. According to evolutionary theory, aggression is not simply a learned phenomenon, nor is it an inevitable instinctual drive. Rather, aggression often has adaptive benefits and is sensitive to social contexts that have the potential to evoke aggressive behavior (Buss & Shackelford, 1997). Though behavior is affected by both learned social and underlying biological attributes, it is only in recent years that researchers have considered both attributes simultaneously. We feel that exploring aggression using evolutionary theory will add to our understanding of it. For this reason, the majority of the introduction will be devoted to this perspective.

Tooby and Cosmides (1992) discuss guidelines for approaching psychology from an evolutionary perspective. They argue that humans have specialized rather than generalized information-processing mechanisms, and these mechanisms are context-specific. To explain a human behavior, the researcher must look at how it is related to an adaptive problem.

Author note. This article is based on the undergraduate research work of Barbara J. Hagenah, Department of Psychology, Linfield College. Christopher Heaps and Eugene Gilden, Department of Psychology, and Michael Roberts, Department of Biology, Linfield College, formed the honors thesis committee and supported the development of this study. A poster of this study was presented at the annual meeting of the Western Psychological Association Convention on April, 15, 2000, in Portland, Oregon.

Correspondence concerning this article should be addressed to Barbara J. Hagenah, c/o Christopher Heaps, Department of Psychology, Linfield College, McMinnville, OR 97128.
and consider the ancestral environment in which natural and sexual selective pressures influenced human behavior (Tooby & Cosmides, 1992). Natural selection is a process that controls variation of physical, physiological, or behavioral traits within a population, and influences life, death, and reproduction. Traits having a genetic basis spread if they are more favorable to survival and reproduction in the face of selection pressures. Sexual selection is the specific process that selects traits influencing reproductive success. The two main tasks involved in reproductive success are finding and attracting mates and providing care to offspring (Geary, 1998). The operation of these forces on human behavior means that the variables we observe in cultures and individuals are “the product of a common, underlying evolved psychology, operating under different circumstances” (Tooby & Cosmides, 1992, p. 45). Moreover, these selective pressures make testable predictions about sex differences in aggression (Buss & Shackelford, 1997).

The evolutionary model argues that aggression has been an adaptive behavior for both sexes in certain situations. The model proposes that men and women experienced the pressures of sexual selection differently in our ancestral past (Barkow, Cosmides, & Tooby, 1992; Buss, 1995; Geary, 1998) and that these differences account for consistent sex differences in type of aggression (Björkqvist, Österman, & Kaukiainen, 1992; Burbank, 1987; Campbell, Sapochnik, & Muncer, 1997; Daly & Wilson, 1988; Eagly & Steffen, 1986). Men and women differ according to the amount of effort and the strategies used in approaching mating and parenting. Specifically, the amount of parental energy that is required to successfully reproduce and rear offspring differs greatly between the sexes. Primarily due to gestation and postnatal care, women must invest more resources in each offspring than men. Men do not have the same gestation requirements, and although male investment of resources after birth typically increases survival of offspring, the amount of effort and resources expended by men is less than the amount expended by women. These factors influence behavioral differences between the sexes (Geary, 1998). As Buss (1995) notes, “Strong sex differences occur reliably in domains closely linked with sex and mating, precisely as predicted by psychological theories based on sexual selection” (p. 166).

Björkqvist, Lagerspetz, and Österman (1992) provide an inclusive model for aggression that includes assessment of three types of aggression: direct physical (e.g., hit, kick), direct verbal (e.g., yell, insult), and indirect (e.g., gossip, ostracize). Variation in the type of aggression men and women use would be related to sex differences in amount of parental energy needed for rearing successful offspring. The reproductive efforts required by women will discourage them from using direct physical aggression because the risk of injury would jeopardize biological resources crucial to success in bearing and rearing offspring (Daly & Wilson, 1988). Therefore, women who utilize indirect aggression could benefit from aggressing with less associated risk of physical injury. Men, having a lower degree of parental investment, would not have the same adaptive pressure to avoid physical aggression (Daly & Wilson, 1988). Consequently, men will employ direct physical aggression relatively frequently. Verbal aggression would be used by both sexes with similar frequency because general verbal ability has similar adaptive advantages for men and women, and a selective pressure does not likely exist for one sex to utilize direct verbal aggression more than the other (Geary, 1998).

Another necessary consideration for understanding aggression is that aggressive behaviors can be understood as “context-sensitive solutions to particular adaptive problems of social living” (Buss & Shackelford, 1997, p. 605). Aggressive behavior may be present when such contexts emerge. Buss and Shackelford (1997) describe seven contexts or situations in which aggression elicits adaptive benefits. These contexts are: intrasexual (same-sex) competition, co-opting the resources of others, negotiating status and dominance hierarchies, deterring long-term mates from infidelity, aggressing toward unrelated children, defense against attack, and deferring rivals from future aggression. These contexts can be better understood when divided into three broader categories based on function.

The first category includes intrasexual competition, co-opting resources from others, and negotiating status and dominance hierarchies. We were interested in these categories for two reasons. First, both men and women gain important benefits from having greater resources, status, or better mates. For this reason, differences in the type of aggression employed to reach these goals should relate to the differing strategies associated with reproduction and not to differences in relative importance of the contexts. Second, the contexts are important social situations underlying mate choice and retention prior to forming long-term attachments or bearing children. Thus, these situations are particularly appropriate for measurement in a college-aged population.

A second category includes the contexts of deterring long-term mates from infidelity and aggression toward unrelated children. We did not include these contexts in this study because they are related
to situations that occur after mating and, therefore, are not generally appropriate for measurement in a college-aged population. A third category includes defense against attack and deterring rivals from future aggression. These contexts deal with reactions either to other’s aggression or with an individual’s aggression in the past. These contexts are removed from the initial aggressive intent, and for this reason we did not include them in this study.

Predictions about sex differences in aggression will differ depending on whether a social learning model or an evolutionary model is used to make the predictions. In the contexts of intrasexual competition, co-opting resources, and negotiating status and dominance hierarchies, a social learning model predicts that overall aggression will be higher for men than women. This prediction is made because the model proposes that each sex has acquired a role in how each should aggress, and these patterns should exist in each context. Direct physical aggression should show the greatest difference, men engaging most frequently, but men should also show more indirect and direct verbal aggression than women (Eagly, 1987).

An evolutionary model predicts that overall aggression in each of these contexts may have similar benefits for both men and women. For this reason, in contrast to the social model, this model predicts that the amount of aggression for both sexes will be similar in each context and across all contexts. However, in light of sex differences in selective pressures associated with reproduction and parental investment, this model proposes that differences will occur in the type of aggression men and women use in these contexts. It predicts that men will consistently engage in more direct physical aggression than women in each context, and utilize lower levels of indirect aggression. It predicts that women will be more indirectly aggressive than men in each context. This model predicts that men and women will have similar levels of high direct verbal aggression in all contexts.

Utilizing evolutionary theory to explain sex differences implies that the differences are not arbitrarily socialized. A sex role inventory that assesses masculinity, femininity, and androgyny will provide insight to participants’ perceptions of their gender identity and attempts to account for biological sex as a predictor of aggression type independently of learned gender roles. Social learning theory, an approach that explains sex differences in aggression by differing learned gender roles, suggests that masculine and feminine gender roles will have an influence on both men’s and women’s aggressive patterns as predicted by the social model. However, an evolutionary model would predict that regardless of sex role, men and women would behave as predicted by the evolutionary model, because sex role does not change the reproductive differences in men and women.

To test these predictions we administered a questionnaire to college-aged men and women assessing reported aggressive frequency in three types of aggression: direct physical, direct verbal, and indirect, in the three contexts of intrasexual competition, co-opting resources, negotiating status and dominance hierarchies. We used the Bem Sex-Role Inventory (Bem, 1978) to assess participants’ perceived sex role.

Method

Participants

One hundred and fifteen college students (49 men, 66 women; mean age = 19.46, SD = 1.24) volunteered to participate. Participants responded to publicity posted in the psychology department and campuswide.

Materials

We adapted the Direct and Indirect Aggression Scale (DIAS) questionnaire from Björkqvist, Lagerqvist, et al. (1992) to assess frequency in aggression. Adaptations required the modification of pronouns to be expressed in the first person. Aggression was defined in three categories: direct physical (e.g., hit, kick), direct verbal (e.g., yell, insult), and indirect (e.g., gossip, ostracize). There were 24 questions (7 physical, 5 verbal, 12 indirect), and participants rated frequency of aggressive behaviors on a scale ranging from 0 (never) to 4 (very often). Appendix A shows the modified version of the DIAS. We included a description of three contexts in which aggression could confer adaptive benefits (Buss & Shackelford, 1997), and reported frequency of aggression was considered in each context. The contexts included intrasexual competition, co-opting resources, and negotiating status; these contexts were counterbalanced. Appendix B shows the context descriptions that were used for this study.

We used the Bem Sex-Role Inventory (Bem, 1978) to assess masculine, feminine, androgynous, and undifferentiated gender types. Sixty gender-descriptive words are listed of which there were 20 masculine (e.g., forceful, willing to take risks), 20 feminine (e.g., affectionate, sensitive to needs of others), and 20 neutral (e.g., moody, reliable) items. Participants self-assessed each item on a scale ranging from 1 (never or almost never true) to 7 (always or almost always true). Masculine sex roles result from a high masculine and a low feminine score. Feminine sex roles result from
a low masculine and high feminine score. Androgynous sex roles result from both high masculine and feminine scores. Undifferentiated is defined as having a low masculine and feminine score such that a gender type cannot be assigned.

**Procedure**

We conducted testing in 4-hr blocks for 3 days; participants came at their convenience. Each participant responded to the three context descriptions and self-reported the likelihood of their aggression for each question in the last 4 years. Following the aggression questionnaire, participants filled out the Bem Sex-Role Inventory.

**Results**

A 2 (sex) × 3 (evolutionary context) × 3 (aggression type) analysis of variance (ANOVA) revealed that men and women aggress with equal frequency across contexts. Table 1 shows the ANOVA. A significant Sex

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex (S)</td>
<td>1.536</td>
<td>1</td>
<td>1.536</td>
<td>0.990</td>
<td>0.322</td>
</tr>
<tr>
<td>Context (C)</td>
<td>9.810</td>
<td>2</td>
<td>4.905</td>
<td>20.769</td>
<td>0.001</td>
</tr>
<tr>
<td>Aggression type (A)</td>
<td>133.523</td>
<td>2</td>
<td>66.761</td>
<td>198.051</td>
<td>0.001</td>
</tr>
<tr>
<td>SC interactions</td>
<td>0.323</td>
<td>2</td>
<td>0.161</td>
<td>0.683</td>
<td>0.506</td>
</tr>
<tr>
<td>SA interactions</td>
<td>4.201</td>
<td>2</td>
<td>2.101</td>
<td>6.232</td>
<td>0.002</td>
</tr>
<tr>
<td>AC interactions</td>
<td>10.875</td>
<td>4</td>
<td>2.719</td>
<td>24.945</td>
<td>0.001</td>
</tr>
<tr>
<td>SCA interactions</td>
<td>0.507</td>
<td>4</td>
<td>0.127</td>
<td>1.163</td>
<td>0.326</td>
</tr>
<tr>
<td>Error</td>
<td>173.810</td>
<td>112</td>
<td>1.522</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>445.024</td>
<td>1</td>
<td>445.024</td>
<td>286.764</td>
<td>0.0</td>
</tr>
</tbody>
</table>

**TABLE 2**

Means and Standard Deviations of Aggression Frequency Scores for Men and Women According to Contexts and Aggression Types

<table>
<thead>
<tr>
<th>Context</th>
<th>Direct physical</th>
<th></th>
<th></th>
<th>Indirect</th>
<th></th>
<th></th>
<th>Context total</th>
<th>Overall total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Men</td>
<td>Women</td>
<td>Men</td>
<td>Women</td>
<td>Men</td>
<td>Women</td>
<td>Men</td>
<td>Women</td>
</tr>
<tr>
<td>Intrasexual</td>
<td>M</td>
<td>.20</td>
<td>.02</td>
<td>.109</td>
<td>.78</td>
<td>.102</td>
<td>.113</td>
<td>.80</td>
</tr>
<tr>
<td>competition</td>
<td>SD</td>
<td>.46</td>
<td>.15</td>
<td>.69</td>
<td>.62</td>
<td>.63</td>
<td>.75</td>
<td>.52</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>.23</td>
<td>.13</td>
<td>.87</td>
<td>.78</td>
<td>.55</td>
<td>.62</td>
<td>.52</td>
</tr>
<tr>
<td>Co-opting</td>
<td>SD</td>
<td>.33</td>
<td>.23</td>
<td>.72</td>
<td>.71</td>
<td>.53</td>
<td>.57</td>
<td>.44</td>
</tr>
<tr>
<td>resources</td>
<td>M</td>
<td>.24</td>
<td>.08</td>
<td>1.21</td>
<td>1.05</td>
<td>.92</td>
<td>1.06</td>
<td>.78</td>
</tr>
<tr>
<td>Negotiating</td>
<td>SD</td>
<td>.45</td>
<td>.17</td>
<td>.76</td>
<td>.77</td>
<td>.65</td>
<td>.77</td>
<td>.51</td>
</tr>
<tr>
<td>status</td>
<td>M</td>
<td>.23</td>
<td>.08</td>
<td>1.06</td>
<td>.87</td>
<td>.83</td>
<td>.93</td>
<td>.70</td>
</tr>
<tr>
<td>Aggression</td>
<td>SD</td>
<td>.38</td>
<td>.13</td>
<td>.60</td>
<td>.59</td>
<td>.51</td>
<td>.60</td>
<td>.44</td>
</tr>
<tr>
<td>type total</td>
<td>M</td>
<td>.14</td>
<td>.95</td>
<td>.89</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall total</td>
<td>SD</td>
<td>.27</td>
<td>.60</td>
<td>.56</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: All superscripts indicate $p < .01$. Comparisons a, b, c, and d are illustrated in Figures 1–4, respectively.
HUMAN SEX DIFFERENCES IN AGGRESSION  □  Hagenah, Heaps, Gilden, and Roberts

× Aggression Type interaction was found in which direct physical aggression was more frequent among men than women in the context of intrasexual competition, $t(113) = -2.9, p < .004$, co-opting resources, $t(113) = -2.04, p < .043$, negotiating dominance and power hierarchies, $t(113) = -2.73, p < .007$, and over all contexts, $t(113) = -2.99, p < .003$. Figures 1, 2, 3, and 4 show the mean frequency scores of direct physical aggression in each context and over all contexts for men and women. Men and women were similar in their level of indirect aggression. No sex differences in frequency of verbal aggression were found. Table 2 shows the mean and standard deviation scores for aggression types in each context, as well as scores for total aggression type over all contexts, for men and women.

There was also a significant effect for aggression type and context in which direct verbal aggression was more common in the contexts of achieving status/dominance than it was in the contexts on intrasexual competition, $t(114) = -.204, p < .001$, and co-opting resources, $t(114) = -.299, p < .001$. Indirect aggression was also more common in the contexts of intrasexual competition, $t(115) = .48, p < .001$, and
status/dominance, $t(115) = -0.413$, $p < .001$, than in the context of co-opting resources.

A one-way ANOVA and Tukey’s test showed there were no significant interactions in which masculinity, femininity, or androgyny was a significant predictor of how men or women of each gender type would behave aggressively. Undifferentiated was not considered in the analysis because this group does not reveal sex role.

**Discussion**

Our results indicate that men and women report aggression with similar frequency across the contexts of intrasexual competition, co-opting resources, and negotiating status and dominance hierarchies. This pattern of results is consistent with the evolutionary model’s prediction and not with that of the social learning model. Men did report engaging in direct physical aggression significantly more than women, as predicted by both models. Men and women reported engaging with similar frequency in indirect aggression. This result was not predicted by either model. Men and women reported similar frequencies of direct verbal aggression as predicted by the evolutionary model and not by the social learning model. Because no predictions were made from either model regarding significant interactions between context and aggression type, these findings, although interesting, do not add clarification to sex differences in aggression. In addition, the interaction between context and aggression type does not directly lend support to one model or the other. For these reasons, these results will not be a part of the discussion.

As suggested by the evolutionary model, our results indicate that sex role identity is not a significant predictor of how men and women will report frequency of aggression or type of aggression. Regardless of whether men or women were masculine or feminine, these gender types did not influence how each sex would report aggression. By comparing the predictions each model makes regarding men’s and women’s aggressive behaviors, we suggest that an evolutionary model better explains our data than the social learning model. For this reason, the remainder of the discussion will examine an understanding of the data from an evolutionary perspective and its implications for understanding human behavior within this framework.

It is important to note again that frequency of aggression in this study was determined by self-report, and this should be taken into account in interpreting the results. Also, the mean frequencies of aggression, calculated on a scale from 1 to 4, did not go above 2, resulting in a population that did not report high frequencies of aggression. Overall aggression in our population was low, but the emphasis of this study was to examine sex differences in aggressive behavior, and these differences were significant.

Because the evolutionary model is based on both biological and social attributes of humans, cross-cultural comparisons are important parts of the evolutionary approach. Using evolutionary theory to interpret our data in conjunction with other studies yields important implications for understanding aggression. First, both men and women reported engaging in a similar amount of aggression. This result suggests that aggression is beneficial to both sexes in many social situations, and that the contexts of intrasexual selection, co-opting resources, and negotiating status and power hierarchies are situations in which aggression may be adaptive for both men and women.

Men and women differed in the frequency with which they reported to engage in the specific types of aggression, and the differences were consistent across the three contexts. In agreement with our results, other studies indicate that in women, physical aggression is less frequent than other types of aggression (Burbank, 1987; Buss & Shackelford, 1997; Eagly & Steffen, 1986). This pattern suggests a tendency for women to avoid direct physical aggression, which is consistent with sexual selection theory in that high amounts of maternal care given to offspring require that women minimize their risk of physical injury (Buss & Shackelford, 1997). In a review of paternal investment, Geary (1998) demonstrated that cross-culturally and historically the mortality of mothers greatly increased the likelihood of the child’s mortality. Thus, the constraints of the female reproductive strategy may have led to a selective pressure for women to engage in low amounts of physical aggression.

In this study, indirect aggression was the type most frequently reported by women, suggesting that there may be a selective advantage for women to use indirect aggression. Other studies have shown that women use significantly more indirect aggression than men (Björkqvist, Österman, et al., 1992; Campbell et al., 1997). High levels of parental care that characterize the female reproductive strategy seem to be strongly connected to utilizing indirect methods of aggression. Indirect aggression allows women to gain benefits such as mates, resources, or status, while simultaneously decreasing their risk of physical injury.

Women also report high frequencies of direct verbal aggression. Eagly and Steffen (1986) found that women frequently use a form of aggression that causes psychological harm (i.e., aggressive acts that are not
direct verbal and indirect aggression. Although more amounts of physical aggression and low levels of type of culture, and men in this society show very high survival is low (Geary, 1998). Chagnon and Bugos’s observations of the Yanomamö indicate this. In societies characterized by polygamous systems with high incidence of warfare, raiding, and male social displays in which physical aggression less than they use direct verbal and indirect aggression. Male investment is higher in societies that are industrial or stratified, where monogamous relationships are predominant (Geary, 1998). Examples of such societies include the United States (in the present study), Finland (Björkqvist, Österman, et al., 1992), and England (Campbell et al., 1997). These studies showed that men use physical aggression more than women, direct verbal and indirect aggression were the types of aggression most frequently reported by men in this study. Consideration of cultural differences in paternal investment can provide understanding of the relation between the amount of male physical aggression relative to the amount of direct verbal and indirect aggression. In contrast to the consistent amount of maternal investment given to offspring, the amount of paternal investment differs between cultures (Geary, 1998). We suggest the differences in amount of paternal investment relate to type of aggression men most frequently use.

In societies with high paternal investment, we expect to see lower levels of direct physical aggression and higher levels of direct verbal and indirect aggression. In contrast, in societies with low paternal investment, we expect higher levels of male direct physical aggression and lower levels of direct verbal and indirect aggression. Male investment is higher in societies that are industrial or stratified, where monogamous relationships are predominant (Geary, 1998). Examples of such societies include the United States (in the present study), Finland (Björkqvist, Österman, et al., 1992), and England (Campbell et al., 1997). These studies showed that men use physical aggression less than they use direct verbal and indirect aggression. In cultures characterized by polygamous systems with high incidence of warfare, raiding, and male social displays in which physical aggression could lead to higher status or numerous mates, direct paternal investment needed to ensure offspring survival is low (Geary, 1998). Chagnon and Bugos’s (1979) observations of the Yanomamö indicate this type of culture, and men in this society show very high amounts of physical aggression and low levels of direct verbal and indirect aggression. Although more work is needed to understand aggressive behavior as a function of parental investment, the studies discussed indicate that this type of relation may apply across cultures.

We view this research as an initial attempt to encourage the value of using an evolutionary framework for understanding human behavior in general and aggression in particular. Further work that approaches aggression in relation to differences in male and female reproductive strategies is a worthwhile consideration. However, to make strong claims about human aggression, direct testing of actual behaviors is necessary. In order to better understand this relation, we suggest that the continuation of cross-cultural research in conjunction with defining aggression as a potentially adaptive behavior that can manifest in either direct or indirect form, will further our understanding of human aggression.

This research has barely touched on the idea of testing predictions made by the evolutionary model with predictions made by a social model. To comprehensively test one model against the other, many factors such as participants’ beliefs about their behavior, other sex role inventories, and more diverse populations are necessary to thoroughly examine the two models. It is apparent that social and cultural factors will shape both individual and societal manifestations of behavior. However, our work suggests that these manifestations cannot be understood without reference to evolutionary theory, and that evolutionary theory provides a powerful basis for understanding general patterns of male and female aggression. Recognizing the trends that exist based on evolutionary selective pressures will aid scientists in understanding a complete picture of the nature of human aggression.

Reference


APPENDIX A

Modified DIAS

Answer the questions by circling the number which seems to tell about your behavior in the closest way.

0 = never, 1 = seldom, 2 = sometimes, 3 = quite often, 4 = very often

1. Hit the person?
2. Shut the person out of the group?
3. Yell at or argue with the person?
4. Become friends with another as a kind of revenge?
5. Kick the person?
6. Ignore the person?
7. Insult the person?
8. Gossip about the person you are angry with?
9. Trip the person?
10. Tell bad or false stories about the person?
11. Say that you are going to hurt the person?
12. Plan secretly to bother the person?
13. Shove the person?
14. Say bad things behind the person’s back?
15. Call the person names?
16. Say to others, “Let’s not be with him/her!”?
17. Take things from the person?
18. Tell the person’s secrets to a third person?
19. Tease the person?
20. Write small notes where the person in criticized?
21. Push the person down on the ground?
22. Criticize the person’s hair or clothing?
23. Pull at the person?
24. Try to get others to dislike the person you are angry with?

Note: Physical aggression is indicated in numbers 1, 5, 9, 13, 17, 21, and 23. Verbal aggression is indicated in numbers 3, 7, 11, 15, and 19. Indirect aggression is indicated in numbers 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, and 24.
APPENDIX B

Context Descriptions

1. **Context description of intrasexual competition:**
   We are interested in the things people do to make other members of their own sex seem *undesirable* to members of the opposite sex. For each of the following items, estimate how often (in approximately the last four years) that you have used each of the methods to make other members of your own *sex undesirables* to members of the opposite sex.

2. **Context description of co-opting resources from others:**
   We are interested in the things that people do to get "resources" from other people. These resources could be money, favors, objects—anything you might want that someone else has. For each of the following items, estimate how often (in approximately the last four years) that you have used each of these methods to get something you wanted from another person or group of people.

3. **Context description of negotiating status or dominance hierarchies:**
   We are interested in the things that people do to increase their status or dominance—how they make themselves "look better" or "seem better," how they may try to appear tougher, smarter, in more control of people or situations, etc. For each of the following items, estimate how often (in approximately the last four years) that you have used each of these methods to increase your status or dominance.
Sincere appreciation is expressed for the hard work on the part of the following individuals who served as reviewers for this issue. Without the assistance of such dedicated professionals, the Psi Chi Journal simply would not be able to function!

—EDITOR

John Broida, University of Southern Maine
Charles I. Brooks, King’s College
Sheila Brownlow, Catawba College
Lisa N. Coates-Shrider, McMurry University
Susan Conradsen, Berry College
David Cox, East Tennessee State University
F. Richard Ferraro, University of North Dakota
Robert W. Flint, Jr., The College of Saint Rose
James N. Forbes, Angelo State University
Gary Gargano, Saint Joseph’s University
John A. Juve, University of Missouri–Columbia
Allen Keniston, University of Wisconsin–Eau Claire
Defne Koraman, Weehawken, NJ
Jason P. Kring, University of Central Florida
John Kulig, Plymouth State College
Psi Chi Research Awards and Grants

Psi Chi annually sponsors national undergraduate and graduate research award competitions, as well as research awards for members submitting the best research for the regional and national paper/poster sessions. Members are encouraged to begin research papers early to submit for presentation at local, state, regional, or national conventions. Chapters are encouraged to provide an opportunity for members to rehearse their papers before an audience prior to presenting them at a convention.

In addition, Psi Chi also sponsors programs to fund student and faculty research. Descriptions of the award/grant competitions follow. Further information and submission forms may be obtained from Psi Chi’s national website (www.psichi.org) or from the Psi Chi National Office, P.O. Box 709, Chattanooga, TN 37401-0709; telephone: (423) 756-2044; e-mail: psichi@psichi.org.

Guilford Awards

All Psi Chi undergraduate members are eligible to submit their research for the Psi Chi/J. P. Guilford Undergraduate Research Awards. Cash awards are $1,000 for first place, $650 for second place, and $350 for third place. In addition, all winners and their faculty research advisors receive award certificates. The abstracts of the winning papers, as well as photographs and brief biographies of the top three winners, are published in Eye on Psi Chi. The deadline for this award is May 1 (postmark).

Allyn & Bacon Awards

The Psi Chi/Allyn & Bacon Psychology Awards, sponsored by Allyn & Bacon Publishers, are open to all undergraduate Psi Chi members and are awarded to those who submit the best overall empirical research papers. The awards are $500 for first place, $300 for second place, and $200 for third place. In addition, all winners and their faculty research advisors receive award certificates. The abstracts of the winning papers, as well as photographs and brief biographies of the top three winners, are published in Eye on Psi Chi. The deadline for this award is April 1 (postmark).

Erlbaum Awards

The new Psi Chi/Erlbaum Awards in Cognitive Science, sponsored by publisher Lawrence Erlbaum Associates, Inc., are open to all Psi Chi undergraduate and graduate Psi Chi members and are awarded to those who submit the best overall empirical studies in the area of cognitive science. The awards are $500 for the first-place graduate student and $500 for the first-place undergraduate student. In addition, the winners and their faculty research advisors receive award certificates. The abstracts of the winning papers, as well as photographs and brief biographies of the top three winners, are published in Eye on Psi Chi.
graphs and brief biographies of the top two winners, are to be published in *Eye on Psi Chi*. The deadline for this award is April 1 (postmark).

**Newman Graduate Award**

All psychology graduate students are eligible to submit their research for the Psi Chi/APA Edwin B. Newman Graduate Research Award. The winner receives: (1) travel expenses to attend the APA/Psi Chi National Convention to receive the award, (2) a three-year subscription to an APA journal of the winner’s choice, and (3) two engraved plaques, one for the winner and one for the winner’s psychology department, as a permanent honor to the winner. In addition, the abstract of the winning paper, as well as a photograph and brief biography of the winner, is published in *Eye on Psi Chi*. This award is the only student research award presented during the prestigious APA/APF Awards ceremony at the APA/Psi Chi National Convention in August. The deadline for this award is February 1 (postmark).

**Regional Research Awards**

All Psi Chi members (undergraduate and graduate) are eligible to submit their research for the Regional Research Awards. Cash awards of $300 each are presented to students submitting the best research papers to Psi Chi sessions at regional conventions. The number of awards in each region vary with the size of the regions; 78 awards of $300 each are available for the 2001–2002 year. Award monies are distributed at the conventions following the presentations. The Psi Chi regional vice-presidents each send a Call for Papers and a letter to the Psi Chi chapters in their respective regions during the fall. These letters include information about the Regional Research Awards, the regional conventions, and submission deadlines for Psi Chi programs. Deadlines for submissions vary according to region and sometimes from year to year; check your fall regional mailing or Psi Chi’s national website ([www.psichi.org](http://www.psichi.org)) for details.

**National Convention Research Awards**

All Psi Chi members (undergraduate and graduate) are eligible to submit their research for the National Convention Research Awards. Cash awards of $300 each are presented to students submitting the best research for Psi Chi sessions at the APA and APS national conventions. Up to eight awards are given: four for the APA Convention and four for the APS Convention. Award monies are distributed at the conventions following the presentations. A Call for Proposals is mailed to all chapters in the fall and is also available from the Psi Chi National Office or the Psi Chi national website ([www.psichi.org](http://www.psichi.org)). The deadline for submissions to the Psi Chi student sessions at both the APA and APS conventions is December 1 (postmark).

**Hunt Research Awards**

All Psi Chi student and faculty members are eligible to apply for a Thelma Hunt Research Award. Up to three awards of $3,000 each will be presented annually to enable members to complete empirical research that addresses a question directly related to Psi Chi, as posed by either (1) the Psi Chi National Council or (2) the researcher submitting a proposal. Unlike other national Psi Chi award/grant programs, the Hunt Awards focus on research directly related to the mission of Psi Chi. The deadline for this award program is October 1 (postmark).

**Undergraduate Research Grants**

All undergraduate Psi Chi members are eligible to apply for these undergraduate research grants. The purpose of this program is to provide funds for members to defray the cost of conducting a research project. Applicants may request up to $1,500 for each project. A total of $45,000 has been allotted for this student grant program. The deadline for this grant program is October 1 (postmark).

**Faculty Advisor Research Grants**

All current faculty advisors and coadvisors who have served an active Psi Chi chapter for at least one year are eligible to apply for these new faculty advisor research grants. The purpose of this program is to provide funds for advisors to defray the direct costs of conducting a research project (no stipends included). Two grants will be awarded annually within each of Psi Chi’s six regions. The maximum amount of each grant will be $2,000. The deadline for this grant program is June 1 (postmark).
Psi Chi Journal of Undergraduate Research

Subscriptions and Back Issues

The Psi Chi Journal of Undergraduate Research is a national, fully reviewed, quarterly journal dedicated to the publication of undergraduate student research. All active Psi Chi chapters receive one complimentary subscription to the journal. We encourage each chapter to see that an additional subscription is obtained for the school library and that other organizations and interested individuals are made aware of its availability. Every effort has been made to provide a high-quality publication and yet offer the journal at affordable subscription rates to ensure its availability to all interested students, faculty members, and institutions. Back issues and bulk orders for classroom use are also available.

Journal Subscription and Back Issue Order Form

Subscriptions and Back Issues

SUBSCRIPTION RATES
Individuals: $20/year (4 issues)
Institutions: $40/year (4 issues)
Subscription orders must be prepaid and are based on the calendar year. Make checks payable to Psi Chi.

BACK ISSUE RATES
Entire Volumes (Individuals): $20
Entire Volumes (Institutions): $40
Single issues (Individuals): $6 per copy
Single issues (Institutions): $12 per copy

CLASSROOM USE
Bulk orders for classroom use are available at a discount. Faculty members should contact the Psi Chi National Office (journal@psichi.org) for availability, rates, and shipping costs.

INTERNATIONAL (OUTSIDE U.S.)
Canada: Add $10 for annual subscriptions; add $2.50 for single issues.
Other International: Contact Psi Chi National Office for rates.

SHIPPING ADDRESS (please type or print):
Name
Address

Check one: ☐ New Subscription ☐ Renewal
Check all that apply: ☐ 2001 ☐ 2002 ☐ 2003 ☐ 2004
________ @ $20 (individual) = __________
________ @ $40 (institution) = __________

BACK ISSUE ORDER
(see rates at left)
Enter number you are ordering beside all that apply:
2000 __Entire volume __Spring __Summer __Fall __Winter
1999 __Entire volume __Spring __Summer __Fall __Winter
1998 __Entire volume __Spring __Summer __Fall __Winter
1997 __Entire volume __Spring __Summer __Fall __Winter
1996 __Entire volume __Spring/Summer __Fall/Winter

Subscription total _________________ = $________
Back Issue total ______________ = $________
International total (see rates at left) __ = $________
Total amount enclosed ________________ = $________

Mail this form to: Psi Chi National Office (Subscriptions), P.O. Box 709, Chattanooga, TN 37401-0709
Other Journals of Student Research in Psychology

**Journal of Psychological Inquiry**
- Founded 1996 by the Great Plains Behavioral Research Association
- Authors must be undergraduate students.
- Publishes empirical studies, literature reviews, and historical articles concerning any topical area in the psychological sciences.
- Submissions must (a) come from students at institutions that sponsor the Great Plains Students’ Psychology Convention and the *Journal of Psychological Inquiry* or (b) have been accepted for or presented at the meeting of the Great Plains Students’ Psychology Convention, the Association for Psychological and Educational Research in Kansas, the Nebraska Psychological Society, or the Arkansas Symposium.
- Contact: Mark E. Ware, Managing Editor
  Department of Psychology
  Creighton University
  2500 California Plaza
  Omaha, NE 68178-0001
  Telephone: (402) 280-3193
  E-mail: meware@creighton.edu
  Web URL: http://puffin.creighton.edu/psy/journal/JPIhome.html

**Journal of Psychology and the Behavioral Sciences**
- Founded 1966
- One issue per year
- Authors may be undergraduate or graduate students with faculty mentor.
- Contact: Dr. Daniel J. Calcagnetti
  Faculty Editor
  Department of Psychology
  Fairleigh Dickinson University
  285 Madison Avenue
  Madison, NJ 07940-1099
  Telephone: (973) 443-8974
  E-mail: robinc@enter.net
  Web URL: http://alpha.fdu.edu/psychweb

**Modern Psychological Studies**
- Founded 1992
- Two issues per year: September and March
- Primary author must be an undergraduate student.
- Preferred submission deadlines: April and October
- Publishes experimental research, but will also consider theoretical papers, literature reviews, and book reviews.
- Contact: Editor, MPS
  Department of Psychology
  University of Tennessee at Chattanooga
  615 McCallie Avenue
  Chattanooga, TN 37403-2598
  Telephone: (423) 785-2238, 755-4262
  E-mail: mpssub@utc.edu

**Journal for Undergraduate Research in Psychology**
- Founded 1997
- Published quarterly (March, June, September, December)
- Online journal dedicated to undergraduate research presented at local, regional, or national conferences.
- Abstracts accepted from any conference.
- Award winners may have full text of paper published online.
- Contact: Dr. Chris Koch
  Director of Undergraduate Studies in Psychology
  Department of Psychology
  George Fox University
  Newberg, OR 97132
  Web URL: http://www.georgefox.edu/jurp/

The journals listed above all solicit and publish research in psychology conducted and written by students. Journals published internally (i.e., which only accept submissions from students within one institution or department) are not listed. If you know of other journals that meet these criteria, please inform the Psi Chi National Office, P.O. Box 709, Chattanooga, TN 37401-0709; Telephone: (423) 756-2044; Fax (toll-free): 1-877-265-1529; E-mail: journal@psichi.org.