ABOUT PSI CHI

Psi Chi is the International 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 students making the study of psychology one of their major interests and who meet Psi Chi’s 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 Association for Psychological Science (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 over 1,100 senior colleges and universities around the world. The Psi Chi Central Office is located in Chattanooga, Tennessee. A Board of Directors, composed of psychology faculty 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. The first of these is the Society’s obligation to provide academic recognition to its inductees by the mere fact of membership. The second is the opportunity of each of the Society’s local chapters to nourish and stimulate the professional growth of all members through fellowship and activities designed to augment and enhance the regular curriculum. In addition, the Organization provides programs to help achieve these goals including conventions, research awards and grants competitions, and publication opportunities.

JOURNAL PURPOSE STATEMENT

The twofold purpose of the Psi Chi Journal of Psychological Research is to foster and reward the scholarly efforts of psychology students as well as to provide them with a valuable learning experience. The articles published in the journal represent the work of undergraduates, graduate students, and faculty. To further support authors and enhance journal visibility, articles are now available in the PsycINFO® and EBSCO Academic Search Complete® databases.

JOURNAL INFORMATION

The Psi Chi Journal of Psychological Research (ISSN 2164-8204) is published quarterly in one volume per year by Psi Chi, Inc., The International 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 Psi Chi Central Office. 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. Psi Chi does not accept paid advertising for Psi Chi Journal of Psychological Research.

Postmaster: Send address changes to the address below.

Copyright 2015 by Psi Chi, The International Honor Society in Psychology.

For more information, contact:
Psi Chi Central Office
Publication and Subscriptions
P.O. Box 709
Chattanooga, TN 37401-0709
(423) 756-2044
www.psichi.org
psichjournal@psichi.org
2  Development of the Test and Examination Anxiety Measure (TEAM)
Byron D. Brooks, Dania Alshafei, and Lloyd A. Taylor*
The Citadel

11  Effects of Gender and Type of Praise on Task Performance Among Undergraduates
Leah Lessard and Andrew Grossman, University of San Diego
Maggie L. Syme*, San Diego State University

18  Anxious Solitary Adolescents’ Coping in Response to Peer Stress
Victor A. Scott and Madelynn D. Shell, The University of Virginia’s College at Wise
Heidi Gazelle*, The University of Melbourne

29  Attachment Moderates the Association Between Mother and Child Depressive Symptoms
Melanie K. Fox and Jessica L. Borelli*
Pomona College

37  Effects of Stimulus Mode and Mode Congruency on Degraded Image Identification for the Picture Superiority Effect
Thomaesa Brundage and Ami L. Barile-Spears*
Mercer University

45  The Interaction of Semantic Relatedness and Emotion: Why Emotion May Not Lead to Better Memory
Christopher Hunt, Janet Trammel*, and Elizabeth J. Krumrei-Mancuso*
Pepperdine University

52  Personality, Social Stress, and Drug Use Among College Students
Jennifer Coleman, Kent State University
Joseph Trunzo*, Bryant University

*Faculty mentor
Test anxiety (TA) is a debilitating response to examinations that often torments students (Salend, 2012) and has been referred to as “the most virulent impediment to effective role functioning in an educational setting” (Myers, 1986, p. 728). Between 25 and 45% of students report the effects of TA (Salend, 2012), and in today’s arena of high-stakes testing, it is not surprising that many students report experiencing TA. In fact, high-stakes testing induces more TA for students than typical classroom testing (Segool, Carlson, Goforth, Embse, & Barterian, 2013). Stankov (2010) suggested that TA levels will rise as countries become more developed. This phenomenon is not going away and past findings disagree on whether TA causes error in the measurement of students’ true ability on tests (Embse & Hasson, 2013; Putwain, 2008; Reeve & Bonaccio, 2008). Although there is disagreement about whether TA affects testing performance, previous empirical studies are in agreement that TA is related to reductions in academic performance as indicated by grade point average (GPA) and final course grades (Bembenutty, 2009; Chapell et al., 2005; Embse & Hasson, 2013).

TA is comprised of emotional, cognitive, and somatic factors, which occur before, during, and after an examination. A review of the literature defines TA as “the set of phenomenological, physiological, and behavioral responses that accompany concern about possible negative consequences or failure on an exam or similar evaluative situation” (Zeidner, 1998, p. 17). Despite TA being comprised of various components, current scales do not assess all aspects of TA because they were developed utilizing a theory stating that TA has only two components. Therefore, these scales lack content validity of this construct. The purpose of the present study was to create a scale that more comprehensively measured TA among students. The Test and Examination Anxiety Measure (TEAM) was developed and administered to undergraduate and graduate students ($N = 362$). The study examined convergent validity with measures of trait anxiety, TA, and academic performance. Results indicated that the TEAM produced optimal reliability ($\alpha = .90$) and validity. The TEAM had significant positive correlations with the State-Trait Anxiety Inventory ($r = .50$, $p < .001$) and the Test Anxiety Inventory ($r = .79$, $p < .001$) as well as a significant negative correlation with grade point average. These outcomes showed that the TEAM has promise. Recommendations to further strengthen the scale are provided.

**ABSTRACT.** Test anxiety (TA) is a multifaceted domain defined as “the set of phenomenological, physiological, and behavioral responses that accompany concern about possible negative consequences or failure on an exam or similar evaluative situation” (Zeidner, 1998, p. 17). Despite TA being comprised of various components, current scales do not assess all aspects of TA because they were developed utilizing a theory stating that TA has only two components. Therefore, these scales lack content validity of this construct. The purpose of the present study was to create a scale that more comprehensively measured TA among students. The Test and Examination Anxiety Measure (TEAM) was developed and administered to undergraduate and graduate students ($N = 362$). The study examined convergent validity with measures of trait anxiety, TA, and academic performance. Results indicated that the TEAM produced optimal reliability ($\alpha = .90$) and validity. The TEAM had significant positive correlations with the State-Trait Anxiety Inventory ($r = .50$, $p < .001$) and the Test Anxiety Inventory ($r = .79$, $p < .001$) as well as a significant negative correlation with grade point average. These outcomes showed that the TEAM has promise. Recommendations to further strengthen the scale are provided.
(Mandler & Sarason, 1952). This description was further expounded upon by Liebert and Morris (1967), who differentiated between the affective and cognitive components presented by Mandler and Sarason, referring to these components as worry and emotionality. According to Mandler and Sarason, worry is defined as the cognitive aspect of TA including negative self-evaluations, interfering thoughts of the consequences of failure, worrisome thoughts, ruminative negative thinking, and doubting one’s ability on the performance of an exam (Enright, Baldo, & Wykes, 2000; Morris & Liebert, 1970). Emotionality is defined as the affective and physiological reactions experienced by the individual in response to the examination (Enright et al., 2000; Morris & Liebert, 1970). This conceptualization led many more test designers to only use that substantive theory as the basis and inspiration for their assessments and even the most widely known measure for assessing TA (Chapell et al., 2005), the Test Anxiety Inventory (TAI) by Charles Spielberger (Taylor & Deane, 2002). Although much empirical work exists to support that worry and emotionality are components of the domain of TA, there is just as much research supporting other components such as distractibility, self-efficacy, metacognition, and trait anxiety (Hodapp & Benson, 1997; Matthews, Hillyard, & Campbell, 1999).

Assessments that measure TA have been built upon the substantive theory provided by these theoretical models, but no one model accounts for the complexity of TA (Zeidner, 1998). If none of the theoretical models explain all of the components of TA, then as a result the assessments built upon these models will not fully capture this complex domain. Criticisms of previous TA measures offered by Zeidner (1998) are that (a) current scales are limited because they do not fully measure the construct of TA, (b) existing scales do not allow for clinicians to differentiate the types of TA, (c) these scales are not relevant for clinical purposes, and (d) current scales do not differentiate between adaptive and maladaptive manifestations of TA (1998).

The purpose of the present study was to address the shortcomings of previous scales and develop an assessment that more comprehensively measured the construct of TA. The scale developed in this study, the Test and Examination Anxiety Measure (TEAM), was intended to resolve these critiques of previous TA scales by measuring more facets of the TA domain, distinguishing the type of TA that an individual has, and allowing clinicians to discern between varying manifestations of TA. In the process of developing the TEAM, this study also provided evidence about TA levels among sex, ethnic group, academic performance, and validation of the short form of the TAI (TAI-5, Taylor & Deane, 2002).

Methods

Participants
The participants were 362 college undergraduate and graduate students from a university in the southeastern United States. Participants were recruited from 15 classes. Undergraduate students made up 77% of the participants and 23% were graduate students. Of the 362 participants, 80% identified themselves as European American, 13% as African American, 3% as Hispanic American, 2% as Asian American, and 1% as other. Men made up 77% of the participants, and the average age was 20.87 years old (SD = 3.67) with a range of 18 to 51 years. Participants were recruited from introductory level and required core courses to obtain more heterogeneity in the sample among college majors.

Measures

The measures used in the present study were the short form of the TAI-5 (Taylor & Deane, 2002), the trait scale of the State-Trait Anxiety Inventory (STAI-T; Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983), self-reported GPA, and the scale developed in the present study, the TEAM.

Past research and assessments have indicated that emotionality and worry are components of TA (Minor & Gold, 1985), but these components needed to be further subdivided and expanded upon to differentiate between types of TA. Worry and emotionality were further subdivided into self-efficacy, metacognition, and social humiliation. Also, trait and state anxiety were included as indicated by past research (Bembenutty, 2009; Bonaccio & Reeve, 2010; Chamorro-Prezuzic, Ahmetoglu, & Funham, 2008; Friedman & Bendas-Jacob, 1997; Hassanzadeh, Ebrahimi, & Mahdinejad, 2012).

In this study, self-efficacy referred to the students’ belief that they could complete a task, in this case an exam. This was included because Hassanzadeh et al. (2012) and Bembenutty (2009) found that students with higher self-efficacy experience less TA. Metacognition was the “beliefs that thoughts are uncontrollable,” which in this case referred to testing situations and was included in the TEAM because Matthews et al. (1999, p.124) proposed that TA traits are related to metacognition.
Development of the TEAM | Brooks, Alshafei, and Taylor

Social humiliation referred to the thoughts related to worry or fear that peers would mock or ridicule a student's performance on an exam. This was included because social derogation is a salient factor of TA (Friedman & Bendas-Jacob, 1997). Trait anxiety was how much anxiety a person generally experiences on a day-to-day basis. State anxiety was based upon how an individual feels in the moment due to their interaction with their current environment. Both were included because research has indicated that TA is comprised of trait and state anxiety (Chamorro-Premuzic et al., 2008; Head & Lindsey, 1983).

The next phase of the present study was writing the prompts for the TEAM and generating an item pool. Using the research as underlying theory, items were generated and written both positively and negatively (e.g., “I have effective test-taking skills” and “I lack effective test-taking skills”). The item pool consisted of 66 items and was reduced to 35 items to make the initial draft of the TEAM. After developing the TEAM, the scale was administered to the participants along with the STAI-T and the short form of the TAI-5.

**TA**

Developed specifically for students, the TAI (Spielberger, 1980) is a long-standing, extensive measure used in assessing TA (Taylor & Deane, 2002). It has demonstrated construct as well as content validity (Spielberger, 1980). The TAI consists of 20 items on a self-report inventory where participants respond to a 4-point Likert-type scale. Participants have four options for responses: (a) *almost never*, (b) *sometimes*, (c) *often*, and (d) *almost always*. It is based upon Liebert and Morris’s (1967) conceptualization of TA as worry and emotionality, and it gives a total score that is determined by summing the subscale scores. The TAI has high internal consistency ($\alpha = .92$), and generally takes 8 to 10 min to complete (Spielberger, 1980). Due to time constraints during the present study, the short form TAI-5 (Taylor & Deane, 2002), a 5-item inventory based on the TAI, was used because it correlates highly ($r = .94$) with the full 20-item inventory and has a high reliability coefficient ($\alpha = .87$). Scores for the TAI-5 were calculated by summing the responses of the five items and multiplying the sum by four in order for the obtained value to be compared to the normative values of the full TAI (Taylor & Deane, 2002).

**Trait Anxiety**

The STAI-T (Spielberger et al., 1983) is used in measuring an individual’s level of current anxiety (state anxiety) and their general level of anxiety (trait anxiety; Szafranski, Barrera, & Norton, 2012). The STAI-T is a 40-item self-report inventory where participants respond on a 4-point Likert-type scale (Julian, 2011). The STAI-T has two subscales: State Anxiety and Trait Anxiety (Bados, Gomez-Benito, & Balaguer, 2010). The Trait Anxiety scale (STAI-T) was used in this study because previous research has supported the notion that trait anxiety is a significant predictor of TA (Chamorro-Premuzic et al., 2008; Head, Engley, & Knight, 1991; Head & Lindsey, 1983). The STAI-T has a high reliability coefficient ($\alpha = .86$; Julian, 2011). Scores for the STAI-T were derived by adding the scores of the 20 items, and the ones that were anxiety-present were scored 1 to 4. Items absent of anxiety were reverse scored 4 to 1 (Spielberger et al., 1983).

**Grades**

Participants’ cumulative GPA was self-reported, and the range at the university participants selected from was 0.0 to 4.0.

**TEAM**

The TEAM (see Appendix) is a 26-item, reduced from 35 items, self-report inventory that measures students’ anxious behaviors related to testing situations. The TEAM has five subscales that discern which manifestations of TA a student possesses. The five subscales of the TEAM are State Anxiety, Distractibility During Exams, Trait Anxiety, Worry, and Rumination. State Anxiety refers to the amount of perceived stress that an individual feels while being in a testing situation. Distractibility During Exams is the individual’s behavioral response before and during an exam. Trait Anxiety in the context of the TEAM is conceptualized as how much anxiety the person experiences on a regular basis because it is a part of their personality. Worry measures how the individual copes with a failed testing situation and their concern of how a testing situation will affect their self-image. Rumination is the individual’s repetitive negative thinking about how much they do not enjoy testing situations. Tables 1 and 2 outline the normative data and factors of the TEAM. Respondents rated their test anxious behaviors on a 5-point Likert-type scale from 1 (*uncharacteristic of me*) to 5 (*characteristic of me*). The TEAM produced a reliability coefficient of .90 and demonstrated concurrent validity with...
other measures of anxiety and TA. Scores were derived by adding the values of the 26 items of the TEAM. Prompts that had anxiety present were scored 1 to 5 and prompts free of anxiety were reverse scored 5 to 1.

**Procedure**

Prior to administering the scales, the institutional review board approved this project. Participants were administered the scales in class on days that the class was not scheduled for any type of assessment (e.g., no tests, exams, or quizzes) to not influence the participants’ anxious behavior. At the beginning of the class, participants were given the informed consent document and the three assessment measures. After obtaining informed consent, participants responded to the prompts on the three assessments. All the data were collected and entered into Statistical Package for the Social Sciences for analysis. Scores from the TEAM, TAI-5, STAI-T, and GPA were evaluated using Cronbach’s alpha, factor analysis, Pearson’s correlation, independent-samples t tests, and an Analysis of Variance (ANOVA).

**Results**

Based on the results produced, prompts on the TEAM were evaluated, and the TEAM was reduced to 26 items. Items were removed based on the item’s correlation with other factors and evaluation of the Cronbach’s alpha value output, which reported what the scale reliability would be if the item was deleted from the scale.

**Factor Analysis and Reliability**

A component factor analysis (direct oblimin rotation) was conducted on the TEAM’s 26 items to determine the components of the TEAM and which prompts loaded on each factor. As summarized in Tables 2 and 3, analyses of the 26 core items in the TEAM ultimately yielded five factors with adequate internal consistency in students who have TA. The Cronbach’s alpha value of the five subscales ranged from .72 to .83. Preliminary analysis of the face validity of the factors suggested that State Anxiety, Distractibility During Exams, Trait Anxiety, Worry, and Rumination factors make up the TEAM.

To determine the reliability of the TEAM, the internal consistency measure of Cronbach’s alpha was calculated on the TEAM scores. The TEAM had a reliability coefficient of .90. Cronbach’s alpha was also calculated on STAI-T and the TAI-5. Those reliability coefficients were .90 and .89 respectively (see Table 3).

**Validity and Correlation**

Construct and concurrent validity of the TEAM was established by correlating the total scores of the TEAM, TAI-5, STAI-T, and GPA with each other using Pearson’s correlations. The TEAM significantly correlated with the TAI-5 at .79 (. < .001). The TEAM significantly correlated with the STAI-T at .50 (. < .001), and the TEAM, STAI-T, and TAI-5 all negatively correlated with GPA. A summary of these results can be found in Table 4. Table 5 outlined the five factors of the TEAM and their relationship to the TAI-5, STAI-T, and GPA.

**Independent t Test and ANOVA**

An independent t test was conducted to determine if TA differed among sex. Women had higher scores on the TEAM, STAI-T, and TAI-5 than men. There was a significant difference between men and women on the TAI-5. No significant difference was found between men and women on the TEAM and STAI-T. A summary of sex differences can be found in Table 6.

An ANOVA was conducted to determine if there were any significant differences in TA among participants’ ethnic groups. The results of the ANOVA revealed no significant differences among ethnic groups in TA on the TEAM and TAI-5 (see Table 7). The results did indicate that participants who described themselves as Asian American had higher levels of trait anxiety based on their responses on the STAI-T.

**Discussion**

The purpose of the present study was to address criticisms of the previous TA scales by developing a scale that more comprehensively measured the construct of TA. Specifically, this study sought to address the criticisms that (a) current scales are

<table>
<thead>
<tr>
<th>Measure</th>
<th>Total Sample</th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEAM</td>
<td>70.61±17.63</td>
<td>69.84±18.02</td>
<td>73.86±16.83</td>
</tr>
<tr>
<td>STAI-T</td>
<td>36.60±8.79</td>
<td>36.23±8.62</td>
<td>37.58±8.98</td>
</tr>
<tr>
<td>TAI-5</td>
<td>40.56±16.24</td>
<td>39.63±16.29</td>
<td>44.22±16.21</td>
</tr>
</tbody>
</table>

Note. TEAM = Test and Examination Anxiety Measure; STAI-T = Trait Scale of the State-Trait Anxiety Inventory; TAI-5 = Short Form of the Test Anxiety Inventory. The total sample size varies due to missing responses on the STAI-T and TAI-5.

Men and Women add to 327 because there were 35 participants with missing data for sex.
limited because they do not fully measure the construct of TA, (b) existing scales do not allow for clinicians to differentiate the types of TA, (c) these scales are not relevant for clinical purposes, and (d) current scales do not differentiate between adaptive and maladaptive manifestations of TA, all offered by Zeinder (1998). To correct for these criticisms by developing a new scale was a daunting task, but the TEAM seems to be on the right path as indicated by the results of this study.

The Cronbach’s alpha value for the TEAM of .90 shows that the scale can be used for clinical applications, the preferred alpha for clinical assessments being .90 or higher (Connelly, 2011; DeVon et al., 2007). The TEAM proved that it measures TA as indicated by its significant correlations with the TAI-5 and the STAI-T. The TEAM’s significant correlations with the TAI-5 and STAI-T provided support for the TEAM having construct and convergent validity due to the TAI-5 and STAI-T being previously validated measures of TA and trait anxiety. Further work could be done in this area to set clinical cutoff scores for the TEAM so that it may be used in formal clinical evaluations.

The TEAM corrected for some previous criticisms of TA scales by including more aspects of the TA domain, therefore measuring more manifestations of TA besides the worry and emotionality, as indicated by the factor analysis results demonstrating five factors. However, the five factors that the TEAM measured did not match up with the components of TA that the prompts were written for based upon the literature review. This indicated that TA is not only a complex domain to measure, but also a very convoluted construct. By including more aspects of TA, the TEAM was able to identify five factors: the type of State Anxiety that a student has when taking an examination, Distractibility During Exams, Trait Anxiety, Worry, and Rumination about examinations. State Anxiety due to examinations, the first factor of the TEAM, was represented by prompts that focused on the perceived stress the students experience when in a testing situation. The prompts of “The thought of an exam makes me anxious,” “When I am faced with an exam, I become anxious,” “Exams generally cause me more anxiety than other items in my life,” “When presented with an exam, I do not sense any physical symptoms of anxiety,” and “Exams do not cause me more anxiety than other things in my life.”

<table>
<thead>
<tr>
<th>Item</th>
<th>Corresponding factor</th>
<th>Factor loading</th>
<th>Communality</th>
</tr>
</thead>
<tbody>
<tr>
<td>I feel anxious the majority of the time.</td>
<td>Trait Anxiety</td>
<td>.72</td>
<td>.59</td>
</tr>
<tr>
<td>I am easily distracted during exams.</td>
<td>Distractibility</td>
<td>.68</td>
<td>.48</td>
</tr>
<tr>
<td>I have a difficult time comprehending the instructions of exams.</td>
<td>Distractibility</td>
<td>.71</td>
<td>.53</td>
</tr>
<tr>
<td>I do not put in effort when it comes to exams because I know I will fail.</td>
<td>Distractibility</td>
<td>.66</td>
<td>.47</td>
</tr>
<tr>
<td>I avoid courses or professors that use a lot of exams.</td>
<td>Distractibility</td>
<td>.59</td>
<td>.37</td>
</tr>
<tr>
<td>I worry about how others will view me if I do poorly on an exam.</td>
<td>Trait Anxiety</td>
<td>.69</td>
<td>.48</td>
</tr>
<tr>
<td>I have performed poorly on an exam, I have a hard time with coping and moving on from that experience.</td>
<td>Trait Anxiety</td>
<td>.69</td>
<td>.62</td>
</tr>
<tr>
<td>I worry about how others will view me if I do poorly on an exam.</td>
<td>Trait Anxiety</td>
<td>.61</td>
<td>.59</td>
</tr>
<tr>
<td>I am hypercritical of myself usually.</td>
<td>Trait Anxiety</td>
<td>.73</td>
<td>.54</td>
</tr>
<tr>
<td>Doing poorly on an exam makes me feel dejected.</td>
<td>Worry</td>
<td>.74</td>
<td>.57</td>
</tr>
<tr>
<td>After an exam, I still continue to worry about how well I did on that exam until I find out for certain.</td>
<td>Worry</td>
<td>.74</td>
<td>.58</td>
</tr>
<tr>
<td>After I have performed poorly on an exam, I have a hard time with coping and moving on from that experience.</td>
<td>Worry</td>
<td>.51</td>
<td>.62</td>
</tr>
<tr>
<td>I worry about how others will view me if I do poorly on an exam.</td>
<td>Worry</td>
<td>.57</td>
<td>.59</td>
</tr>
<tr>
<td>I worry about how an exam will affect my success in the future.</td>
<td>Worry</td>
<td>.67</td>
<td>.52</td>
</tr>
<tr>
<td>I view exams as a negative part of the education system.</td>
<td>Rumination</td>
<td>.79</td>
<td>.66</td>
</tr>
<tr>
<td>Worrying about my performance on an exam affects my performance on an exam.</td>
<td>Rumination</td>
<td>.61</td>
<td>.53</td>
</tr>
<tr>
<td>I wish there were ways to measure my knowledge of material other than exams.</td>
<td>Rumination</td>
<td>.71</td>
<td>.58</td>
</tr>
<tr>
<td>Exams are a way for me to demonstrate my knowledge.</td>
<td>Rumination</td>
<td>.76</td>
<td>.65</td>
</tr>
</tbody>
</table>

Note: TEAM = Test and Examination Anxiety Measure. Factor loadings and communalities based on a principal component analysis with oblimin rotation for the items from the TEAM are also reported (n = 362).
not put in effort when it comes to exams because I know I will fail,” and “I avoid courses or professors that use a lot of exams” displayed being unfocused or avoidance before or during the testing situation.

Trait Anxiety, the third factor of the TEAM, measured on the students’ normal level of experienced anxiety on a daily basis. Prompts like “I feel anxious the majority of the time” and “I am hypercritical of myself” usually represented the students’ typical level of anxiety. The fourth factor, Worry, was based on how the students coped with a failed testing situation and their concerns about how the situation would affect them. “Doing poorly on an exam makes me feel dejected,” “After an exam, I still continue to worry about how well I did on that exam until I find out for certain,” “After I have performed poorly on an exam, I have a hard time with coping and moving on from that experience,” “I worry about how others will view me if I do poorly on an exam,” and “I worry about how an exam will affect my success in the future” demonstrated how the students continue to experience anxiety after the testing situation. The students’ concern about how the exam will affect how their peers view them and how the exam will affect their future is measured by this factor.

The last factor of the TEAM was Rumination, which centered on the students’ negative thinking about examinations and how they do not enjoy testing situations. “Worrying about my performance on an exam affects my performance on an exam” focused solely on the students’ negative thinking about examinations. “I view exams as a negative part of the education system,” “I wish there were other ways to measure my knowledge of material other than exams,” and “Exams are a way for me to demonstrate my knowledge” represented how the students were pessimistic about testing situations. The preliminary results demonstrated that the TEAM measured several facets of TA, but further revision may be needed to enhance the TEAM ability to comprehensively measure the domain of TA.

More work needs to be done with the TEAM prior to being able to compete with the most widely used measure of TA, the TAI by Spielberger. The TEAM appears to measure more facets of TA when compared to the TAI, which measures only emotionality and worry (Spielberger, 1980). The TEAM also measures trait anxiety as well as state anxiety, and the TAI only measures state anxiety in testing situations (Spielberger, 1980). Another criticism of the TAI presented by Szafranski et al. (2012) is that the normative data for the TAI has not been updated since being published in 1980. Their study demonstrated that the norms published in 1980 are significantly lower than the college population now. Since then, academic settings have become far more diverse in all manners, and the TAI is falling behind in staying up-to-date with the changing times.

The present study provided evidence toward other findings in TA research. This study found

<table>
<thead>
<tr>
<th>TABLE 3</th>
<th>Descriptive Statistics for the Five TEAM Factors (N = 362)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor</td>
<td>No. of items (n)</td>
</tr>
<tr>
<td>State Anxiety</td>
<td>7</td>
</tr>
<tr>
<td>Distractibility During Exams</td>
<td>6</td>
</tr>
<tr>
<td>Trait Anxiety</td>
<td>4</td>
</tr>
<tr>
<td>Worry</td>
<td>5</td>
</tr>
<tr>
<td>Rumination</td>
<td>5</td>
</tr>
</tbody>
</table>

Note: TEAM = Test and Examination Anxiety Measure. The sum of the number of items is greater than the number of items due to multiple items loading highly on multiple factors.

<table>
<thead>
<tr>
<th>TABLE 4</th>
<th>Correlations Between the TEAM, the TAI-5, STAI-T, and GPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEAM</td>
<td>TAI-5</td>
</tr>
<tr>
<td>TEAM</td>
<td>.79**</td>
</tr>
<tr>
<td>TAI-5</td>
<td>.46**</td>
</tr>
<tr>
<td>STAI-T</td>
<td>-.14**</td>
</tr>
</tbody>
</table>

Note: TEAM = Test and Examination Anxiety Measure; STAI-T = Trait Scale of the State-Trait Anxiety Inventory; TAI-5 = Short Form of the Test Anxiety Inventory; GPA = Grade Point Average. ** Significant correlation at the 0.01 level (1-tailed).

| TABLE 5 | Correlations Between the Five Factors of the TEAM, TAI-5, STAI-T, and GPA |
|---|---|---|---|---|---|
| | TEAM | TAI-5 | STAI-T | GPA |
| State Anxiety | .45** | .53** | .58** | .71** | .36** | .08 |
| Distractibility During Exams | .37** | .33** | .63** | .60** | .35** | -.43** |
| Trait Anxiety | .53** | .37** | 1 | .77** | .33** | .50** | .57** | .00 |
| Worry | .58** | .33** | .77** | 1 | .37** | .53** | .40** | .06 |
| Rumination | .53** | .63** | .33** | .37** | 1 | .68** | .33** | -.38** |
| TAI-5 | .72** | .80** | .50** | .53** | .68** | 1 | .46** | -.28** |
| STAI-T | .38** | .35** | .57** | .40** | .33** | .46** | 1 | -.14** |
| GPA | .08** | -.43** | -.00 | -.06 | -.38** | -.23** | -.14** | 1 |

Note: TEAM = Test and Examination Anxiety Measure; STAI-T = Trait Scale of the State-Trait Anxiety Inventory; TAI-5 = Short Form of the Test Anxiety Inventory; GPA = Grade Point Average. ** Significant correlation at the 0.01 level (1-tailed).
a significant difference in TA levels among sex indicated by scores on the TAI-5, which is consistent with past literature finding that women exhibit higher levels of TA (Akanbi, 2013; Akca, 2011; Chapell et al., 2005; Hembree, 1988). The TEAM negatively correlated with GPA which is consistent with other studies (Bembenutty, 2009; Chapell et al., 2005). Taylor and Deane’s study reported an alpha of .87 for the TAI-5, and the present study found a Cronbach’s alpha value of .89 for the TAI-5, which is significant because the TAI-5 appeared to be more reliable in this population and the TAI-5 has only one other study which replicated their results.

The TEAM is a reliable and valid measure, but the present study contained limitations. Although the sample used in this study was normative for the population it was drawn from, it does not represent the population of college students in the United States. Men made up 77.3% of the participants in this sample, and that figure is significantly greater than the 45% average reported by the U.S. Census Bureau (2012a). The sample was representative of European American and African American college students in the United States, but the sample was not representative of Hispanic American students in the United States. The U.S. Census Bureau (2012b) reported that 12% of the college population are Hispanic American, but only 3% of the sample identified as Hispanic American in the present study.

The TEAM shows true potential of being a TA measure that clinicians and educational counseling centers may use to identify students with TA. It can be a very practical measure to utilize because it corrects for many of the past issues that have afflicted other TA measures. Future research with the TEAM is needed to replicate the results found in this study. The reliability of the TEAM is optimal for clinical use, but other measures of reliability like parallel forms or test-retest need to be established as well. To make the TEAM scale more practical for clinical use, further work should be done to determine clinical cutoff scores for the TEAM to better assist clinicians in assessing TA. The TEAM needs to be administered to a more diverse population and gain a larger sample size, allowing the standardization and further development of the scale. The TEAM could add more prompts to include more facets of the TA construct as well. Future research may also investigate how students’ scores on the TEAM fluctuate before, during, and after an examination.

In conclusion, TA is a ubiquitous problem that will become more prevalent as Stankov (2010) suggested and the measurement of TA is essential in treating this problem. The TEAM demonstrated potential in being an assessment that will help clinicians and researchers more accurately and comprehensively measure TA. As suggested by Zeidner (1998), TA should not be diagnosed on the basis of a single score from any measure, and that score should not be utilized to provide an explanation for someone’s behavior. With further research, the TEAM could be an instrumental tool of complete TA triage assessment to provide evidence of what manifestation of TA a student has, and allow a clinician to better tailor and select interventions that will apply to that student’s type of TA.

### TABLE 6

<table>
<thead>
<tr>
<th>Scores</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>t</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEAM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>253</td>
<td>69.84</td>
<td>18.02</td>
<td>1.72</td>
<td>325</td>
<td>.09</td>
</tr>
<tr>
<td>Women</td>
<td>74</td>
<td>73.86</td>
<td>16.83</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STAI-T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>235</td>
<td>36.23</td>
<td>8.62</td>
<td>1.11</td>
<td>299</td>
<td>.27</td>
</tr>
<tr>
<td>Women</td>
<td>74</td>
<td>37.58</td>
<td>8.98</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAI-5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>248</td>
<td>39.63</td>
<td>16.29</td>
<td>2.12</td>
<td>319</td>
<td>.04*</td>
</tr>
<tr>
<td>Women</td>
<td>73</td>
<td>44.22</td>
<td>16.21</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. TEAM = Test and Examination Anxiety Measure; STAI-T = Trait Scale of the State-Trait Anxiety Inventory; TAI-5 = Short Form of the Test Anxiety Inventory. * Significance at the .05 level.

### TABLE 7

<table>
<thead>
<tr>
<th>Ethnicity</th>
<th>European American</th>
<th>African American</th>
<th>Hispanic American</th>
<th>Asian American</th>
<th>Other</th>
<th>df</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEAM</td>
<td>70.54</td>
<td>67.40</td>
<td>78.20</td>
<td>69.00</td>
<td>80.75</td>
<td>4,281</td>
<td>1.15</td>
<td>.34</td>
</tr>
<tr>
<td>STAI-T</td>
<td>36.28a</td>
<td>35.30a</td>
<td>37.75a,b</td>
<td>45.67b</td>
<td>41.25a,b</td>
<td>4,307</td>
<td>2.18</td>
<td>.07</td>
</tr>
<tr>
<td>TAI-5</td>
<td>40.51</td>
<td>38.40</td>
<td>48.00</td>
<td>47.33</td>
<td>49.00</td>
<td>4,299</td>
<td>1.22</td>
<td>.30</td>
</tr>
</tbody>
</table>

Note. TEAM = Test and Anxiety Examination Anxiety Measure; STAI-T = Trait Scale of the State-Trait Anxiety Inventory; TAI-5 = Short Form of the Test Anxiety Inventory. a,b = When F is significant, letters are used to indicate which means were comparable versus different according to Duncan post-hoc test. Those with common letters were comparable.
References
Julian, L. (2011). Measures of anxiety: State-Trait Inventory (STAI), Beck Anxiety Inventory (BAI), and Hospital Anxiety and Depression Scale-Anxiety (HADS-A). *Arthritis Care and Research, 63*, 467–472. doi:10.1002/acr.20561

Author Note. Byron D. Brooks, Dania Alshafei, Lloyd A. Taylor, Department of Psychology, The Citadel, SC.
Correspondence concerning this article should be addressed to Byron D. Brooks, Department of Psychology, The Citadel, Charleston, SC 29409. E-mail: bbrooks2@citadel.edu
**APPENDIX**

**Test and Examination Anxiety Measure**

<table>
<thead>
<tr>
<th>Sex (circle one):</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethnicity (circle one):</td>
<td>European American</td>
<td>Native American</td>
</tr>
<tr>
<td></td>
<td>African American</td>
<td>Hispanic</td>
</tr>
<tr>
<td></td>
<td>Asian</td>
<td>Other</td>
</tr>
<tr>
<td>Class Standing (circle one):</td>
<td>First-Year</td>
<td>Sophomore</td>
</tr>
<tr>
<td></td>
<td>Junior</td>
<td>Senior</td>
</tr>
<tr>
<td>GPA:</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Please read each statement below and consider how characteristic it is of you. Rate each statement using the following scale and record your answer in the space provided.

1 = Uncharacteristic of me  
2 = Somewhat uncharacteristic of me  
3 = Neither uncharacteristic nor characteristic of me  
4 = Somewhat characteristic of me  
5 = Characteristic of me

1. The thought of an exam makes me anxious.
2. Doing poorly on an exam makes me feel dejected.
3. After an exam, I still continue to worry about how well I did on that exam until I find out for certain.
4. When someone finishes an exam when I am halfway done with an exam, I become anxious.
5. I have effective test taking skills.
6. I often feel relaxed and laid-back.
7. I view exams as a negative part of the education system.
8. Worrying about my performance on an exam affects my performance on an exam.
9. When presented with an exam, I begin to sense the physical symptoms of anxiety (sweating, increased heart rate, muscle tension, difficulty breathing).
10. During an exam, I become flustered, and my mind goes blank.
11. When I am faced with an exam, I become anxious.
12. Exams generally cause me more anxiety than other items in my life.
13. I am easily distracted during exams.
14. I have a difficult time comprehending the instructions of exams.
15. When I am well-prepared for an exam, I do not feel anxious about it.
16. I feel anxious the majority of the time.
17. I am hypercritical of myself usually.
18. After I have performed poorly on an exam, I have a hard time coping and moving on from that experience.
19. I worry about how others will view me if I do poorly on an exam.
20. I worry about how an exam will affect my success in the future.
21. I wish there were other ways to measure my knowledge of material other than exams.
22. I do not put in effort when it comes to exams because I know I will fail.
23. When I present myself with an exam, I do not sense any physical symptoms of anxiety (sweating, increased heart rate, muscle tension, difficulty breathing).
24. Exams are a way for me to demonstrate my knowledge.
25. I avoid courses or professors that use a lot of exams.
26. Exams do not cause me more anxiety than other things in my life.
Effects of Gender and Type of Praise on Task Performance Among Undergraduates

Leah Lessard and Andrew Grossman, University of San Diego
Maggie L. Syme, San Diego State University

ABSTRACT. Research has suggested that receiving process praise (e.g., “You’re working hard”) increases children’s academic performance and that person praise (e.g., “You’re smart”) can have detrimental effects. However, few studies have examined how these findings relate to young adults. This experimental study examined the effects of type of praise (person vs. process) and gender on young adults’ task performance. Forty-eight undergraduates were introduced to hidden-item puzzles by the experimenter, who completed the first set with them. Participants were then given 1 min to work on a similar task. Upon completion, they were given either person praise or process praise from the experimenter. After receiving the praise, participants completed a final set of 6 puzzles, which served as the dependent measure. The results indicated that participants who received process praise ($M = 8.08, SD = 3.04$) significantly outperformed those who received person praise ($M = 6.54, SD = 1.56$), $F(1, 44) = 5.00, p = .03, \eta^2 = .10$. However, there was no significant effect for gender and no interaction between type of praise and gender. These findings suggested that process praise may be an effective method of improving academic performance in undergraduate students. Implications for classroom practice and the need for further research that considers longitudinal designs and larger sample sizes are discussed.

Despite the fact that the United States is considered to be a world power, when compared globally, American students rank 25th in math and 14th in reading (Jones, 2010). Clearly, even in our achievement-based society, our students are not up to par. As adults and educators, it is our responsibility to provide the best possible education for children and do all that we can to set them up for academic success. However, many commonly used practices may actually deter students from reaching their potential by affecting their motivation and mindset for learning. Teaching strategies intended to enhance learning can cause students to develop maladaptive self-theories about the nature of intelligence, ultimately hindering the learning process and contributing to gender and racial achievement gaps (Aronson & Steele, 2005).

Why do some students quit at the slightest confrontation with difficulty, when others persevere despite all odds against their success? Students who give up easily may have extrinsic motivation, meaning they are driven by external forces or controls (Henderlong & Lepper, 2002). On the other hand, students who work through challenges may possess intrinsic motivation, which can be defined as engagement due to inner gratification. These students are motivated to master educational tasks for the value of learning itself, rather than for a high grade, accolades, or some other extrinsic reward. Intrinsic motivation has been shown to have many positive outcomes including persistence, creativity, and academic achievement (Henderlong & Lepper, 2002).
Epistemic beliefs of intelligence have similarly been shown to influence academic achievement (Mueller & Dweck, 1998). Specifically, if an individual has an entity view of intelligence (i.e., fixed mindset), they believe that intelligence is immutable and is more likely to think effort plays little to no role in outcome. Students with fixed mindsets are more likely to develop a fear of failure (Kamins & Dweck, 1999; Mueller & Dweck, 1998), resulting in the avoidance of intellectual tasks (Elliott & Dweck, 1988) and giving up in the face of difficulty (Dweck, 1999). The rationale is that effort is unnecessary if individuals are smart because everything should come easy, and if individuals are not smart, hard work cannot compensate for this deficiency.

On the other hand, students who believe that intelligence is adjustable have an incremental view of intelligence (i.e., growth mindset). Students with this growth mindset believe that intelligence is the result of hard work and the use of appropriate strategies. Rather than giving up in the face of failure or challenge, those with a growth mindset interpret setbacks as inevitable for learning to take place. Because they are not worried that exertion of effort is a reflection of lack of intelligence, they are not afraid to work hard, resulting in a superior performance compared to their fixed mindset peers (Dweck, 2007). To illustrate, Blackwell, Trzesniewski, and Dweck (2007) found that, even after 2 years, students with an incremental view of intelligence academically outperformed students with an entity view of intelligence.

It is clear that intrinsic motivation and the development of an incremental view of intelligence are critical to achievement, but where do these conceptions come from? One factor that is controllable and has been shown to be influential is the type of praise that students receive (Hattie, 2003; Hattie & Timperley, 2007). Verbal praise is often administered as a way to reinforce the performance or behavior of students. Although there may be positive intentions, some types of praise can have debilitating effects for the recipient, specifically depending on what the praise is directed toward. For example, person praise focuses on the individual, similar to an affirmation of self-worth such as, “Wow, you’re so smart.” Because it applauds the individual by applying a label or an unchangeable characteristic, person praise promotes an entity view of intelligence (Pomerantz & Kempner, 2013). Students are being rewarded, through praise, for their performance based on their ability. On the other hand, process praise focuses on the actions taken by the individual, especially their effort and problem-solving strategies such as, “Great job! You’re working really hard.” Process praise reinforces the association between success and effort (or behavior) rather than a fixed ability, which cultivates a more adaptive, incremental view of intelligence.

Significant research has linked person praise with a fixed mindset and process praise with a growth mindset. Mueller and Dweck (1998) randomly assigned fifth graders to receive either person or process praise while they solved a set of progressive matrices. The researchers found that those who were given person praise had worse task performance, made more low-ability attributions, reported less task enjoyment, and exhibited less task persistence than children who were given process praise. Additionally, the children given person praise were more focused on performance goals and described intelligence as a fixed trait. This is in contrast to the children given process praise, who were more focused on learning goals and described intelligence as a fluid concept.

Type of praise not only affects immediate behaviors, beliefs, and emotions, but it has also been shown to have long-term consequences. Specifically, it affects how individuals deal with future difficulties and their willingness to apply effort to challenges that may come their way (Gimpian, Arce, Markman, & Dweck, 2007; Corpus & Lepper, 2007; Dweck, 1999; Henderlong & Lepper, 2002; Pomerantz & Kempner, 2013). For example, Kamins and Dweck (1999) showed that person praise is more likely to promote helpless responses to subsequent failures than process praise.

It is important to note that type of praise is not a clear-cut predictor of performance. The individual giving the praise, the recipient’s beliefs, and other contextual factors also need to be taken into account and can play a significant role in the strength of the effect (Maclellan, 2005). Brummelman and colleagues (2014) found that adults are more inclined to give person praise to children whom they perceive to have low self-esteem and give process praise to children whom they perceive to have high self-esteem. The gender of the individual receiving the praise might also be an important characteristic. Research exploring the effect of gender, however, is somewhat conflicting. Koestner, Zuckerman, and Koestner (1989) found that the performance of fifth- and sixth-grade boys was enhanced on a hidden-figure
task when they were given person praise, whereas the performance of their female peers was enhanced by process praise. In their meta-analysis, Henderlong and Lepper (2002) reported the results of Henderlong’s unpublished doctoral dissertation work with upper-elementary students, indicating that process praise enhanced intrinsic motivation in all children, whereas person praise decreased intrinsic motivation in girls and did not significantly affect boys. Thus, the way in which gender interacts with praise to affect motivation and performance remains in question.

The majority of the research conducted on the effects of praise has focused on school-aged children, and little is known about whether the influences remain consistent in young adulthood. The few studies that have involved the undergraduate population have suggested that a similar effect exists. However, the findings remain inconsistent. Hancock (2000) examined the impact of praise on the amount of time college students spent on homework. The results showed that those who were given verbal praise by their professors spent significantly more time on their homework than those that did not receive any praise. The students who received praise also outperformed their peers on the final exam. Nevertheless, the study did not differentiate between types of praise. Haimovitz and Corpus (2011) found that, among undergraduate students, process praise increased intrinsic motivation and perceived competence, whereas person praise decreased overall academic motivation. Koestner, Zuckerman, and Koestner (1987) asked undergraduate students to complete a set of hidden-item puzzles after receiving person praise, process praise, or no praise by the experimenters. The results showed that participants who received process praise spent significantly more time playing with hidden-item puzzles during a solitary free-choice period than those who received person praise or no praise at all. These studies suggested that praise, specifically process praise, may be an effective strategy for enhancing performance in undergraduates.

The present study expanded upon past findings to explore how gender and type of praise (person vs. process) affected young adults’ performance on a hidden-item puzzle task. To isolate the effect of praise on task performance and avoid the potential confounds of prior exposure and ability, the hidden-item puzzle task was chosen as the dependent measure. As a characteristic effort task, trying harder should improve performance on hidden-item puzzles (Baumeister, Hutton, & Cairns, 1990).

By investigating the role of gender, we hoped to shed more light on the gender gap in achievement and learn about contextual factors involved in praise. Consistent with the literature regarding the effect of type of praise on children, we hypothesized significantly higher task performance following process praise as compared to person praise. We expected to observe no gender difference in task performance, a finding that has been substantiated by past research (Butler & Nisan, 1986; Pintrich & de Groot, 1990). Finally, we also predicted an interaction between type of praise and gender, such that process praise would be significantly more effective for women than men. The results have important implications for the classroom environment and suggest pedagogical techniques to improve students’ learning.

Methods

Participants
This study involved 48 undergraduate students from a small private university in southern California. There were 24 (50%) women and 24 (50%) men enrolled in the study. Their ages ranged from 18 to 25 ($M = 20.96$, $SD = 1.44$). Most students were White ($n = 34$, 70.8%) and upper-middle class ($n = 19$, 39.6%), but a range of race/ethnicities and socioeconomic backgrounds were represented (34 White, 4 Asian, 4 Hispanic or Latino, 3 Black, 2 multi racial, 1 Middle Eastern). The sample of students included a range of student achievement levels and academic majors. Participants were recruited either from the introductory psychology participant pool or through informal advertisement of the study in on-campus dormitories. Six (12.5%) were from the introductory psychology pool, and received course credit for their participation. Otherwise, no incentives were offered to participants.

Measures

Demographic questionnaire. Participants were asked to self-report their birth date, the current date, their gender, race/ethnicity, annual household income, grade, major, and GPA on a questionnaire developed by the researchers.

Nina puzzles. Several related studies have used Nina puzzles as a measure of intrinsic motivation and task performance of undergraduates (e.g., Harackiewicz, 1979; Koestner et al., 1987). The puzzles are a compilation of 10 cartoon style
Gender, Praise, and Task Performance | Lessard, Grossman, and Syme

hidden-item drawings by artist Al Hirschfeld. In each of the drawings, the word Nina is hidden anywhere from 2 to 40 times. To measure task performance, for each of the pictures, the participants are asked to circle the word Nina as many times as they can find it. A high number of Nina’s circled is indicative of higher performance.

Design
The present study employed an experimental, between-subjects design to examine the effect of different types of praise and participant gender on task performance. To eliminate the potential confound that type of praise may have different effects when elicited by the opposite gender, we gender-matched the experimenter and participant, such that a female experimenter tested a female participant, and vice versa. This design was also conducted by Baumeister et al. (1990). Participants were randomly assigned within gender to receive either person praise or process praise by a gender-matched experimenter. Person praise was operationalized as, “Great! You are really good at these!” Process praise was operationalized as, “Great! You are really working hard!” The outcome variable of task performance was operationalized as the score on the Nina puzzles. Gender was self-reported by participants on the demographics questionnaire.

Procedure
All procedures in this study were approved by the Institutional Review Board of the University of San Diego. To ensure gender congruency, a female experimenter met individually with female participants, and a male experimenter met individually with male participants. The experimenters were peers of the participants and were aware of the study’s hypotheses.

First, participants were asked to give informed consent and then filled out a demographics questionnaire. The experimenter then introduced the Nina puzzle task and assisted the participant in solving two practice Nina puzzles to ensure understanding. The participant was then given two additional puzzles and asked to find and circle as many Nina words as possible over the course of 1 min. After the minute passed, in the person praise condition, the experimenter stated, “Great! You are really good at these!” In the process praise condition, the experimenter stated, “Great! You are working really hard!” For both conditions, the feedback was in reference to the participants’ performance solving the Nina puzzles thus far. Finally, for the dependent measure, the experimenter gave the participant a set of six Nina puzzles to solve, instructing him or her to find and circle as many Nina words as possible in 3 min. Once this was complete, the experimenter debriefed the participant as to the actual purpose of the study. The total number of Nina’s circled by the participant across the six puzzles was subsequently counted by the experimenter, representing the outcome variable of interest.

Results
A 2 x 2 Analysis of Variance was conducted to test the effects of two different types of praise and gender on task performance in undergraduate students. It was hypothesized that students given process praise would perform significantly better than students given person praise. We predicted no significant gender differences on task performance. Additionally, it was hypothesized that task performance of those in the different praise conditions would differ by gender.

The results confirmed the first hypothesis, yielding a main effect for type of praise, such that the average number of Nina’s found across the final set of puzzles was significantly higher for participants who received process praise (M = 8.08, SD = 3.04) than for those who received person praise (M = 6.54, SD = 1.56), F(1, 44) = 5.00, p = .03, η² = .10. Effect size of these differences indicated that 10% of the difference in the number of Nina’s the participants found was due to the type of praise they received. The second hypothesis related to gender effects was similarly confirmed because no significant difference in task performance between genders was found, F(1, 44) = 2.66, p = .11, η² = .06.

In addition to the main effects of type of praise and gender, this study also examined the interaction between type of praise and gender. It was hypothesized that both genders would perform similarly when receiving person praise, but that women would perform significantly better than men when receiving process praise. The means and standard deviations for task performance as a function of type of praise and gender are presented in Table 1. The interaction effect was found to be not significant, F(1, 44) = 0.30, p = .59, η² = .01, indicating that type of praise had a similar effect on men and women.

Discussion
In the present study, we explored the effect of person and process praise, as well as gender, on
task performance in young adults. We hypothesized that individuals who received process praise would perform significantly better on the Nina puzzle task than those who had received person praise. Our research findings were consistent with this hypothesis. The results also confirmed our expectation of no differences between men and women on task performance. We also predicted that gender differences would contribute to the effect of praise. However, the data collected failed to support this hypothesis.

In the broadest sense, our results were consistent with Hancock (2000), suggesting that praise does influence the performance of young adults. However, our findings extended this work, suggesting that there is more to this concept of positive reinforcement with the effect of praise being significantly dependent upon the type of praise given.

Consistent with the many studies conducted with children (Corpus & Lepper, 2007; Kamins & Dweck, 1999; Mueller & Dweck, 1998), the superior effects of process praise, compared to person praise, appeared to continue along the developmental trajectory into young adulthood. Similarly, Haimovitz and Corpus (2011) found that, when studying undergraduates, process praise increased intrinsic motivation and perceived competence, whereas person praise decreased students’ motivation. These results, along with our own findings, supported the basic tenets of Dweck and colleagues’ theory of attribution, motivation, and achievement (Blackwell et al., 2007; Dweck, 2007; Henderlong & Lepper, 2002; Kamins & Dweck, 1999; Mueller & Dweck, 1998; Pomerantz & Kemper, 2013).

Undergraduate students in this study receiving process praise might have attributed their successes to the controllable factor that they were praised for, effort. It is likely that they acted in accordance with this belief by responding to challenge by increasing effort, resulting in superior task performance. In contrast, students receiving person praise might have attributed their successes or failures to a fixed variable or unchangeable trait. This might have resulted in less motivation to work hard because the person praised participants believed the outcome was out of their hands (Weiner, 2010).

The lack of interaction between type of praise and gender in our results conflicted with past findings that task performance and intrinsic motivation is significantly more enhanced for women after receiving process praise than for men (Koestner et al., 1987; 1989). In their work with undergraduates, Haimovitz and Corpus (2011) reported a marginally significant interaction of gender and type of praise in terms of intrinsic motivation, suggesting that process praise was especially beneficial for women. Our failure to replicate similar results may be explained by our small sample size or other dependent variables such as time spent on the task, which might be more sensitive to gender differences. We suggest that future research conduct a power analysis to explore whether a larger sample size might reveal a clearer pattern and clarify this discrepancy.

In the real world, individuals often receive praise from the opposite gender. Thus, more ecologically valid classroom research is needed to explore whether gender congruency between the individual offering the praise and the recipient of the praise influences the effect of type of praise. Further research should examine how contextual features such as the gender, age, and familiarity of the individual offering praise impact the effect of type of praise. If contextual features are found to mediate the effect of praise type, then this would have important implications for feedback procedures in the classroom.

The results of our study have been limited by the fact that the participants may not be representative of the larger population of undergraduate students. Although this study utilized an experimental design with random assignment, the participants were recruited from a very specific population (e.g., most were White, upper middle class students), representing little demographic diversity. Therefore, our findings can only be generalized to a similarly limited population. Future research should explore the possibility of a larger and more diverse sample of participants, which may display stronger trends with conclusions that can be more confidently applied across settings. Additionally, it is important to address the dependent measure used. Although hidden-item puzzle tasks, particularly Nina’s, have been used in the past as
an indicator of task performance (Baumeister et al., 1990), a more comprehensive measure, clearly indicative of academic performance is recommended as a way to capture the variable of interest.

Because the experimenters were not naïve to the study’s hypotheses, it is possible that the process praise was delivered more enthusiastically or believably than the person praise. However, to control for this as much as possible, and to avoid any potentially influential spontaneous conversation, the experimenters followed a script. Future research should involve blind experimenters to reduce bias.

Perhaps the most significant limitation was the single session intervention. Although our procedure did reveal that just a few minor words have the power to influence participants’ performance, we measured the effect immediately after the feedback was provided. Would students who were praised for their effort outperform their counterparts who were praised for their skill over a long-term period? How does this effect hold up across domains (e.g., classroom, homework)? Knowing whether performance is affected by type of praise, and also how strongly this effect pervades educational outcomes, would be valuable for research on achievement motivation and contribute to educators’ ability to provide the most optimal learning environment for students.

Despite these significant limitations, this study provided important insight into a population that has been, for the most part, omitted from the large body of research on the effects of praise. Because our results were generally consistent with the findings on children, it is likely that the praise young adults receive everyday is shaping their performance in a variety of areas. Specifically, process praise contributes to a mastery learning orientation (Kamins & Dweck, 1999; Pomerantz & Kempner, 2013), whereas person praise not only harms long-term academic performance, but is also associated with shame, lower self-esteem, and lower intrinsic motivation (Brummelman et al., 2014; Heyman, 2008). If these malleable factors play such a significant role in achievement, then educators do students a disservice by offering them feedback based on fixed characteristics, or those that could be perceived as fixed. Rather, our findings suggested that an effective way to maximize performance would be to extend process praise to students immediately before they begin a task. Ultimately, educators should be mindful of the underlying messages they are communicating when offering praise to their students.

References


References
Lessard, Grossman, and Syme | Gender, Praise, and Task Performance


Author Note. Leah Lessard and Andrew Grossman, Department of Psychological Sciences, University of San Diego, CA; Maggie L. Syme, Department of Psychology, San Diego State University, CA.

Correspondence concerning this manuscript should be addressed to Leah Lessard, Department of Psychological Sciences, University of San Diego, San Diego, CA 92110. E-mail: leahmlessard@sandiego.edu
Anxious Solitary Adolescents’ Coping in Response to Peer Stress
Victor A. Scott and Madelynn D. Shell, The University of Virginia’s College at Wise
Heidi Gazelle*, The University of Melbourne

ABSTRACT. This investigation explored coping strategies in response to peer stress among anxious solitary (AS) adolescents who were shy but desired peer interaction. A total of 195 students (56% girls) completed surveys during the fall of 6th grade. Adolescents self-reported anxious solitude and identified how they coped in response to peer problems. Coping responses were coded along 2 dimensions (voluntary vs. involuntary and engaged vs. disengaged), which created 4 categories of coping. Results demonstrated that AS adolescents reported experiencing more peer-related stress than non-AS adolescents ($p = .02$). However, there were also group differences in responses to these stressors. AS as compared to non-AS adolescents were more likely to use both engaged ($p = .01$) and disengaged ($p = .001$) involuntary coping. Although voluntary engagement was the most common strategy for both groups, AS adolescents used voluntary engagement less than non-AS adolescents ($p < .001$). Results suggested that AS adolescents often have uncontrollable responses to their heightened peer stress, and these responses may contribute to further peer mistreatment and anxious solitude.
Workman, & Allan, 2010). As a result of this affective-behavioral vulnerability, AS adolescents may respond less actively to social challenges or setbacks (Gazelle & Druhen, 2009; Stewart & Rubin, 1995). In particular, their social evaluative concerns may lead to heightened sensitivity in the face of ambiguous peer feedback and increased fear of negative peer responses, causing AS adolescents to select less effective coping strategies compared to non-AS adolescents. Therefore, the individual vulnerability of anxious solitude may lead to more avoidant or less direct responses to the same peer stressor.

**Coping in Response to Social Stress**

Coping is a dynamic cognitive, affective, behavioral, or physiological response to a situation that is perceived as taxing to an individual’s resources (Lazarus & Folkman, 1984). In adolescence, peer stressors such as bullying, conflict, peer pressure, or social exclusion could elicit a coping response. Connor-Smith et al. (2000) have identified several types of coping that can be classified into two dimensions: voluntary-involuntary and engagement-disengagement (see Figure 1). The voluntary-involuntary dimension concerns the level of conscious control over coping. Voluntary coping is a controlled effort to regulate the source of stress or reaction to stress, whereas involuntary coping is an automatic response that is not under the individual’s volitional control. Instead, an involuntary response may be a result of the individual’s temperamental predisposition or a conditioned response to a stimulus (Connor-Smith et al., 2000). Although in some literature coping refers only to controlled (voluntary) responses to stressors (Compas et al., 2001), for the purposes of this investigation, coping will refer to both voluntary and involuntary responses to such conflicts. The engagement-disengagement dimension focuses on the level of engagement with, or responses to, the stressor. Engaged responses are approach responses directed toward a stressor. Disengaged responses are oriented away from the stressor and are typically avoidant responses such as ignoring the problem. These two dimensions create four types of coping: voluntary engaged, involuntary engaged, voluntary disengaged, and involuntary disengaged.

Coping strategies change over time based on developmental period and stress. In early childhood, children tend to engage in more involuntary coping, and they develop more cognitive control over their responses to stress throughout childhood (Compas et al., 2001). Across adolescence, active engagement becomes more common, and although disengagement increases in early adolescence (possibly in response to elevated stress), it tends to decrease again in midadolescence (Seiffge-Krenke, Aunola, & Nurmi, 2009). Furthermore, coping in childhood is often dependent on the child’s mood or the specific stressor. In contrast, by late adolescence and early adulthood, individuals have developed trait-like patterns of coping that are consistent across time and situations (Grant et al., 2003; Sontag, Graber, Brooks-Gunn, & Warren, 2008). Nonetheless, these early coping strategies are important because they predict coping and stress over time. Adolescents who used engaged coping in one context in early adolescence reported less perceived stress overall in later adolescence (Seiffge-Krenke et al., 2009). In contrast, adolescents who used disengaged coping in one context later reported more stress within the same context (although not overall, Seiffge-Krenke et al., 2009). Early adolescence may be a particularly important time point in the development of coping because adolescents may be in the process of developing more permanent and well-established patterns of responses to stress.

The coping strategies that adolescents select in response to stress may also have social and emotional consequences in the moment and over time. Problem-focused voluntary engagement appears to

**FIGURE 1**

![Voluntary-Involuntary and Engagement-Disengagement Dimensions Form Four Types of Coping](image_url)
Anxious Solitary Coping

Anxious Solitude and Coping in Response to Peer Stressors

AS adolescents’ social evaluative concerns may lead to heightened arousal during peer conflict (Tan et al., 2012), and AS adolescents may use less direct and more avoidant coping strategies as a result (Findlay, Coplan, & Bowker, 2009). Furthermore, many previous observational and peer-reported studies have identified that AS adolescents experience poorer peer treatment (Gazelle & Ladd, 2003; Gazelle et al., 2005). Therefore, it is likely that AS adolescents self-report higher rates of peer stress (Markovic, Rose-Krasnor, & Coplan, 2013; Seiffge-Krenke et al., 2009), which has been associated with the endorsement of less effective coping strategies (Sonntag et al., 2008). Elevated rates of peer mistreatment and heightened arousal during these stressors may lead AS adolescents to engage in less effective coping strategies, which could result in more negative social and emotional outcomes in the long term.

Voluntary engagement. Voluntary engagement involves a conscious effort to assess the situation and address the problem. AS adolescents’ social evaluative concerns and history of peer mistreatment (Gazelle & Ladd, 2003) may lead them to be less likely to engage in proactive solutions to peer challenges because they fear negative peer responses to a direct attempt to resolve the issue.

Consistent with this idea, observational studies have indicated that socially withdrawn adolescents displayed fewer socially assertive problem-solving strategies (Stewart & Rubin, 1995) and were less likely to initiate conversations with peers (Spence, Donovan, & Brechman-Toussaint, 1999). Furthermore, in response to perceived behavioral rejection by a friend, AS children who were excluded by peers were less likely than normative children to demonstrate voluntary engagement by selecting a new friend (Gazelle & Druhen, 2009). Finally, in response to hypothetical peer conflict, shy adolescents were less likely than others to use approach forms of coping such as problem solving or eliciting social support (Markovic et al., 2013). These findings have suggested that AS adolescents’ social fears may prevent them from using more direct (and possibly more effective) voluntary engagement.

Involuntary engagement. In contrast to voluntary engagement, involuntary engagement occurs when an individual is affected by a stressor (physiologically, cognitively, or emotionally), but these effects occur automatically and are beyond the individual’s conscious control. Elevated rates of peer mistreatment, in addition to social evaluative concerns, may lead AS adolescents to experience more automatic responses to peer conflict (Sonntag et al., 2008). For example, AS adolescents have greater and more sustained physiological responses to negative peer experiences compared to their peers (Erath, Tu, & El-Sheikh, 2012; Gazelle & Druhen, 2009; Tan et al., 2012). Furthermore, AS versus non-AS adolescents report more uncontrollable thoughts in the form of worrying, self-pity, self-blame, and emotional upset in response to a peer challenge (Gazelle & Druhen, 2009; Kingsbury, Coplan, & Rose-Krasnor, 2013). In addition, rumination, or excessive uncontrollable dwelling on a problem, is a form of involuntary engagement that has been linked to heightened anxiety (Burwell & Shirk, 2007; Nolen-Hoeksema, Wisco, & Lyubomirsky, 2008). Previous work has suggested that adolescents higher in internalizing symptoms such as anxiety engage in more co-rumination (excessively discussing and revisiting problems with a friend; Rose, Carlson, & Waller, 2007). Likewise, when faced with a peer challenge, AS adolescents, who on average experience heightened internalizing symptoms (Gazelle & Rudolph, 2004; Gazelle et al., 2010), may be particularly likely to ruminate on the problem individually as well (Jose & Weir, 2013). Although anxiety may increase rumination, rumination also exacerbates internalizing
symptoms (Rose et al., 2007; Tan et al., 2012), contributing to continued anxious solitude. Thus, in response to peer challenges, AS adolescents appear to experience more uncontrollable cognitive, emotional, and physiological responses compared to their peers.

**Voluntary disengagement.** Voluntary disengagement involves making a conscious decision to avoid a problem. AS adolescents are characterized as having normative approach motivation, meaning that they have moderate levels of social interest and desires to engage in social interaction, but also high social avoidance motivation, indicating that social interactions can be anxiety-provoking and stressful, and therefore may be avoided (Asendorpf, 1990; Coplan & Armer, 2007). As a result, AS adolescents may be particularly likely to choose to avoid or disengage from a stressful peer situation. Consistent with this hypothesis, shy compared to nonshy adolescents are more likely to endorse avoidant coping in response to hypothetical social challenges (Burgess, Wojiławowiec, Rubin, Rose-Krasnor, & Booth-LaForce, 2006; Kingsbury et al., 2013), although this work has not distinguished between voluntary and involuntary avoidance. However, some evidence has suggested that voluntary disengagement may be more effective in relieving anxiety in AS adolescents because distraction in response to a peer challenge predicted lower emotional upset for clinically anxious adolescents, but not controls (Tan et al., 2012). Therefore, although it is clear that AS adolescents use more disengagement, whether this disengagement is voluntary or involuntary is somewhat unclear. It is plausible that both AS and non-AS adolescents are equally likely to use voluntary disengagement in coping with peer-related stress, although voluntary disengagement could have more positive emotional consequences for AS adolescents.

**Involuntary disengagement.** Involuntary disengagement involves automatic and uncontrollable responses that are oriented away from the problem such as emotional numbing. The heightened avoidance reported by AS adolescents (e.g., Kingsbury et al., 2013) may not be under their conscious control, and AS adolescents’ social evaluative concerns could lead to heightened avoidance as an automatic response to fear associated with peer difficulties (Connor-Smith et al., 2000). Therefore, they could be more likely than their peers to experience cognitive interference or emotional numbing when faced with peer challenges. The distinction between controlled versus automatic avoidance (i.e., voluntary or involuntary) could predict how effective these strategies are in making the child feel better following a peer conflict.

**Sex.** Previous evidence has suggested main effect sex differences in coping strategies such that, in both children and adults, girls and women report more voluntary engagement compared to boys and men (Burwell & Shirk, 2007; Connor-Smith & Compas, 2002). Although these effects may be moderated by individual vulnerability such as anxious solitude, results of investigations exploring the relationship between anxious solitude, sex, and coping have been mixed (Kingsbury et al., 2013; Sandstrom, 2004; Sontag et al., 2008). The present study tested for main effects of sex as well as interactions between sex and anxious solitude.

**The Present Study**

The present study explored how AS adolescents cope in response to peer stressors in the fall of sixth grade. Early adolescence was expected to be a particularly important time in the development of coping because of the stress associated with the middle school transition. The study compared the coping strategies of AS versus non-AS students using self-reported measures of anxious solitude and coping in response to problems with peers. We hypothesized that AS adolescents would report more peer stressors than non-AS adolescents. Furthermore in response to these stressors, we expected that AS adolescents would report less voluntary engaged coping and more involuntary and disengaged coping strategies than non-AS adolescents. Sex main effects and interactions between sex and anxious solitude were also tested.

**Method**

**Participants**

A sample of 195 students in 36 middle schools in the southeastern United States was used in this study. Participants were on average 11.69 years old (range: 10.99–14.15 years, SD = 0.53 years). Most participants were girls (56%, n = 110) and European American (55.4%, n = 108; 24.6% Latino(a) American, n = 48; 18.5% African American, n = 36; 1.5% Asian American, n = 3). Of the current sample, 63 participants (32.3%) received free or reduced school lunch, indicating low socioeconomic status (SES), and 132 (67.7%) did not qualify for free or reduced lunch, indicating average/high SES.
These participants were selected from a pool of 230 adolescents participating in a larger longitudinal study and were selected based on completeness of data in the fall of sixth grade, just after the transition from elementary to middle school. Longitudinal participants were selected from a screening sample of 688 students in 46 elementary school classrooms in the fall of third grade, and were followed through the spring of seventh grade. The longitudinal participants were selected because they scored high on peer nominations of anxious solitude at the initial time point or were a demographically matched control.

Prior to conducting the study, an institutional review board (protocol number 05279) approved the data collection and methodology. An additional institutional review board at another institution (protocol number 2013-0401) approved analysis of the deidentified data for the purposes of this project.

**Measures**

Participants completed self-report surveys in small groups with one research assistant. The research assistant read questions aloud as students recorded their answers on paper.

**Anxious solitude.** The Motivation for Solitude and Sociability Scale (MSS, developed for this study) is a self-report scale used to measure motives for solitary play. Anxious solitude was computed based on a composite score of seven items (Cronbach’s α = .80). Students responded to the items on a 4-point Likert-type scale, ranging from 0 (never) to 3 (always). The adolescents rated statements about “What I’m like when I’m with other kids” such as “I’m more shy and quiet than other kids, and talk less than they do” and “I’m afraid I will embarrass myself around other kids.” The mean response was 0.46 (SD = 0.44), indicating that adolescents reported low levels of anxious solitude on average. Participants who scored at the 75th percentile (.86) or above on the Anxious Solitude subscale were classified as AS (n = 46), and all others were classified as non-AS (n = 149).

**Coping.** Self-reports on the Response to Stress Questionnaire for Adolescents (RSQ-A) were used to measure coping in response to peer stress (Connor-Smith et al., 2000). In order to measure peer stress, participants identified which of nine peer stressors they had experienced since the start of the school year (e.g., being around kids who are rude, having trouble with friends, peer pressure, being teased, conflict, and romantic or other peer rejection). After selecting which stressors they had experienced, participants rated perceived distress as a result of these stressors in response to the question “How stressful, or how much of a hassle, have these problems been for you” (1 = not at all to 4 = a lot). Next, participants rated how often they acted or felt a certain way in response to these problems (1 = not at all to 4 = a lot). Composite scores were calculated for each type of coping. Reliability and validity of the RSQ-A have been established in previous publications (e.g., Connor-Smith et al., 2000), and Cronbach’s alphas for each type of coping in this sample are listed below.

Voluntary Engagement (9 questions, α = .82) involved acknowledging a problem and making a conscious decision on how to deal with the issue (e.g., “I try to think of different ways to change the problem or fix the situation”). Involuntary Engagement (15 questions, α = .89) involved an automatic or uncontrollable response with a focus on the problem (e.g., “I can’t stop thinking about what I did or said” and “I can’t stop thinking about them when I sleep, or I have bad dreams about them”). Voluntary Disengagement (6 questions, α = .66) involved a conscious effort to engage in an avoidant response (e.g., “I try not to think about it, to forget all about it”). Finally, Involuntary Disengagement (12 questions, α = .85) involved an automatic response that did not directly address the problem (e.g., “I just have to get away when I have problems with other kids. I can’t stop myself” and “My mind just goes blank. I can’t think at all”).

Because AS adolescents experience more peer mistreatment than non-AS peers on average (Gazelle & Ladd, 2003; Gazelle et al., 2005), it was possible that they would endorse all coping items more than non-AS adolescents. To control for differing endorsement rates, RSQ-A responses for each coping type were standardized within participant as recommended by Connor-Smith et al. (2000, see also Connor-Smith & Compas, 2002; Sontag et al., 2008). Scores for each coping type were computed by dividing mean endorsement for the coping type by the mean overall RSQ-A score, indicating their endorsement of coping on average. A score of one on voluntary engagement would indicate that voluntary engagement was endorsed as frequently as all other strategies. Scores greater than one indicated that voluntary engagement was endorsed more than other strategies, and below one indicated that voluntary engagement was endorsed less than other strategies. Means and standard deviations for each coping
type are available in Table 1.

Results
First, correlations between all variables were explored (see Table 1). Adolescents who reported more peer stressors also reported greater perceived distress as a result of these stressors. Furthermore, those with more peer stressors reported less use of voluntary engagement and more involuntary engagement and involuntary disengagement. Adolescents who reported more perceived distress reported more involuntary engagement, but there were no significant correlations between perceived distress and other coping strategies. SES was included in initial analyses but was not a significant predictor of coping and therefore was excluded from further analyses.

Perception of Peer Stress
Two-way (sex, AS group) Analyses of Variance were conducted to test for group differences in the extent to which adolescents experienced peer stress. As expected, AS adolescents reported experiencing more peer-related stress than non-AS adolescents (M_{AS} = 1.98, SD = 1.82; M_{non-AS} = 1.21, SD = 1.40), F(1, 191) = 6.01, p = .02, partial η_p^2 = .030. In addition, girls reported more peer-related stress than boys (M_{girls} = 1.63, SD = 1.74; M_{boys} = 1.07, SD = 1.13), F(1, 191) = 8.57, p = .004, partial η_p^2 = .043. Similar differences emerged with regard to perceived distress as a result of peer stress. AS adolescents reported marginally more distress than non-AS adolescents (M_{AS} = 2.15, SD = 0.94; M_{non-AS} = 1.80, SD = 0.94), F(1, 191) = 3.31, p = .07, partial η_p^2 = .017, and girls reported significantly more distress than boys (M_{girls} = 2.01, SD = 1.00; M_{boys} = 1.71, SD = 0.86), F(1, 191) = 5.29, p = .02, partial η_p^2 = .027.

Coping
A two-way (sex, AS group) Multivariate Analysis of Variance was conducted on the use of the four coping strategies. Overall, AS and non-AS adolescents used voluntary engagement more than the other coping strategies. However, as expected (see Figure 2), AS adolescents reported less voluntary engagement than non-AS adolescents, (M_{AS} = 1.09, SD = .16; M_{non-AS} = 1.20, SD = 0.20), F(1, 191) = 14.56, p < .001, partial η_p^2 = .071, and girls were more likely than boys to use voluntary engagement (M_{girls} = 1.19, SD = 0.18; M_{boys} = 1.14, SD = 0.21), F(1, 191) = 6.63, p = .01, partial η_p^2 = .034. In contrast, involuntary engagement and involuntary disengagement were not commonly used by adolescents overall. However, as expected, these strategies were more commonly used by AS adolescents who used involuntary disengagement (M_{AS} = 0.90, SD = 0.11; M_{non-AS} = 0.83, SD = 0.15), F(1, 191) = 11.09, p = .001, partial η_p^2 = .034, as well as involuntary engagement (M_{AS} = 0.95, SD = 0.12; M_{non-AS} = 0.88, SD = 0.14), F(1, 191) = 6.68, p = .01, partial η_p^2 = .055, significantly more than non-AS adolescents. Finally, contrary to expectations, there were no differences in the use of voluntary disengagement between AS and non-AS adolescents.

Discussion
This investigation explored the role of anxious solitude in adolescents’ responses to peer stressors. Consistent with hypotheses, AS compared to non-AS adolescents reported experiencing more peer stressors and more distress as a result of these stressors. Furthermore, in response to these stressors, AS adolescents reported less voluntary engaged coping and more involuntary engagement and involuntary disengagement. Thus, AS adolescents experienced more peer stress and used more involuntary coping strategies in response to these stressors.

<table>
<thead>
<tr>
<th>Table 1: Intercorrelations Among All Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>Child Sex</td>
</tr>
<tr>
<td>Peer Stress</td>
</tr>
<tr>
<td>Coping</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>AS</td>
</tr>
<tr>
<td>Peer Stress</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>Coping</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>AS</td>
</tr>
<tr>
<td>Peer Stress</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>Coping</td>
</tr>
</tbody>
</table>

Note. N = 195, *p < .05, **p < .01, ***p < .001.
Perception of Peer Stress
The findings that AS adolescents report more peer-related stress than non-AS adolescents were consistent with previous observational, peer-, and teacher-reported findings indicating that AS adolescents experience more peer exclusion and victimization and are less likely to receive positive responses from peers (Cillessen, Van IJzendoorn, Van Lieshout, & Hartup, 1992; Gazelle & Ladd, 2003; Gazelle et al., 2005; Spence et al., 1999). Additionally, AS adolescents reported marginally more distress in response to these stressors compared to non-AS adolescents. AS adolescents’ heightened sensitivity to rejection could contribute to elevated distress in response to peer mistreatment. However, more stressors and more distress in response to these stressors could also contribute to continuing maladjustment in AS adolescents over time (Gazelle & Rudolph, 2004).

Voluntary Engagement
Consistent with hypotheses, results indicated that non-AS adolescents used voluntary engagement significantly more than their AS peers. However both AS and non-AS adolescents used this more than any other type of coping. This suggests that AS adolescents regularly employed positive and proactive coping strategies, although they did so less frequently than non-AS adolescents. The results supported previous findings indicating that shyness and peer mistreatment were negatively related to approach-oriented forms of coping such as problem-solving and emotion regulation (Markovic et al., 2013; Sontag et al., 2008). This effect could be a result of AS adolescents’ previous experiences of peer rejection contributing to feelings of social helplessness within a conflict situation. AS adolescents may not employ voluntary engagement because they fear that such an active strategy could lead them to be perceived as socially awkward or incapable. However, this reluctance to engage in an active coping strategy may further contribute to their peer difficulties.

Involuntary Engagement
Although AS adolescents reported less voluntary engagement compared to their non-AS peers, they endorsed more involuntary engagement. Thus, they appeared to be affected by the stressor but unable to control their responses. The findings were consistent with previous investigations that found that AS adolescents used more involuntary physical, emotional, and cognitive engagement in response to both behavioral rejection and hypothetical peer stressors (Erath et al., 2012; Gazelle & Druhen, 2009; Kingsbury et al., 2013; Rose et al., 2007; Tan et al., 2012). AS adolescents could employ involuntary engagement more frequently because of a heightened anxiety and sensitivity to rejection (Tan et al., 2012), which could lead AS adolescents to experience uncontrollable distraction by a situation (Rose, 2002). Uncontrollable rumination and physiological arousal in response to peer stress could contribute to continued anxious solitude over time because AS adolescents may avoid peer interactions in an attempt to avoid these negative experiences.

Voluntary Disengagement
Contrary to hypotheses, there were no significant differences in the rates at which AS and non-AS adolescents employed voluntary disengagement and chose to avoid the stressor. This appeared to be a relatively common strategy for all adolescents and it may be an appropriate strategy when adolescents are faced with a problem that they do not have the cognitive, social, or emotional resources to directly cope with.
**Involuntary Disengagement**

Although there were no group differences in voluntary disengagement, AS compared to non-AS adolescents reported more involuntary disengagement in response to peer stress. Involuntary disengagement may be more likely to have negative outcomes because the lack of control may contribute to increased internalizing symptoms (Compas et al., 2001). Furthermore, involuntary disengagement may be particularly difficult to change in favor of more positive coping strategies because adolescents’ responses are automatic and not under their conscious control. AS group differences in involuntary, but not voluntary, disengagement suggested that differences in avoidant coping found in previous research (Kingsbury et al., 2013) might have been a reflection of group differences in involuntary disengagement. These results suggested that, although AS adolescents voluntarily choose disengagement at times compared to non-AS adolescents, they are also more likely to display automatic or uncontrolled disengagement in response to a stressor.

**Sex Differences**

This study replicated previous findings indicating that, although girls reported elevated peer stressors, they also reported coping more effectively with these stressors (Connor-Smith & Compas, 2002; Groer, Thomas, & Shoffner, 1992; Seiffge-Krenke, 2006; Wagner & Compas, 1990). The elevated stress levels in girls compared to boys could be specific to this time point because girls typically reach puberty earlier than boys (Dorn & Biro, 2011), and the physical and emotional changes associated with puberty combined with a school transition may lead to greater stress (Wagner & Compas, 1990). Furthermore, girls may experience elevated social stress specifically because, compared to boys, they are more focused on, and motivated by, relationships (Gnagey, 1980; Groer et al., 1992). This relationship focus could also make girls more motivated to resolve conflict, leading to higher endorsements of voluntary engagement. Some researchers have hypothesized that this motivational difference develops as a result of gender intensification in adolescence, when traditional sex-typed expectations become more pronounced. As a result, girls learn to express more emotions and be more sensitive (leading to more voluntary engagement in the face of peer stress), whereas boys focus on developing independence (Groer et al., 1992).

However, it is also possible that the self-reported measurement of peer stress and coping could have influenced these findings. Many previous studies using peer reports and other methods have indicated that early adolescent boys experience peer stress, particularly victimization (Hoglund, 2007; Russell, Kraus, & Ceccherini, 2010; Shell, Gazelle, & Faldowski, 2014), but they may be unwilling to report distress at such problems. Alternatively, although boys may objectively experience more peer mistreatment, girls may experience more subjective distress in response to the mistreatment they receive. Finally, it is important to note that there were no sex by anxious solitude interactions in coping patterns, demonstrating that the patterns for anxious solitude hold regardless of sex.

**Contributions and Limitations**

This study made several important contributions to the literature on AS adolescents’ responses to negative peer treatment. Consistent with previous research, AS adolescents reported more peer stressors and more distress as a result of these challenges. However, this study added to previous work by making an important distinction between voluntary and involuntary avoidance (disengagement). Although AS adolescents were more likely to use involuntary disengagement, they were no more likely than their peers to consciously choose to disengage from peer conflict. Finally, findings demonstrated that AS adolescents reported a consistent pattern of involuntary or automatic responses to peer stressors compared to their non-AS peers. Previous research has suggested that higher levels of involuntary coping may develop as a result of early temperament and go on to exacerbate anxious solitude. Behaviorally inhibited toddlers who demonstrate attention bias toward threat (e.g., angry faces) are more likely to develop social withdrawal in early childhood (Pérez-Edgar et al., 2011). In adolescence, this heightened focus on threat or stress could lead AS adolescents to engage in more involuntary automatic responses to peer stress. Furthermore, children’s responses to perceived threat may influence well-being. Behaviorally inhibited toddlers who engaged in low levels of attention shifting were unable to voluntarily control their attention away from threat and were more likely to later develop anxiety symptoms in preschool (White, McDermott, Degnan, Henderson, & Fox, 2011). Our findings contributed to the behavioral inhibition literature by...
extending the links between social withdrawal and involuntary engagement into early adolescence.

Several strengths also contributed to the validity of these findings. The analysis of proportions of endorsements, rather than raw endorsements, was advantageous in that it allowed for comparisons of relative endorsements across coping styles in addition to between-group comparisons. Additionally, the developmental timing of this study (immediately after the middle school transition) was important because the middle school transition may be associated with new peer challenges resulting from the restructuring of peer groups (Rudolph, Lambert, Clark, & Kurlakowsky, 2001; Shell et al., 2014). These challenges could prompt early adolescents to engage in coping more often than they did in elementary school. As peers become increasingly important during the period of early adolescence, peer stressors may become more frequent, and responses to these stressors may have important influences on later peer relations. Furthermore, coping becomes more stable and consistent in adolescence (Sontag et al., 2008). Therefore, it is important to identify and intervene with potential problems in coping early, before they are fully established.

These findings were important because identifying factors that put adolescents at risk for negative coping strategies could help determine possible intervention strategies to improve peer relations and coping in response to social stress. AS adolescents may be caught in a mutually exacerbating cycle of peer difficulties and negative coping. AS adolescents who avoid peer stress and do not directly attempt to resolve the issues may experience more negative treatment from peers within these encounters. In turn, this negative treatment may be a confirmation of AS adolescents’ social fears and could lead to feelings of social helplessness and less effective coping strategies. In particular, elevated rates of involuntary coping among AS adolescents suggest that this may be an area in which they could benefit from direct instruction. Teaching them to notice when they are engaging in involuntary coping processes, and then helping them consciously choose how to respond to a problem may decrease AS adolescents’ social evaluative concerns and could improve their peer relations. This is particularly important to do in early adolescence before AS adolescents develop more permanent patterns of coping in late adolescence and early adulthood (Sontag et al., 2008).

Despite these strengths, there were several limitations to this investigation. First, this study assessed coping strategies at a single point in time and did not explore changes over time. Longitudinal investigations are crucial to understanding developmental change in coping over time and how environmental changes such as a school transition may affect coping. In addition, longitudinal investigations could help further disentangle the relationship between peer mistreatment and coping (i.e., does negative peer treatment contribute to less effective coping or vice versa). Second, this study did not investigate the effectiveness of coping strategies in resolving peer conflict and improving peer relations and emotional well-being. Previous investigations (Compas, Connor-Smith, & Jaser, 2004; Erath et al., 2012; Gazelle & Druhen, 2009) have suggested that involuntary coping may concurrently exacerbate peer difficulties. However, this has yet to be directly tested. Finally, coping and anxious solitude were both assessed through self-reports, and therefore results could have been influenced by shared-method variance. Future studies should include peer or teacher reports to better understand the relationship between coping, anxious solitude, and peer treatment.

Nonetheless, this study highlighted several important findings with regard to early adolescents’ coping strategies in response to peer challenges. First, both AS and non-AS adolescents used voluntary engagement more often than other coping strategies when they faced peer stress. Therefore, even adolescents who have heightened sensitivity to peer rejection are more likely to directly address a peer problem than to use less effective coping strategies. Second, compared to their peers, AS adolescents coped less effectively with peer stress. They experienced more peer stress and used less voluntary engagement and more involuntary strategies when coping with these conflicts. Finally, this study made an important distinction between voluntary and involuntary disengagement. AS adolescents engaged in more involuntary (but not voluntary) avoidance. Therefore, they did not consciously choose to disengage from peer stress. Overall, although AS adolescents used positive coping strategies more than other strategies, they often used less effective coping strategies than their peers when faced with peer conflict.

References
Anxious Solitary Coping | Scott, Shell, and Gazelle


Author Note. Victor A. Scott and Madelynn D. Shell, Department of Social Sciences, The University of Virginia’s College at Wise; Heidi Gazelle, The Melbourne School of Psychological Science, The University of Melbourne, Australia.

This research was supported by grant 1K01MH076237 from NIMH to Heidi Gazelle.

Thanks to the children, parents, and teachers who participated in the project as well as the members of the Social Development Lab at the University of North Carolina at Greensboro.

Correspondence concerning this article should be addressed to Victor A. Scott, Department of Social Sciences, The University of Virginia’s College at Wise, Appalachia, VA 24216. E-mail: vasc3f@uvawise.edu
Among the myriad factors that influence the development of depression in children, maternal mental health has garnered considerable interest (Boyle & Pickles, 1997; Hammen & Brannan, 2003). Self, parent, and teacher reports all converge, revealing that preadolescent children of depressed mothers display higher levels of internalizing symptoms than children of nondepressed mothers (Billings & Moos, 1983; Breslau, Davis, & Prabucki, 1988; Lee & Gotlib, 1989). Considering the grave impacts that depression can have on children in the academic, social, and emotional domains (Diamond, Siqueland, & Diamond, 2003), understanding the development of depression in children is crucial.

Researchers have generated numerous explanations for the correlation between mother and child depression. Some posited that the link is biologically based, citing genetic heritability and alterations of neuroregulatory mechanisms during early development as the root cause (Allen, 1976; Goodman & Gotlib, 1999). However, twin and adoption studies have illustrated that genetics can only partially explain the relationship. For example, in one study, 60% of monozygotic twins were discordant for unipolar depression (Allen, 1976; Burbach & Borduin, 1986). Others maintained that both mother and child depression may occur due to shared experiences such as marital conflict, poverty, or abuse (Kendler, Karkowski, & Prescott,

**ABSTRACT.** Children of depressed mothers are more likely to develop depression than their counterparts whose mothers are not depressed. However, not every depressed mother raises a depressed child. Therefore, other factors likely moderate the probability that a child of a depressed mother will develop depression. Child attachment security to their mothers may be one factor that moderates the association between mother and child depression. Specifically, insecure attachments may place children of depressed mothers at a heightened risk for developing depression, and secure attachments may serve as a buffer against the onset of depression by promoting children’s emotion regulation and effective coping strategies during times of stress. In the current study, we investigated this possibility among a subsample of mother-child dyads in which the mother scored above the clinical cutoff for mild depression on the Beck Depression Inventory ($n = 30$, derived from a larger sample of $N = 107$) Child participants (ages 8–12 years old) completed the Child Attachment Interview and the Children’s Depression Inventory, and mothers completed the Beck Depression Inventory. Results indicated that attachment security moderated the association between mother and child depressive symptoms such that higher levels of maternal depressive symptoms were associated with higher levels of child depressive symptoms for children with lower attachment security ($\beta = .93$, $p = .02$), but not for children with higher attachment security ($\beta = .09$, $p = .65$). Results are discussed in terms of their contribution to the understanding of risk factors for child depression.

**Attachment Moderates the Association Between Mother and Child Depressive Symptoms**

Melanie K. Fox and Jessica L. Borelli*
Pomona College

---

*Faculty mentor

COPYRIGHT 2015 BY PSI CHI, THE INTERNATIONAL HONOR SOCIETY IN PSYCHOLOGY (VOL. 20, NO. 1/ISSN 2164-8204)
Maternal Depression, Child Depression, and Attachment

However, significant variability exists in the psychological functioning of children with depressed mothers, indicating that the relationship between mother and child depressive symptoms may be qualified by other important factors.

Maternal parenting behavior may be one such factor that aids in the transmission of maternal to child depression. Depression has been shown to impede many mothers’ ability to be affectionate and consistent in their responses to their children (Downey & Coyne, 1990). Depressed mothers may rely on strategies requiring less cognitive effort more frequently than do nondepressed mothers such as inflexibly enforcing obedience or withdrawing when facing child rebellion (Cohn, Matias, Tronick, Connell, & Lyons-Ruth, 1986). If these cognitive strategies are used, children may be at greater risk for developing many psychosocial problems themselves including depression (Downey & Coyne, 1990). Therefore, in the current study, we hypothesized that maternal depressive symptoms would be related to child depressive symptoms, in that the greater the maternal depressive symptoms, the greater the child’s depressive symptoms.

Importantly, not all children with depressed mothers develop depression. Therefore, it is useful to examine individual differences in mother-child relationship quality that impact the development of depression in children. Attachment theory, one framework for understanding variations in parent-child relationship quality, may provide insight into differences in the development of depression among children with depressed mothers. According to attachment theory, the quality of a caregiver’s response to their child’s requests for support and protection shapes the child’s behavior within the parent-child relationship as well as his or her thoughts and feelings related to attachment relationships in general (Bowlby, 1973). Parents of secure children are more sensitive and consistently responsive to their children’s needs and distress than are parents of insecure children (Ainsworth, Blehar, Waters, & Wall, 1978; see De Wolff & van IJzendoorn, 1997, for a meta-analysis). As a result, securely attached children learn that their distress and negative emotions can be resolved effectively by relying on their caregiver for support. Indeed, secure attachment in infancy is associated with better emotion regulation and the ability to maintain healthy relationships in the preschool years (Sroufe, Carlson, Levy, & Egeland, 1999; Vondra, Shaw, Swearingen, Cohen, & Owens, 2001). Insecure attachment in infancy and childhood predicts conduct problems in school, as well as adjustment and internalizing symptoms such as depression in childhood and adulthood, ostensibly because these children are less effective at regulating emotion and maintaining a balance of negative and positive emotional experiences (Cassidy, 1994; Downey & Coyne, 1990; Moss, Cyr, & Dubois-Comtois, 2004; Moss, Rousseau, Parent, St-Laurent, & Saintonge, 1998; Shaw & Vondra, 1995). Therefore, in the current study, we predicted that child attachment security would be related to child depressive symptoms, in that lower attachment security would be associated with more child depressive symptoms.

According to attachment theory, secure children, by virtue of their mothers’ sensitive parenting, learn that their distress is manageable and has a resolution (Bowlby, 1973). Therefore, it is possible that, when faced with maternal depression, children with a secure representation are able to regulate their emotions effectively and thus are no more likely to develop depression than their counterparts without depressed mothers. However, when faced with maternal depression, insecure children may be vulnerable to developing the disorder themselves because they have not learned to regulate their emotions and distress successfully early in life (Borelli, David, Crowley, & Mayes, 2010; Cassidy, 1994). Thus, it is possible that a secure attachment representation formed in infancy serves as a buffer against affective problems as the child ages such that maternal and child depressive symptoms may be more strongly associated among insecure children and less strongly associated among secure children.

Despite a plethora of research linking insecure attachment to emotion dysregulation and maternal depression to child depression (Allen, 1976; Cassidy, 1994; Downey & Coyne, 1990; Goodman & Gotlib, 1999; Moss et al., 2004), to date, only one study has examined how maternal depressive symptoms and attachment security interact in predicting child depressive symptoms. Milan, Snow, and Belay (2009) found that the severity of maternal depression interacted with children’s attachment security (measured by the Strange Situation at age 3; Ainsworth et al., 1978) to predict children’s self-reported depressive symptoms as they entered adolescence. Securely attached children who had severely depressed mothers displayed fewer depressive symptoms than insecurely attached children whose mothers were similarly depressed.

In the current study, we built upon previous
work by assessing attachment security and depression in middle childhood. To date, most research on attachment and depression has focused on the adolescent years, the developmental period associated with profound increases in the prevalence of depression (Hankin & Abramson, 2001; Nolen-Hoeksema & Girgs, 1994). Less is known about depressive symptoms in middle childhood although short-term longitudinal studies have suggested that middle childhood depressive symptoms and associated constructs such as rumination are associated with worsening depressive symptoms over time (Abela, Aydin, & Auerbach, 2007; Abela, Brozina, & Haigh, 2002). Late middle childhood (i.e., ages 8–12), the developmental period preceding adolescence, may represent an important phase for studying risk factors for the development of depression (Hankin et al., 1998). If the development of depressive symptoms can be better understood during the phase preceding adolescence, interventions may be able to stem the tide of emotion dysregulation, acting as preventative, rather than reactive.

Further, only recently has middle childhood become a prominent focus of attachment research (Kerns, 2008), primarily owing to the advent of assessment tools designed to probe attachment in middle childhood (Kerns, Abraham, Schlegelmilch, & Morgan, 2007; Kerns, Aspelmeier, Gentzler, & Grabill, 2001; Target, Fonagy, & Shmueli-Goetz, 2003). In terms of relationships, the transition from early middle childhood to adolescence is marked by a dramatic increase in the influence of peer interactions (Qualter, Brown, Munn, & Rotenberg, 2010), which has the potential to be a complicated transition for some youth. Therefore, the links between attachment and depressive symptoms in this age range may provide important insight regarding risk for relational and psychiatric distress during adolescence.

In the present study, we used a cross-sectional design to test two central hypotheses regarding the concurrent interrelations of maternal depressive symptoms, child depressive symptoms, and child attachment security. Consistent with extant literature, we hypothesized that maternal depressive symptoms and attachment security would be uniquely associated with children’s depressive symptoms in middle childhood and that attachment security would moderate the association between mother and child depressive symptoms, such that maternal and child depressive symptoms would be more strongly positively associated among children with lower levels of attachment security in middle childhood.

**Method**

**Participants**

Mother-child dyads (N = 107) participating in a larger study of children’s socioemotional adjustment were recruited through flyers and Internet postings in southern California. Parents and children each received a $50 compensation for their time. Our full sample included children (53% girls) who ranged in age from 8 to 12 years old (M = 9.74, SD = 1.48). The sample was socioeconomically diverse, with 40% of participants having a household income of less than $40,000 per year. Further, 29% of participants had a college degree or higher, and 57% of the sample was married or cohabiting with a romantic partner. Our sample was comprised of 34% Latino, 30% European American, 16% African American, 8% biracial, and 12% of participants who declined to report their ethnicity. Due to our desire to examine our hypotheses among mothers reporting high levels of depressive symptoms, we restricted our sample to mothers whose scores on the Beck Depression Inventory (BDI-II; Beck, Steer, & Brown, 1996) fell above the clinical cutoff for mild depression (> 14). On average, the mothers in our clinical subsample were 40 years old (M = 39.14, SD = 6.69), had two children (M = 2.37, SD = 0.85), were married or had a domestic partner (52%), and had not been divorced (55%).

**Procedure**

The present study was part of a larger cross-sectional study that involved two 1.5 hr sessions, approximately one to two weeks apart. Institutional review board approval was received for the study. During the first session, participants and their parents completed assent, consent, and demographic forms. Children then completed the Child attachment Interview (CAI) with a trained research assistant. Afterward, both child and parent completed a battery of questionnaires by hand including self-report assessments of depressive symptoms (Children’s Depression Inventory [CDI] and Beck Depression Inventory [BDI-II]).

**Measures**

**Child attachment.** The CAI (Shmueli-Goetz, Target, Fonagy, & Datta, 2008) is a semistructured interview for 8- to 13-year-olds and is parallel in structure to the Adult Attachment Interview (George, Kaplan,
Maternal Depression, Child Depression, and Attachment | Fox and Borelli

& Main, 1984). It consists of 19 questions regarding the children’s current and past experiences with their primary caregivers (e.g., “Have you ever felt like your parents don’t really love you?” and “What happens when mom and/or dad gets upset or angry?”). The interview is approximately 30 min long and is videotaped, transcribed, and coded on eight 9-point scales based on the quality of the narrative the child generates (e.g., narrative coherence, balance of positive/negative references to attachment figures). The narrative coherence score can be used as a continuous measure of attachment security (Borelli, Crowley, et al., 2010), which is how it was used in the current study. CAI classification is correlated with child attachment security (Borelli, Crowley, et al., 2010), which is how it was used in the current study. CAI classification is correlated with child attachment security as measured in the Separation Anxiety Test (Wright, Binney, & Smith, 1995) and has been associated with variations in children’s socioemotional functioning (Borelli, Crowley, et al., 2010; Borelli, David et al., 2010).

In the present study, CAIs were administered by trained undergraduate research assistants and were coded by a researcher who had been certified as a reliable CAI coder. A subset (n = 16) of the CAIs were double-coded by another certified CAI coder, and the interrater reliability was high: four-way Kappa = .91, p < .001, Intraclass Correlation Coefficient for narrative coherence scale = .83, p < .001.

**Parental depressive symptoms.** Parents completed the BDI-II, a 21-item self-report scale, which assesses cognitive, behavioral, affective, and somatic dimensions of depression. Each question has four response choices (e.g., “I do not feel sad,” “I feel sad,” “I am sad all the time and I can’t snap out of it,” and “I am so sad and unhappy that I can’t stand it”). Each item is scored on a 4-point scale from 0 (no signs of symptom) to 3 (full endorsement of the symptom). The BDI-II has been found to demonstrate high internal consistency (α = .92; Beck et al., 1996). Adequate validity has been demonstrated, and diagnostic discrimination has been established (Scogin, Beutler, Corbishley, & Hamblin, 1998). In this sample, Cronbach’s α was .88.

For the current analyses, due to our non-clinical pool, we limited our sample to mothers who scored above the mild clinical cut-off for depression (> 14) on the BDI-II because the link between maternal depressive symptoms and child depressive symptoms may be most important to understand among mothers experiencing clinically significant depressive symptoms.

**Child depressive symptoms.** Children completed the CDI (Kovacs, 1992), a 27-item questionnaire that assesses the behavioral, cognitive, emotional, and physiological features of depression in children ages 7 to 17. Children chose one of three statements which best described their symptoms over the past 2 weeks (e.g., “I am sad once in a while,” “I am sad many times,” “I am sad all the time”). Responses were coded on a scale ranging from 0 to 2, with higher scores indicating more severe depressive symptoms. A score of 13 or above is thought to indicate the cutoff for clinical depressive symptoms (Kovacs, 1992). The psychometric properties of the CDI have been reported in the literature (Kovacs, 1992). Cronbach’s α for this sample was .86.

## Results

### Descriptive Data

Table 1 includes means and standard deviations for all primary variables in the current study. We computed zero-order correlations for all study variables. Importantly, child age and sex were unrelated to child or maternal depressive symptoms and attachment. Due to the fact that child age and sex were unrelated to our independent or dependent variables, we did not control for them in subsequent analyses. In the sample at large (N = 107 dyads), mothers reported an average of 9.84 (SD = 7.91) depressive symptoms, and children reported an average of 8.12 (SD = 6.17) depressive symptoms.

### Concurrent Associations Among Mothers and Children

We hypothesized that, for mothers with depressive symptoms, maternal depressive symptoms would predict child depressive symptoms. Furthermore, we predicted that, for children of mothers with clinically significant depressive symptoms, child attachment security would be associated with child depressive symptoms. Results of a hierarchical linear regression revealed that mothers’ scores on

<table>
<thead>
<tr>
<th>TABLE 1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sample Characteristics</strong></td>
</tr>
<tr>
<td>Demographic characteristics</td>
</tr>
<tr>
<td>Parent’s age</td>
</tr>
<tr>
<td>Child’s age</td>
</tr>
<tr>
<td>Maternal depressive symptoms</td>
</tr>
<tr>
<td>Child depressive symptoms</td>
</tr>
</tbody>
</table>

**Maternal Depression, Child Depression, and Attachment | Fox and Borelli**
the BDI-II were not a significant predictor of child depressive symptoms, $R^2 = .04, p = .28$. Similarly, a second regression revealed that attachment security was not significantly associated with child depressive symptoms, $R^2 = .01, p = .91$

**Moderating Variable**

Next we conducted a hierarchical linear regression to evaluate whether attachment security moderated the association between maternal depressive symptoms and child depressive symptoms. Consistent with Holmbeck’s (2002) recommendations for evaluating moderators, prior to conducting analyses, we centered continuous independent variables. The main effects of maternal depressive symptoms and attachment security were entered as a first step, and the interaction between maternal depressive symptoms and attachment security was entered as a second step. Analyses revealed a significant Maternal Depressive Symptoms x Attachment Security interaction effect for the prediction of child depressive symptoms, $\Delta R^2 = .12, p = .03$. Holmbeck’s (2002) guidelines for post-hoc probing of significant moderational effects were used to examine the nature of this interaction. This included (a) computation of new product terms at different levels of the moderator variable (e.g., for high security and low security), (b) computation of simple slope estimates by including these new product terms in reduced regression models (e.g., including covariates and only significant predictors), and (c) examining the statistical significance of these slopes at different levels of the moderator variable. For the interaction effect of Maternal Depressive Symptoms x Attachment Security, results revealed a significant slope for children with low security, $\beta = .93, p = .02$, but not for children with high security, $\beta = .09, p = .65$, indicating that higher maternal depressive symptoms were associated with higher child depressive symptoms for children with low security, but not for children with high security (see Figure 1 and Table 2).

**Discussion**

We conducted the present study to provide a cross-sectional extension of Milan and colleagues’ (2009) paper examining whether child attachment security moderated the association between maternal and child depressive symptoms. In the present project, we examined whether concurrent associations between school-aged children’s and mothers’ depressive symptoms were moderated by child attachment security.

With respect to our first hypothesis, results indicated that attachment security was not associated with child depressive symptoms. Although our analyses focused on the subsample of mothers with mild depressive symptoms, the lack of association between child attachment security and child depressive symptoms held for the larger sample. These findings, inconsistent with the literature (Allen, 1976; Borelli, Crowley, et al., 2010; Cicchetti & Toth, 1995; Goodman & Gotlib, 1999; Lyons-Ruth, Easterbrooks, & Cibelli, 1997), are somewhat difficult to reconcile. One potential explanation consists of the demographic make-up of the sample. In contrast to previous work examining attachment and depressive symptoms in middle childhood (Borelli, Crowley, et al., 2010), as well as the link between maternal and child depressive symptoms in middle childhood (Burbach & Borduin, 1986), the present study comprised a socioeconomically diverse sample. Previous research has hinted that the association between maternal behavior, attachment security, and child psychopathology may be different in samples with lower socioeconomic status (SES). For example, De Wolff and van IJzendoorn (1997) reported that the association between maternal sensitivity and attachment security was moderated by SES; the lower class samples generated a smaller effect size than did middle class samples. Research has also highlighted the need to consider the accumulation of social risk factors in the prediction of child psychopathology.

**FIGURE 1**

Post-Hoc Test of the Interaction Between Maternal Depressive Symptoms and Attachment Security on Child Depressive Symptoms

Note. BDI-II = Beck Depression Inventory.
Maternal Depression, Child Depression, and Attachment

likely to have effective skills for regulating emotion high attachment security, and therefore are more interest because it suggested that, if children have insecure attachment preceded the concordance in depressive symptoms. However, this finding is of note that this study was cross-sectional. Thus, we cannot infer that child insecure attachment preceded the concordance in depressive symptoms. However, this finding is of interest because it suggested that, if children have high attachment security, and therefore are more likely to have effective skills for regulating emotion quality of verbal responsiveness in 1-year-olds, even more than involved parenting did. These findings suggested that the physical environment, as well as relational factors, could contribute in important ways to attachment and emotional experience in lower SES samples. One study attempted to examine these cumulative stressors as predictors of attachment security in a low-income sample (Shaw & Vondra, 1995). The authors did not find any significant associations between individual stressors and attachment security. Further, the results suggested that families with two stressors were more likely have secure infants, whereas families with three or four stressors were more likely to have insecurely attached infants. These studies indicated that the influence of attachment security on psychosocial functioning may operate differently among lower SES families.

Consistent with previous research (Milan et al., 2009), attachment security moderated the relationship between mother and child depressive symptoms. Post-hoc analyses revealed that maternal depressive symptoms were only positively associated with children’s depressive symptoms when children had low attachment security with respect to their attachment relationships. It is of note that this study was cross-sectional. Thus, we cannot infer that child insecure attachment preceded the concordance in depressive symptoms. However, this finding is of interest because it suggested that, if children have high attachment security, and therefore are more likely to have effective skills for regulating emotion and/or reduced emotion reactivity (Borelli, David, et al., 2010; Cassidy, 1994), they may be more likely to develop depressive symptoms than children with healthy mothers. These results highlighted the importance of considering moderating variables in the transmission of clinical disorders between parents and children. Our findings added to the current literature in providing the first cross-sectional support for Milan and colleagues’ (2009) longitudinal study. Further, our study provided an important contribution to the literature in that we assessed attachment security in middle childhood, an understudied developmental phase. Understanding depressive symptoms during middle childhood may help prevent or better treat adolescent and adult depression. Further, our study illuminated the limitations of purely hereditary or event-based models for the transgenerational transmission of depression because it indicated that moderating factors, in particular relationship quality, play a key role in the development of depression in children.

Although our findings were preliminary, if replicated, they would suggest that interventions for depressed children that target attachment relationships may be particularly important. For example, family-focused therapy for depressed children such as Attachment-Based Family Therapy (ABFT; Diamond et al., 2003) may be useful. ABFT is a developmentally based, semi-structured intervention tailored to the particular needs of depressed children. Because parent-child relationships can influence the development of depressive symptoms in children, therapy focused on restructuring and enhancing these relationships is important.

Limitations of the current study should be noted and addressed in future research. First, it would be of interest to measure the moderating effect of child attachment among a clinical sample of mothers in addition to community samples. Second, it would be interesting to assess depression using standardized diagnostic interviews, rather than relying on self-report data exclusively, which can be influenced by many uncontrollable factors such as desire to please or lack of insight (Podsakoff & Organ, 1986). To further explore issues raised in the present study, it would also be fruitful to measure the course of the mother’s depression at several time points throughout the child’s development to determine if length of depressive episodes or the age of the child during the mother’s episodes helps to explain the transmission of symptoms as well. Furthermore, given that previous research
has suggested that the link between relationship quality and psychopathology may be more complex among lower SES families, future research should attempt to consider the unique stressors and factors influencing lower SES families to better understand patterns of psychopathology. It would also be useful to examine the effects of fathers’ depressive symptoms on the model as well.

When considering the findings of the present study, it should be noted that mothers with depressive symptoms and insecure children may differ in other ways from mothers who have depressive symptoms and secure children. These differences could explain the observed link between mother and child depressive symptoms among insecure children, but not among secure children. For example, perhaps mothers who have secure children have greater support systems, more parental training, or better treatment available. Alternatively, perhaps mothers who currently exhibit depressive symptoms and have secure children were not depressed during their children’s early years of life.

Despite these limitations, the current study contributed to the literature by suggesting the importance of examining moderating factors in understanding the intergenerational transmission of depressive symptoms from mother to child. Mother-child relationship quality, specifically high levels of attachment security, may serve as a buffer for the transmission of symptoms. This study also extended the current attachment and depression literature to middle childhood. If replicated, these findings would have implications for interventions, therapy, and relational strategies for depressed mothers and children. Future work should further explore this phenomenon throughout development.

References


Assessment approaches and associations with mood and emotion regulation. *Attachment and Human Development, 9*, 33–53. doi:10.1080/14616730901151441


**Author Note.** Melanie K. Fox and Jessica L. Borelli, Department of Psychology, Pomona College, CA.

Correspondence concerning this article should be addressed to either Melanie K. Fox, 217 Theresa St., San Francisco, CA 94112. E-mail: melaniekfox@gmail.com, or Jessica L. Borelli, Department of Psychology, 647 N. College Way, Claremont, CA 91711. E-mail: Jessica.borelli@pomona.edu
Researchers who have studied human memory have found that pictures are remembered better than words (Paivio, Rogers, & Smythe, 1968). This phenomenon is known as the picture superiority effect. This effect is exemplified when participants are presented with a picture of an apple or the word apple, and, when given a memory test, they recall the picture more often than the word. The fact that memory for pictorial stimuli is superior to that of verbal stimuli implies that pictures have some quality that renders them easier to remember than verbal material, even if said material is equal in detail (Joseph, Waln, & Stone, 1984). The picture superiority effect has been observed across varied memory tests including free recall (Joseph, Joseph, & Beasley, 1982; Paivio & Csapo, 1973), cued recall (Intraub & Nicklos, 1985; Nicolas, 1995), and recognition (Durso & Johnson, 1980; Hockley, 2008). However, evidence has also shown how the effect can be reversed and/or eliminated (Weldon & Roediger, 1987). Collectively, prior research has provided important information that can be used to construct viable theories about how human memory works and it is important to conduct further research on the picture superiority effect to give a clearer and more in-depth understanding of memory processes.

Early theories attempting to explain picture superiority focused on the encoding stage of memory. For example, the dual-encoding hypothesis (Paivio, 1971) assumed that pictures are remembered better than words because they are encoded twice, once in image form and again in verbal form. Another model, proposed by Nelson,
Reed, and Walling (1975), introduced the sensory-semantic model of memory, which posits that the strong sensory code and sensory distinctiveness makes pictures stand out in a participant’s mind more than words and produces superior recall when compared to words.

Eventually, researchers moved toward explaining the picture-superiority effect through examining the retrieval stage. For example, Weldon and Roediger (1987) conducted four experiments in which they focused their attention on memory retrieval. In doing so, they gathered support for the transfer-appropriate processing model of memory. This model posits that, if encoding and retrieval cues do not match, recall will be diminished. For instance, when participants read words during a study phase, they have greater recognition for the words at test if they are presented in visual form as opposed to an auditory form. Weldon and Roediger’s first experiment demonstrated how altering the retrieval demands can reverse the picture-superiority effect. Participants studied either words or pictures and then completed a word-fragment completion task (an implicit memory task). The participants who had studied pictures completed fewer word fragments than participants that received words to study. Thus, picture superiority was changed to word superiority by only changing the retrieval task. Such findings supported the transfer-appropriate processing model of memory.

Those who support the transfer-appropriate processing model have contended that the picture superiority effect is eliminated because study and test tasks do not match either physically or conceptually. In other words, when conducting a priming task, the type of stimuli given at study must match those given at test or picture superiority will not be realized. The matching of encoding and retrieval tasks and stimulus mode at study and test is called congruency and several studies supported the notion that congruency is an important factor in bringing about the picture superiority effect (Sternberg, 2006; Weldon & Coyote, 1996; Weldon & Roediger, 1987; Weldon, Roediger, & Challis, 1989).

However, congruency is only one factor that contributes to picture superiority. The kind of retrieval task and how the task is constructed can affect whether the effect is produced or eliminated. When conducting their fourth experiment, Weldon and Roediger (1987) used word-fragment completion and picture-fragment identification tasks that were thought to be equally focused on the perceptual characteristics of the stimuli. However, the word-fragment tasks were different because they omitted whole letters from the words, whereas the picture-fragment tasks degraded the picture by eliminating parts of it. Word fragments and picture degradations may not be equal in perceptual characteristics. In fact, it has been suggested that the method by which the researchers created their retrieval tasks was flawed because they were not equally degraded (Sternberg, 2006).

Sternberg (2006) recognized that removing letters from a word was not a perceptually based change, and he conducted an experiment in which he attempted to remedy the flaw in methodology. To degrade his pictures and words, he removed a strip of the item from the top and the bottom and the participants had to determine if the remaining strip had been reversed or not. He found that participants in the picture-congruent condition had a higher rate of identifications than those in the word-congruent condition. However, it is possible that Sternberg did not find a picture superiority effect when using degraded pictures because of the methodology he used to degrade the pictures.

There is a way to systematically create picture degradation in a manner that is consistent with the word degradation by using the recognition-by-components model of image processing. Biederman (1987) composed a theory known as recognition-by-components to attempt to explain how this happens. This theory was based on the premise that individuals see things in parts and put them together to form whole objects. The components are called geons, and they consist of simple compositions of concavities, vertices, edges, and other nonaccidental properties (i.e., edges that do not change with orientation). Humans put these geons together implicitly to recognize more complex three dimensional objects. It is the components that make up the geons such as the concavities and nonaccidental properties, which are essential to the process by which humans integrate the geons and subsequently recognize the object. Without those properties, object recognition could not take place. Biederman and Cooper (1991) discovered that a lack of crucial components such as concavities and nonaccidental properties led to longer identification reaction times and more identification errors.

The present experiment used these findings to find a more accurate way to degrade pictures and address Weldon and Roediger’s (1987) flawed methodology. Systematically degrading pictures...
and words equally by removing the concavities and nonaccidental properties was expected to ensure that the picture identification task was truly equated to the word identification task in terms of recognizability based on surface features because only the perceptual properties of the stimuli would be altered.

The present study also explored what role congruency plays in the picture superiority effect. Weldon and Roediger (1987) specifically explored the congruency factor in one of their experiments. In one experimental condition, they gave participants who had studied pictures word-fragment completion tasks at test. Therefore, the modes of the presentation stimulus were not similar between study and test. In doing so, they found that pictures did not prime for degraded words and words did not prime for degraded pictures.

Due to the findings by Weldon and Roediger (1987) and Sternberg (2006), it is clear that congruency and retrieval tasks that equally focus on the perceptual features of the stimuli are both important to producing picture superiority on implicit recognition tasks. The present experiment aimed to determine whether an implicit task could result in the picture superiority effect if the task was constructed correctly. In the experiment, two factors were manipulated: the mode of the stimuli, pictures or words, presented at study and at test, and whether the stimuli were congruent at study and at test. The pictures and words were intact at study and degraded at test, and the participants were asked to identify them. The accuracy with which the object or word was identified in its degraded state was measured along with the time it took (reaction time) for the participants to recognize the degraded image.

A main effect for picture superiority was hypothesized in that participants in picture-congruent and picture-noncongruent conditions should perform better (a higher proportion of correct recognitions and shorter reaction times) than the participants in word conditions. There should also be a main effect for congruency, which would be achieved when the stimuli matched at study and test. We expected that the participants would perform better when the stimuli matched at study and test. Finally, we hypothesized that an interaction would occur whereby the picture-congruent condition would show the highest recognition score and best reaction time when compared to all other conditions. However, we predicted that the word-noncongruent condition would show the poorest performance (the lowest proportion of correct recognitions and the longest reaction times) when compared to all other conditions. The interaction was expected to occur because picture superiority and congruency both work in favor of the picture-congruent condition, whereas neither of these factors existed in the word-noncongruent condition.

**Method**

**Participants**

Participants were 39 undergraduate students from introduction psychology classes at a private southeastern university. Twenty-seven participated in the experiment and 12 participated in the preliminary norming task. Six participants in the preliminary norming task and six participants in the experiment participated voluntarily after being invited by the researcher. The remaining 27 participants received one credit toward their grade in introduction to psychology for their participation. All participants were 18 years old or older and had normal or corrected vision. The average age of the participants was 19.9 years ($SD = 1.9$). Of the 39 total participants, 12 (31.9%) were men and 27 (69.2%) were women. Most participants were European American ($n = 19, 48.7%$). The remaining participants were African American ($n = 11, 28.2%$), Asian ($n = 8, 20.5%$) and American Indian ($n = 1, 2.6%$).

**Design**

A 2 (picture, word) x 2 (congruent, noncongruent) within-subjects design was employed in the present experiment. The independent variables were stimuli mode and mode congruency. The dependent variables were reaction time and number of correct identifications.

**Materials**

A list of 80 words and pictures were selected from the Snodgrass and Vanderwart (1980) corpus to use in the study and test phases. The pictures were 40 simple line drawings of common objects such as a *balloon* or a *tomato*, and the 40 words were the names of these pictures.

**Preliminary norming phase.** Fifty pictures and the corresponding 50 words (100 items) from the 260 were selected from Snodgrass and Vanderwart (1980). These 100 items were degraded based on Biederman’s (1987) object recognition theory. The concavities and the vertices, which are nonaccidental properties, were removed. To do this, each...
Degraded Image Identification | Brundage and Barile-Spears

picture from an electronic copy of the Snodgrass and Vanderwort (1980) corpus was copied and then pasted into Microsoft Paint®. Then it was magnified, and the eraser was used to remove the vertices and the parts that were concave. To degrade the verbal stimuli, the entire alphabet was typed in capital letters in 72-point Arial font into a text box in the paint program. Then the eraser was used to remove the concavities and the vertices of all the letters. The letters from the degraded alphabet were put beside each other to make the degraded words (see Figure 3 for examples of degraded images).

The participants for the preliminary norming task were presented with these degraded images on the computer screen. The participants were not primed for these images, and they were asked to identify them. The usable images were selected from the pictures that had between 50% and 83% recognizability during the preliminary norming phase (the center of the distribution). From these images, reaction times that were identified as outliers because of extreme scores that were far greater than the other scores were removed. The final 20 pictures had an average of 71.65% recognizability. Ten of these pictures were chosen to be in picture form during the study phase and the other 10 were selected to be in word form. During the study phase, all participants saw the same pictures and words, but they were presented in a different random order for each participant. The 20 pictures and the corresponding 20 words were used in the test phase of the actual experiment. Another 40 items (20 pictures and 20 words) were used as distracter items during the test phase. The recognizability for the remaining 20 items (10 pictures and 10 words) was not considered because they were used during the delay test task.

Study materials. Participants studied the list of 10 words and 10 pictures that were selected during the preliminary norming phase. The pictures and words were unrelated. Only the pictures and words that were normed for recognizability in the degraded form were used in the study task. At study, the stimuli were clear and free of degradation.

Test materials. At test, the participants were presented with the same list of 10 words and 10 pictures that they received at study. However the words and pictures were presented twice, once in word form and once in picture form, in a random order. At test, all items were degraded. Forty distracter items were mixed in with the test items. The distracter items were degraded in the same way as the other test items but their ability to be recognized was not considered because these stimuli were not primed. All of the participants received the same list of degraded pictures, words, and distractor items. Each list was presented in a different random order.

Delay test task stimuli. The participants were given a delay test task so that they would not become aware of the fact that the experiment was focused on memory. During the delay test task, participants were presented with 20 images that were degraded in the same way as the test stimuli. These 10 pictures and 10 words appeared nowhere else in the experiment. These stimuli were presented for 5 s per item with an interstimulus interval of 1 s.

Apparatus
Study and test items were presented on 17 in. monitors of Dell Optiplex™ 760 desktop computers. The instructions and the stimuli were presented using the computer software MediaLab™. To take the most accurate reaction times, computer software called DirectRT™ was employed.

Procedure
Participants entered the computer lab and sat at a computer of their choice. They were given instructions that asked them to read words and view pictures that were presented on the screen. They were not informed that they would be given a memory test. The participants were told that they were being tested on object recognition.

During the study phase, participants were shown a list of 10 pictures and 10 words, randomly presented. The items were presented at a rate of 5 s per item with an interstimulus interval of 1 s. After the study phase, participants were given the delay test task in which they were presented with and asked to identify a set of degraded pictures and words. Finally, the participants were randomly presented with the test items and distracter items for identification. The study, delay, and test phases were presented continuously, one after the other.

During the test phase, the participants’ task was to identify the degraded item by typing their response into the computer. At the first key stroke of their response, the reaction time for the item was recorded. The computer program also recorded the actual response. Participants responded to all of the stimuli in less than 30 s, and all of the responses are included in the results.

When calculating the proportion of correct responses, any word that exhibited a typing error was still counted as correct if the word appeared
to be correct despite the error. Also, synonyms were counted as correct in the picture conditions. For example, if the picture was of a couch and the participant typed sofa, the response was recorded as correct. Synonyms were not counted as correct in the word conditions because, if the participant saw the degraded word couch and typed sofa, it would be incorrect to identify the word read as sofa. After the completion of data collection, participants were debriefed. All procedures were approved by the university’s institutional review board.

**Results**

To determine the effects of stimulus mode and mode congruence on the participants’ ability to identify degraded images, the mean reaction times and the proportion of correct identifications were measured. The scores from the proportion correct data could range from zero to one. The data were analyzed to test the three hypotheses: a main effect for picture superiority, a main effect for congruency, and an interaction effect. A two-way repeated measures Analysis of Variance (ANOVA) was employed to analyze the reaction time data. A second two-way repeated measures ANOVA was employed to analyze the correct identification data. The alpha level was set at .05.

All of the participants’ reaction time data for each identification task were analyzed for outliers. Because reaction time analyses are sensitive to long reaction times (Ratcliff, 1993), all reaction time values whose distance from the nearest quartile was greater than 1.5 times the interquartile range were removed. Then the mean reaction times for each participant were computed across conditions. The means for each condition were then computed. These results are displayed in Figure 1.

The picture-congruent condition \((M = 2,890, SD = 489)\) had a slower mean reaction time (in ms) than the picture-noncongruent condition \((M = 2,710, SD = 576)\). The word-congruent condition \((M = 2,481, SD = 555)\) had a faster mean reaction time than the word-noncongruent condition \((M = 3,231, SD = 631)\). The results of the ANOVA showed a significant interaction, \(F(1, 26) = 30.12, p < .001, \eta_p^2 = .54\), and a significant main effect for congruency, \(F(1, 26) = 49.05, p < .001, \eta_p^2 = .65\). There was no significant main effect for stimuli mode, \(F(1, 26) = 0.88, p = .36, \eta_p^2 = .03\).

To investigate the interaction further, a series of paired-samples \(t\) tests revealed a significant difference among all pairs of comparisons except between the picture-congruent and picture-noncongruent conditions (see Table 1). The significance level was split between the six tests to control for Type I error using the Bonferroni correction method for determining significance.

Mean accuracy across conditions was also analyzed. Figure 2 displays the mean accuracy for each condition. When the pictures matched \((M = 0.93, SD = 0.10)\) the mean of proportion correct responses was higher than when the pictures did not match \((M = 0.98, SD = 0.04)\), and the word matching condition \((M = 0.99, SD = 0.02)\) had a high mean of proportion correct responses than when the words did not match \((M = 0.83, SD = 0.12)\). The results of the ANOVA showed a significant interaction, \(F(1, 26) = 18.00, p < .001, \eta_p^2 = .41\), and a significant main effect for stimuli mode, \(F(1, 26) = 6.76, p = .015, \eta_p^2 = .21\).

Given that a similar pattern of results was found for proportion correct scores, once again a series of paired-samples \(t\) tests, using the Bonferroni correction, was calculated to further understand the interaction. The tests revealed a significant difference between the means of the picture-congruent condition and the word-congruent condition, the picture-congruent condition and the word-noncongruent condition, the picture-noncongruent condition and the word-noncongruent condition.
condition, and the word-congruent condition and the word-noncongruent condition. There was no significant difference between the means of the picture-congruent condition and the picture-noncongruent condition, or the picture-noncongruent condition and the word-noncongruent condition (see Table 2).

**Discussion**

The transfer-appropriate processing model of memory posits that changing the retrieval task will eliminate or reverse the picture superiority effect (Weldon & Roediger, 1987). In the present experiment, this model of memory was examined further. The purpose of this study was to show that a picture superiority effect could be produced when using an implicit memory test if the images were degraded properly. Prior research had no theoretical support for the ways in which they degraded the images used in their experiments. However, the present experiment used the theory of object recognition (Biederman, 1987) as the theoretical bases of degrading the images. By removing the concavities and the nonaccidental properties of both the words and pictures, all the images were degraded based on their perceptual properties and not their conceptual properties.

The results showed that congruency in the picture condition had little effect because the participants’ performance did not differ. However, the results also showed that congruency was the defining factor in the word condition. This was evidenced by the fact that the word-noncongruent condition yielded the worst performance of any of the conditions. Conversely, the word-congruent condition yielded the best performance. Thus, these data demonstrated that congruency could be the key factor that determined whether participants in these conditions were able to remember the previously viewed stimuli.

Another explanation for the performance in the word-congruent condition could be that reading processes operate differently than picture processing. Because reading is an automatic process, the word matched condition performance was at an advantage. In fact, Carr, McCauley, Sperber, and Parmelee (1982) suggested that word naming and picture naming are fundamentally different in terms of processing mechanisms. They contended that word naming is a perceptual process and that picture naming uses the semantic route to identify pictures. Thus, words are read faster because they are able to bypass semantic processing, accounting for the better performance in the word matched condition in the current study. It is evident that, once congruency was removed, the participants’ performance suffered significantly. Thus, the main effect for congruency that was exhibited for both dependent measures might only have been an artifact of the interaction.

These results demonstrated that degrading images by removing concavities and nonaccidental properties had a significant effect on the participants’ ability to quickly and accurately identify the

---

**TABLE 1**

<table>
<thead>
<tr>
<th>Paired-Samples t Tests for Reaction Time Data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Comparison</strong></td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>Picture congruent vs. Picture noncongruent</td>
</tr>
<tr>
<td>Picture congruent vs. Word congruent</td>
</tr>
<tr>
<td>Picture congruent vs. Word noncongruent</td>
</tr>
<tr>
<td>Picture noncongruent vs. Word congruent</td>
</tr>
<tr>
<td>Word congruent vs. Word noncongruent</td>
</tr>
</tbody>
</table>

*Note:* Indicates significance adjusted for using the Bonferroni method.
pictures and words. Although the main effect for pictures versus words was not significant for the reaction time data, the post-hoc $t$ tests revealed that the picture-superiority effect was still realized. This was because the pictures, after participants were primed for them, were easy to identify regardless of whether they were congruent or noncongruent at test. The pictures clearly presented sensory properties that allowed them to be remembered efficiently. Nelson et al. (1975) first posited that pictorial stimuli stand out in participants’ minds more because of the distinct properties of the image. The pictorial images in the current experiment evidently stood out in the minds of the participants more because participants had better memory for the pictures. This finding suggested that picture superiority is important and has a great effect on the memory of participants that it overrides the congruency manipulation.

Despite the fact that pictorial stimuli had a clear effect on the participants’ ability to remember a presented image, the present study, like the studies of other researchers (Sternberg, 2006; Weldon & Coyote, 1996; Weldon & Roediger, 1987), did not produce a true picture superiority effect. This effect would only be exhibited if both the picture conditions (congruent and noncongruent) had better results than both word conditions. However, the effect was not totally eliminated. The influence of the pictorial stimuli was still strong enough to get better results than the word-noncongruent condition and the picture superiority effect was not reversed.

The main limitation of the present study was the sample size. Having only 27 participants in the actual experiment likely influenced the results of the study. If the sample size had been larger, a true picture superiority effect might have been achieved. Having only 12 participants in the preliminary norming phase might have influenced the recognizability statistics for each stimulus.

The results from the present experiment have significant implications for the transfer-appropriate processing model of memory. This theory, first posited by Weldon and Roediger (1987), states that, if encoding and retrieval cues do not match, recall will be diminished. The results of the present experiment partially supported this theory. The support was shown by the poor performance of the word-noncongruent condition. The mismatched encoding and retrieval cues in this condition did, indeed, diminish the participants’ ability to remember the stimuli. The theory was not supported by the fact that there was no difference between the two picture conditions. Even when the study and test stimuli were mismatched, the participants’ performance for both dependent measures remained about the same. Their ability to remember the stimuli was not significantly diminished by the incongruent encoding and retrieval modes of the stimuli.

Now that it has been demonstrated that congruency has less of an impact on picture memory than it has on memory for words, future researchers

<table>
<thead>
<tr>
<th>TABLE 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paired-Samples $t$ Tests for Correct Identification Data</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Comparison</th>
<th>$t$</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
<th>Cohen’s $d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Picture congruent</td>
<td>-2.56</td>
<td>26</td>
<td>.100</td>
<td>0.49</td>
</tr>
<tr>
<td>Picture noncongruent</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Picture congruent</td>
<td>-3.34</td>
<td>26</td>
<td>.020*</td>
<td>0.64</td>
</tr>
<tr>
<td>Word congruent</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Picture congruent</td>
<td>3.42</td>
<td>26</td>
<td>.010*</td>
<td>0.65</td>
</tr>
<tr>
<td>Word noncongruent</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Picture noncongruent</td>
<td>-2.13</td>
<td>26</td>
<td>.260</td>
<td>0.41</td>
</tr>
<tr>
<td>Word congruent</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Picture noncongruent</td>
<td>5.54</td>
<td>26</td>
<td>.000*</td>
<td>1.07</td>
</tr>
<tr>
<td>Word noncongruent</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Word congruent</td>
<td>6.98</td>
<td>26</td>
<td>.000*</td>
<td>111.34</td>
</tr>
</tbody>
</table>

Note: * Indicates significance adjusted for using the Bonferroni method.

**FIGURE 3**
Examples of Degraded Images

Note: These images were used in the test phase of the experiment. They were altered according to the recognition-by-components theory. The images on the left are the pictures before they were degraded. The images on the right are the same pictures that were degraded by removing the concavities and the nonaccidental properties.
should conduct follow-up studies to investigate the possibility of creating a true picture superiority effect with an implicit memory test. Creating this effect would produce sufficient evidence to support a new theory of memory that supersedes the transfer-appropriate processing model of memory. Creating a better theory could lead to a greater understanding of human memory.

References


Author Note: This article is dedicated to the memory of Dr. Ami Barile-Spears. She was a passionate educator and an extraordinary mentor. This article could not have been completed without her support and guidance.
Thomaeas Brundage and Ami L. Barile-Spears, Department of Psychology, Mercer University, GA.
Correspondence concerning this article should be addressed to Thomaeas Brundage, Department of Psychology, Mercer University, Macon, GA 31207. E-mail: thomaeas@gmail.com
The Interaction of Semantic Relatedness and Emotion: Why Emotion May Not Lead to Better Memory
Christopher Hunt, Janet Trammel*, and Elizabeth J. Krumrei-Mancuso*
Pepperdine University

ABSTRACT. The present study examined what effect the combination of semantic relatedness and emotion had on the free recall of words. We hypothesized that emotion-induced priority-binding mechanisms (Mackay et al., 2004) could impair relational processing of gist traces (Brainerd & Reyna, 2002), thus leading to reduced recall of a semantically related emotional list compared to a semantically related neutral list. Seventy-two undergraduate participants viewed and recalled four 20-item pure word lists: a semantically related neutral list, a semantically related emotional list, an unrelated neutral list, and an unrelated emotional list. An Analysis of Variance revealed a significant interaction between emotion and semantic relatedness, $F(1, 71) = 6.75, p = .01, \eta^2 = .087$, such that emotion impaired recall in semantically related but not unrelated lists. There was also a main effect of semantic relatedness, $F(1, 71) = 263.56, p < .001, \eta^2 = .79$, such that semantically related lists were recalled better than unrelated lists, and a main effect of emotion, $F(1, 71) = 7.49, p = .008, \eta^2 = .09$, such that neutral lists were overall recalled better than emotional lists. These findings indicated that emotion may impair memory for overall meaning, which may be relevant to fields like education and advertising.

Research on word recall has systematically shown effects of organization (i.e., semantic relatedness) and emotion of words on their recall (Colombel, 2000; Matthews & Waring, 1972). Although the effects of these two factors have been demonstrated individually, little research has examined the interaction of these characteristics on recall, and the research that has examined them has produced inconsistent results (Buchanan, Etzel, Adolphs, & Tranel, 2006; Palmer & Dodson, 2009). Determining what effect the interaction of these two factors has could provide greater insight into how to make information more memorable, which would be relevant to fields like education and advertising where success is often measured in terms of memory for the message. For example, many educational and advertising messages employ the use of emotionally arousing anecdotes because of the commonly held belief that emotion always makes events more memorable. However, if emotion actually impairs organizational processing, then embedding the message with emotion actually detracts from the memorability of its overall meaning. For this reason, research on how emotion affects the memorability of organized material is crucial to understanding how to make educational and advertising messages optimally memorable.

The Independent Effects of Emotion and Semantic Relatedness
The superior recall of semantically related words compared to unrelated words has been demonstrated consistently in the general population (Matthews & Waring, 1972), as well as among deaf (Liben, Nowell, & Posnansky, 1978), intellectually disabled (Nolan, Cottle, & Walker, 1985), and older (Kahana & Wingfield, 2000) populations. Brainerd and Reyna’s (2002) fuzzy trace theory accounts for the superior recall of semantically related words compared to unrelated words on the basis that gist traces, which are the concepts behind the words that involve processing the word relations and patterns,
are shared by words in semantically related word lists. For example, the words dog, cat, and mouse would have a stronger gist than the words dog, sink, and shoe because the former grouping is related to an overarching concept (i.e., mammals). When items on a list are connected to an overarching concept, the brain can better process the overall context or gist of the list, which can aid subsequent recall of individual items.

Another fairly consistent finding in memory research has been that emotional valence leads to greater word recall because many studies have found words with emotional valence to be recalled better than neutral words (Colombel, 2000; LaBar & Phelps, 1998; Rubin & Friendly, 1986), regardless of whether the valence is positive or negative (Siddiqui & Unsworth, 2011). Although a variety of reasons have been postulated for why emotional words are remembered better than neutral words, one promising theory is MacKay et al.’s (2004) binding theory and corollary priority-binding hypothesis. According to binding theory, reactions to emotional words facilitate the binding of two or more elements associated with that word such as the word meaning and the word’s contextual information (e.g., where it occurs in a list). Furthermore, this binding serves as a retrieval cue for later recall (MacKay et al., 2004). Additionally, the corollary priority-binding hypothesis states that the emotional reactions caused by the presentation of emotional words causes impairment of memory of neutral words before and after the emotional word because the emotional word engages a direct amygdala-hippocampus reaction that takes precedence over the encoding of nonemotional information (Hadley & MacKay, 2006).

The Combination of Emotion and Semantic Relatedness

The few studies that have looked at the interaction of emotion and semantic relatedness have produced mixed results. First, Buchanan et al. (2006) found that semantic relatedness and physiological arousal were additive such that a list of taboo words that were both semantically related and emotionally arousing were recalled better than a list of scholastic words, which were only semantically related. This contrasted directly with the results of Palmer and Dodson (2009), who found a subtractive effect when emotion and semantic relatedness were combined; lists of semantically related emotional words were recalled significantly worse than semantically related neutral words. Furthermore, Talmi and Moscovitch (2004) found mixed results in four experiments comparing the recall of semantically related neutral words and semantically related emotional words. The lists were recalled equally in Experiments 1 and 2, but better recall was found for the semantically related neutral list than the semantically related emotional list in Experiments 3 and 4.

To reconcile the inconsistent results of these studies, it is important to theorize what mechanisms are involved when words that are both semantically related and emotional are presented. Specifically, we suggested that a synthesis of fuzzy trace theory and binding theory was needed to properly explain the combination of semantic relatedness and emotion in memory. In addition to gist traces, fuzzy trace theory also states that words have a verbatim trace, which is the word’s superficial form. Furthermore, verbatim traces are an important link to combining fuzzy trace and binding theories, given that verbatim traces include precisely the sort of information that becomes bound to word meaning through emotion according to binding theory. However, the binding of verbatim traces to emotional content does not occur without a cost. Namely, this cost is the impaired processing of information that is not facilitated by emotional binding. In a mixed list (i.e., both emotional and neutral valence) of unrelated words, this impaired information is the neutral words (Hadley & MacKay, 2006). In a pure list of semantically related emotional words, this impaired information should be the gist relations between words, given that this type of information is not one found to be facilitated by emotional binding. Such a mechanism would impair the recall of semantically related information.

Although, to our knowledge, this particular synthesis of fuzzy trace and binding theories has not previously been suggested, the idea of gist trace or context impairment due to emotionality has been supported in past literature. For instance, gist impairment was cited by Palmer and Dodson (2009) as an explanation for the decrease in recall they found in their semantically related emotional list compared to their organized neutral list because they reasoned that the emotionality of the words in the semantically related lists caused impaired processing of the gist, or relation between items in the list that would have aided in recall. In addition, Kensinger and Schacter (2006) demonstrated that, although arousing items had an item memory advantage over neutral words (i.e., a memory of the
word itself), they did not have the same advantage when it came to source memory (i.e., a memory for context). Also, Mather (2007) suggested an object-based framework whereby emotional arousal increased binding for within-object features such as color and detail (i.e., verbatim trace information), but either had no effect or impaired binding for between-object relations (i.e., gist trace information). In this way, a synthesis of fuzzy trace and binding theories supports the idea that emotion would impair recall in a semantically related list.

**Time-Frame Establishment**

Although the synthesis of fuzzy trace and binding theories predicts a subtractive effect of semantic relatedness and emotional valence in memory, this effect is extremely sensitive to both presentation length and recall delay. For the first factor, past research has found that priority-binding mechanisms only affect recall given a presentation length of less than 250 ms, after which time there are sufficient binding resources for all items to be processed without impairment (Hadley & MacKay, 2006). This relatively short threshold is due to the ease of processing unrelated words because only verbatim trace and the individual word meaning need to be processed. In contrast, semantic related words require relational processing, which has been found to be a lengthy process. For example, Folarin (1976) found that the memory advantage for semantic-related words over unrelated words continued to increase between 1,000 to 6,000 ms, showing that relational processing continued to take place during that time. Thus, given their longer processing time compared to unrelated items, priority-binding mechanisms may still be applicable to semantically related items long past the 250 ms threshold established for unrelated items.

This explanation is consistent with the fact that Palmer and Dodson (2009) and Talmi and Moscovitch (2004) found a subtractive effect of semantic relatedness and emotion, despite the latter’s presentation length being a full second longer. However, in Buchanan et al. (2006), whose presentation length was 4000 ms, there was likely enough time for complete or close to complete relational processing. This may explain why Buchanan et al. (2006) found an additive effect for emotion and semantic relatedness because there was enough time to process all gist connections between words sufficiently. However, it should be noted that Buchanan et al. (2006) also used mixed lists of words (e.g., semantic-related neutral were placed next to semantically related emotional). This means priority-binding mechanisms might have still been applicable, but the semantically related emotional words were simply draining resources from other lists, and not from the relational pattern of their own list, thus leading to their superior recall.

Although presentation length differences can help explain the differing results of past studies, it is still possible that a second factor, recall delay (i.e., the amount of time between presentation and recall), may also aply explain the differences. Recall delay may affect the recall of semantically related items due to the finding that memory gradually becomes more reliant on gist traces. Thus, any relational processing impairment that took place during presentation would be expected to become more pronounced with later time intervals because reliance on the impaired process would increase. In Palmer and Dodson (2009), recall occurred immediately after presentation, so this was likely not a factor, and the result can be interpreted through the presentation-length sensitive mechanisms previously discussed. However, in Talmi and Moscovitch (2004), recall occurred between 40 to 55 min after presentation, and thus it is difficult to verify if the longer presentation, the recall delay, or both are responsible for the subtractive effect of emotion and semantic relatedness.

In the present study, we compared the recall of semantically related neutral and semantically related emotional items given a relatively long presentation length of 1000 ms and immediate recall. This combination allowed us to determine whether the impairment of relational processing by the priority-binding mechanisms of emotion (i.e., fuzzy trace and binding theory synthesis) can be observed immediately after the items enter memory. We predicted that such mechanisms would impair relational processing given that 1000 ms, although enough time for individual items to be processed without impairment, is not enough time for gist connections between related words to occur without impairment. This impairment of relational processing by emotion priority-binding was expected to be seen through a significant interaction between semantic relatedness and emotion whereby the ability of semantic relatedness to increase recall would be reduced when paired with emotion. Furthermore, we believed that this explanation would be verified by the presence of more semantically related intrusions (i.e., words closely related to the semantic pattern that were
Semantic Relatedness and Emotion

Two factors.

Moreover, we also predicted that both semantically related word lists would be better recalled than both unrelated word lists and that the unrelated emotional word list would be recalled better than the unrelated neutral word list, given the previously established independent benefit for both of these factors in memory. These findings would provide legitimacy to the theory that a synthesis of fuzzy trace theory and binding theory is apt in explaining the interaction of emotion and semantic relatedness in memory, and would be capable of explaining inconsistencies in past literature that studied the combination of these two factors.

Method

Participants

An a priori power analysis determined that 112 participants would be required to achieve the desired .05 alpha level, .95 power level, and 0.14 effect size. However, only 72 undergraduate students were successfully recruited to participate in the present study. Participants were recruited using SONA systems, a website that allows students to sign up for research studies to receive research credit as part of a requirement in an introductory psychology course. The sample was composed of 50 women and 22 men (Mage = 18.87, SD = 1.18). Most participants (50.0%) identified themselves as European American, and 26.4% identified themselves as Asian.

Design

The study utilized a 2 (emotional valence: emotional vs. neutral) x 2 (semantic relatedness: related vs. unrelated) within subjects experimental design. There were four conditions: a neutral unrelated word list, a neutral semantically related word list, an emotional unrelated word list, and an emotional semantically related word list. Order effects were controlled between the word lists by randomizing the list order for each participant. To prevent performance fatigue, only lists of negative emotional valence were included because an inclusion of positive words would require both two additional word lists and two additional Raven’s Matrices exercises (see Materials). Furthermore, the exclusion of positive emotional lists was validated by the fact that the memory benefit for emotional valence was not found to be significantly different in past studies (Palmer & Dodson, 2009).

Materials

Stimuli consisted of four 20-item words lists (see Appendix). The neutral semantically related word list and the emotional semantically related word list was modified from Palmer and Dodson (2009) by including category names as items and shortening the list to 10 items each. The emotional unrelated word list was composed of words selected from pilot data from Palmer and Dodson (2009), and the neutral unrelated word list was composed of words selected from the University of South Florida Free Association Norms website (Nelson, McEvoy, & Schreiber, 1998). In selecting the words for the negative and neutral unrelated word lists, each word was chosen to ensure that it was not related to any other words in its list or the other lists. All words in these lists had a relatedness value of less than 0.05 (i.e., one word selected as the first associate for another word by less than 5% of participants) when paired with any other word in the study. Additionally, both semantically related lists did not differ from each other on relatedness, t(38) = 0.88, p = .38, d = 0.29.

Procedure

The present study was approved by the university’s institutional review board. Participants gave informed consent and subsequently provided demographic information. They then viewed a practice word list followed by a practice recall portion to familiarize them with the study procedure. Participants then viewed the 20 words in each list one at a time on a computer screen for 1000 ms each. After all words from a list were presented, participants typed as many words as they could remember with no time limit. After each recall portion, participants completed Raven’s Matrices for 5 min as a distraction task before beginning the next list (Raven, Raven, & Court, 1998). Words within each list and list order were randomized for each participant. Participants were debriefed after completing the fourth list.

Results

An Analysis of Variance (ANOVA) revealed that the interaction of semantic relatedness and emotion was significant, F(1, 71) = 6.75, p = .01, η² = .087, such that emotion impaired recall of the semantically related list, but not the unrelated list (see Figure 1). The main effect of semantic relatedness on recall was also significant, F(1, 71) = 263.56,
\( p < .001, \eta^2 = .79 \), such that semantically related lists \((M = 10.84, SD = 0.26)\) were recalled better than unrelated lists \((M = 7.21, SD = 0.21)\). Surprisingly, the main effect of emotion was significant in the opposite direction as predicted (see Figure 2), \( F(1, 71) = 7.49, p = .008, \eta^2 = .09 \), because neutral lists \((M = 9.33, SD = 0.24)\) were recalled better than emotional lists \((M = 8.73, SD = 0.24)\).

Pairwise comparisons showed that the semantically related neutral list \((M = 11.43, SD = 2.71)\) was recalled better than the semantically related emotional list \((M = 10.25, SD = 2.72, p = .001)\), although both semantically related lists were recalled significantly better than both the unrelated emotional list \((M = 7.21, SD = 2.15, ps < .001)\) and the unrelated neutral list \((M = 7.22, SD = 2.05, ps < .001)\). There was no significant difference in recall between the unrelated neutral list and the unrelated emotional list.

In addition to recall, data on semantically related intrusions was analyzed to help determine if the reason for the difference in recall was due to the salience of the respective semantic patterns. Because semantically related intrusions are semantically related words that were not presented, they can reveal how well the participants recognized the semantic pattern. A paired-samples \( t \) test revealed that participants gave more semantically related intrusions (recall for the semantically related neutral list, \( M = 0.29, SD = 0.46 \)) than the semantically related emotional list \((M = 0.11, SD = 0.32)\), \( t(71) = 2.84, p = .06, \eta^2 = .10 \) (see Figure 2).

To investigate the possibility that the difference in recall between lists was due to printed frequency, an ANOVA and Tukey HSD test was conducted. The ANOVA was significant such that lists differed on printed frequency, \( F(3, 76) = 3.16, p = .03, \eta^2 = .11 \), but a Tukey HSD test revealed that the only significant difference was between the semantically related neutral list \((M = 48.09, SD = 66.75)\) and the semantically related emotional list \((M = 14.09, SD = 21.01, p = .026)\). A regression analysis found that printed frequency was not a significant predictor of recall within each list \((ps > .05)\) except in the semantically related emotional list \((p = .035)\) where more frequent items were recalled better.

Similarly, an ANOVA indicated that the lists differed significantly on word length, \( F(3, 76) = 4.07, p = .01, \eta^2 = .14 \), but the only significant difference was between the semantically related emotional \((M = 6.60, SD = 2.06)\) and the semantically related neutral lists \((M = 5.05, SD = 1.28, p = .013)\). Furthermore, word length was not found to be a significant predictor of recall within any of the individual lists \((ps > .05)\).

### Discussion

The purpose of the present study was to examine how organization (e.g., semantic relatedness) and emotion of words interacted to affect recall, and to clarify inconsistencies in previous research. The results indicated that emotion had a detrimental effect on the recall of semantically related, but not unrelated words. This finding was consistent with the prediction that gist trace relational processing would be impaired by emotion. Furthermore, the finding that the semantically related neutral list had a greater amount of semantically related intrusions than the semantically related emotional list also supported our prediction because it indicated that the semantically related emotional list’s semantic...
Semantic Relatedness and Emotion  |  Hunt, Trammel, and Krumrei-Mancuso

pattern was not as recognizable, leading to reduced recall of related but nonpresented items. These results were consistent with the findings of Palmer and Dodson (2009) and Experiments 3 and 4 of Talmi and Moscovitch (2004), who also found a subtractive effect of emotion and semantic relatedness when combined.

These results also supported the prediction that gist relational processing would be sensitive to priority-binding mechanisms given a 1,000 ms/word presentation length, and that impairment by these mechanisms can be observed immediately after recall. Although Hadley and MacKay (2006) suggested that priority-binding mechanisms only affected recall at a presentation rate of 250 ms/word or faster, our results indicated that gist trace relations were still sensitive to the priority-binding mechanisms of emotion at a 1,000 ms/word presentation rate, which was likely due to their relatively lengthy processing time (Folarin, 1976). Results also indicated that gist relational processing impairment was apparent immediately after recall, which is important given that Talmi and Moscovitch (2004) had a 40 to 55 min interval between learning and recall, during which time greater impaired gist trace reliance might have led to the reduced recall of their respective words.

Although the results supported the predictions made for this experiment, one could argue that the higher printed frequency of the semantically related neutral list compared to the semantically related emotional list was the reason for the former’s superior recall. However, individual regressions between printed frequency and recall found that only within the semantically related emotional list did printed frequency significantly predict recall. In this list, words with the highest printed frequency and recall values (i.e., *kill*, *sad*, and *murder*) were also the most highly related words in the list, which greatly increased their chance of being recalled. Specifically, *kill* and *sad* were by far the most common first associates for every single word in their respective lists, and *murder* was the second most common associate for eight out of the nine words in the *kill* list. In this way, the most highly related words within the semantically related emotional list also happened to be the words with the highest printed frequency values, perhaps causing printed frequency to seem like it was affecting recall when, in reality, relatedness was the cause. Furthermore, this explanation was consistent with the fact that printed frequency was not a significant predictor of recall in any of the other word lists, which also did not have relatedness outliers.

However, we emphasize that, although this explanation rests on the fact that relatedness significantly affected recall of individual words within lists, it cannot explain recall differences between semantically related lists, given that the lists did not differ on relatedness. Rather, we hold that list type (i.e., whether semantically related, emotional, both, or neither) affected recall for the reasons outlined by combination of fuzzy trace and binding theories. In further support, Palmer and Dodson’s (2009) lists were equated for word length and frequency, with similar findings. Thus, it is unlikely that word frequency or length alone explains the impairment of semantically related emotional lists.

A surprising finding of this study was that the neutral and emotional unrelated word lists did not differ in recall. Although this contradicted prior research (Colombel, 2000; LaBar & Phelps, 1998; Rubin & Friendly, 1986), a possible explanation for this comes from binding theory, which states that emotional words only gain an advantage in recall over neutral words by being prioritized ahead of them (Hadley & MacKay, 2006). If this is true, emotion’s priority-binding would only lead to its superior recall when lists are mixed and not pure. This is because, in emotional lists, the emotional words cannot gain the priority over the other emotional words the same as they can over neutral words, thus eliminating their superior recall in mixed lists. Still, although binding theory can offer an explanation for no difference in recall between the emotional and neutral unrelated lists, the fact that past studies have found a superiority in recall for emotional word lists means this topic requires further research.

One major limitation of the present study was the lack of sufficient time and resources to recruit enough participants to test our predictions with varied presentation length and recall delays. Although our results indicated that a synthesis of fuzzy trace and binding theories adequately explained the subtractive effect of emotion and semantic relatedness, these variables must be directly manipulated in the same experiment for verification. Specifically, one would expect longer presentation lengths to lead to greater recall of semantically related emotional items compared to semantically related neutral items because there would be gradually more time for relational processing to be completed. Contrastingly, one would expect longer recall delay to further reduce the recall of semantically related emotional items compared to semantically related
neutral items because the reliance on impaired gist traces would increase as time goes on (Brainerd & Reyna, 2002). Future experiments should test these predictions through manipulating these variables, ideally both at the same time to determine if an interaction between them is present (i.e., effect of recall delay depends upon presentation length).

Our findings indicated that information may be most memorable when it is neutral and organized (e.g., semantically related), and not when it is emotional and organized, given that emotionality can impair the relational processing of information that makes organization memorable. In fields where information must be made as memorable as possible (e.g., education and advertising), this finding advised caution when it comes to incorporating emotional stimuli into a message. Although making an overall message more emotional may likely make it more memorable, making individual features of the message (e.g., words in an organized list) emotional may cloud the overall meaning of a message, and thus detract from its overall memorability. Given these findings, teachers and advertisers alike should use caution when employing the use of emotion in their messages. Specifically, they should utilize emotion only when the message is relatively simple and short, when they only wish for their audience to remember a specific detail, and when organizational processing of the message’s various parts is required.

References

Buchanan, T. W., Etzel, J. A., Adolphs, R., & Tranel, D. (2006). The influence of emotion in their messages. Specifically, they finding advised caution when it comes to incorporating emotional stimuli into a message. Although making an overall message more emotional may likely make it more memorable, making individual features of the message (e.g., words in an organized list) emotional may cloud the overall meaning of a message, and thus detract from its overall memorability. Given these findings, teachers and advertisers alike should use caution when employing the use of emotion in their messages. Specifically, they should utilize emotion only when the message is relatively simple and short, when they only wish for their audience to remember a specific detail, and when organizational processing of the message’s various parts is required.

APPENDIX

<table>
<thead>
<tr>
<th>Word Lists and List Type With Category Words in Bold</th>
</tr>
</thead>
<tbody>
<tr>
<td>Categorical/Negative</td>
</tr>
<tr>
<td>Kill</td>
</tr>
<tr>
<td>Murder</td>
</tr>
<tr>
<td>Homicide</td>
</tr>
<tr>
<td>Melancholy</td>
</tr>
<tr>
<td>Upset</td>
</tr>
<tr>
<td>Sleep</td>
</tr>
<tr>
<td>Tired</td>
</tr>
<tr>
<td>Relax</td>
</tr>
<tr>
<td>Rocking</td>
</tr>
<tr>
<td>Sofa</td>
</tr>
<tr>
<td>Regret</td>
</tr>
<tr>
<td>Apple</td>
</tr>
<tr>
<td>Cloud</td>
</tr>
<tr>
<td>Newsprint</td>
</tr>
<tr>
<td>Square</td>
</tr>
</tbody>
</table>
Substance abuse among college students is a widespread concern for colleges and universities across the United States (Palmer, McMahon, Moreggi, Rounsaville, & Ball, 2012). Drug use on campuses can pose significant health, emotional, and behavioral risks (Johnston, O’Malley, Bachman, & Schulenberg, 2011). Students attending college undergo a crucial developmental time period characterized by significant interpersonal and professional transitions that influence their adult development. With college student drinking on the rise, it is imperative to be aware of the significant risk factors and predictors associated with substance use. Furthermore, the current understanding of the risk factors related to college student substance use will assist in establishing effective on-campus prevention and treatment programs (Brook, Morojele, Pahl, & Brook, 2006).

According to the five-factor model of personality (Digman, 1990), the full range of personality traits can be well-defined in terms of five basic dimensions. These dimensions are extraversion, agreeableness, conscientiousness, neuroticism, and openness (Goldberg, 1993; Luo, Kranzler, Zuo, Wang, & Gelernter, 2007). Personality traits are one of the many factors that have been implicated in contributing to the development of drug use and abuse. Premorbid personality traits such as impulsivity, thrill-seeking, rebelliousness, irresponsibility, and nonconformity appear to play a central role in the development of substance dependence (Sher, Walitzer, Wood, & Brent, 1991). In contrast, other personality traits, especially negative emotionality (anxiety, inhibition, moodiness, and unhappiness), may be a consequence rather than a cause of substance dependence (Schuckit, Irwin, & Brown, 1990).

Beginning in early childhood, personality trait differences can be seen in those children who abuse substances versus those who do not (Anderson, Tapert, Moadab, Crowley, & Brown, 2007). Those children who abused drugs at a young age had significantly higher scores on neuroticism, and lower scores of agreeableness and conscientiousness. These traits appear to be consistent across the lifespan. Personality traits related to neuroticism and disinhibition have been consistently associated with substance use disorders (Grekin, Sher, & Wood, 2006). Because of its wide applicability and consistency across the lifespan, knowledge of the personality traits that render a person susceptible to...
substance abuse can prove beneficial in the prevention and early intervention of drug using behavior.

Personality dimensions may be related to both a propensity to engage in substance abuse and to use different kinds of substances (Hopwood, Baker, & Morey, 2008). As such, personality differences among subtypes of drug abusers are a significant and important issue to explore. Although limited, previous research has illustrated that key personality differences may exist between those individuals who abuse alcohol and those who abuse narcotics (McGue, Slutske, & Iacono, 1999). Specifically, Butler (2004) reported that participants who identified alcohol as their primary drug of choice scored significantly higher on neuroticism, agreeableness, and conscientiousness than participants who identified cocaine as their drug of choice. Studying the individual differences among substance abusers may serve as a tool to enhance clinical assessment in substance abusing populations.

In addition to personality traits, stress serves as a risk factor for substance use (Gurley & Satcher, 2003). There is increasing evidence that environmental and economic stressors can have adverse effects on families and children, indicating that stress is an important risk factor in substance use (Brook et al., 2006). Moreover, the literature has consistently found stress to be a significant risk factor for the development of addiction to drugs and/or alcohol and relapse vulnerability (Bamberger & Bacharach, 2006; Sinha, 2008). Furthermore, stressful life events have been found to be associated positively with depression, poor physical health, and substance use, suggesting that stress may be a precipitant of drug use behavior (Unger, Kipke, & Simon, 1998). Lloyd and Turner (2008) investigated the relationship between stress and alcoholism, and lent support to this hypothesis. The researchers found that lifetime stress exposure exhibits a pattern of association with alcohol dependence.

Personality traits and stressful life events serve as potential risk factors for substance use and abuse in college students. The Big Five neuroticism factor has been studied by stress researchers and has been found to be related to many poor health outcomes (Williams & Wiebe, 2000). Neurotic individuals perceive stressors as more stressful, which may underlie some of the health risks (Guenole, Chernyshenko, Stark, McGregor, & Ganesh, 2008). Substance abuse is a form of destructive behavior that people may engage in when experiencing high levels of stress. Identifying personal traits along with improved treatment methods may enable people to more constructively manage their stress.

The present study aimed to construct a profile predictive of drug use among college students. First, we hypothesized that significant personality differences would exist between users of alcohol and marijuana (minor drug use) and users of narcotics (major drug use). Specifically, we hypothesized that major drug users would score higher on neuroticism and lower on conscientiousness. Second, we hypothesized that increased levels of stress would be positively associated with total drug use. Last, we hypothesized that the interaction between stress, high levels of neuroticism, and low levels of conscientiousness would be the greatest predictor of major drug use. Multiple regression analyses and Multivariate Analyses of Variance (MANOVA) were used to investigate these relationships.

**Method**

**Participants**
The participants were 202 undergraduate and graduate university students between the ages of 18 and 24. The sample consisted of 110 (54%) women and 92 (46%) men. Most of the sample was European American, (n = 148), with other ethnicities represented as follows: African American (14%; n = 29), Latino (10%; n = 21), and other (2%; n = 4). Participants reported a mean income of approximately $45,000 (range $9,000–$150,000). Being college students, most of the population were single (94%) or cohabitating but not married. Most participants had some college education (85%), with the rest either high school grads or some graduate school. Participants were not paid or compensated for their time or participation. However, dependent upon the professor, some students might have received extra credit in class for their participation in research. There were no exclusion criteria for participating in the study. The study was approved by the university’s institutional review board.

**Design and Procedure**
The experimenter visited classrooms and residence halls, and asked students to complete packets of self-report measures. Most participants were recruited from the residence halls on campus. Participation was voluntary, and participants were able to withdraw at any time for any reason without penalty. Participants were given two to three days to complete the packets and return the forms to the
Personality, Social Stress, and Drug Use

researcher. The confidentiality of participant data was emphasized, with signed informed consents being obtained prior to the participants’ completion of the packet.

Materials

Demographics. Background data such as age, sex, marital status, household income, ethnicity, and education level were collected using a self-report demographic questionnaire.

Personality. Personality dimensions were assessed in all participants using the 60-question, self-report NEO Five-Factor Inventory (NEO-FFI; Costa & McCrae, 1992; McCrae & Costa, 1997). The NEO-FFI measures the five major domains of personality: openness, conscientiousness, extraversion, agreeableness, and neuroticism. The items are scored on a five-point Likert-type scale, ranging from 0 (strongly disagree) to 4 (strongly agree). Through comparative intercultural structures, evidence so far has suggested that the five-factor structure is very similar across cultures (McCrae & Costa, 1997; McCrae, Costa, Del Pilar, Rolland, & Parker, 1998). The replicability and ubiquity of the Big Five have led many personality psychologists to advocate this structure as a basic framework for personality description and assessment across cultures (Zukauskiene & Barkauskene, 2006). The scale demonstrated good internal consistency and validity across various diverse samples (Bjornsdottr et al., 2014).

Stress. The College Life Stress Inventory (Renner & Mackin, 1998) was used as the measure to investigate stress. The scale is a 50-question self-report form that asks about stressors and daily hassles that are most relevant to college students. The participants checked off each of stressors that they experienced within the past year. The items are scored on an ordinal scale with each stressor assigned to a number value between 20 and 100, with 100 being the most severe stressor. The total number values are added together, and a higher score on the scale indicates a higher level of stress.

Drug Use. The Drug Use Screening Inventory (Tarter, 1990) was used as a quantitative self-report instrument to measure frequency and type of substances used. Questions were asked regarding the type of drugs used and the frequency of use of the drugs in the last year. Used most commonly as a qualitative measure, the Drug Use Screening Inventory was converted into a quantitative measure for statistical purposes. Each of the 11 drug categories (e.g., alcohol, cocaine, ecstasy) were number coded. The five categories for frequency of the drug use were assigned a number to reflect the amount of drug usage. To quantitatively report the total amount of drug use, the frequency of the use of each drug used was added together, with a higher score indicating a greater amount of drug use. Participants who reported use of alcohol and/or marijuana were coded as “minor” drug users, and participants who reported use of other illicit drugs and narcotics were coded as “major” drug users.

Results

Group Personality Differences

We hypothesized that there would be significant differences in neuroticism and conscientiousness between major and minor drug users. Looking at all five of the personality dimensions, a one way MANOVA was conducted, Wilks’ Lambda = .68, F(5, 193) = 17.83, p < .001. There were significant personality differences among the two groups of drug users. Major drug users scored significantly higher on neuroticism, F(1, 197) = 54.33, p < .001, η² = .21, and lower on conscientiousness, F(1, 197) = 35.15, p < .001, η² = .15, than minor drug users.

Predictors of Drug Use

Standard multiple regressions were conducted to determine the accuracy of the independent variable social stress in predicting total drug use. Regression results indicated that the overall model significantly predicted total drug use, R² = .29, adjusted R² = .29, F(1, 195) = 80.59, p < .001. This model accounted for 29.2% of the variance in total drug use. Regression analyses also revealed that the interaction of stress and neuroticism significantly predicted total drug use, R² = .46, adjusted R² = .45, F(3, 190) = 54.53, p < .001. This model accounted for 46.3% of the variance in total drug use. We examined the model of stress, neuroticism, and extraversion predicting total drug use, and found that it was also significant, R² = .46, F(3, 190) = 53.21, p < .001.

Discussion

The study’s findings strengthened previous research supporting the importance of personality and stress as crucial factors in drug use and abuse. As hypothesized, we did observe differences between types of drug users. Users of illicit street drugs and narcotics (major drug users) displayed higher levels of neuroticism and lower levels of conscientiousness. Consistent with previous literature, stress served as a risk factor for general substance
use, regardless of drug use of choice (Sinha, 2008).

Results of the present study supported the growing research examining the typology variance among different drug users (Ersche, Turton, Pradhan, Bullmore, & Robbins, 2010). The findings indicated that those individuals who experience large amounts of stress and possess neurotic personality traits may be more susceptible to abusing substances. Additionally, findings suggested significant differences in personality traits between major and minor drug users, which may have broad implications for future research and treatment methods.

Limitations of the study suggested that the findings be analyzed with caution. The correlational design of the study limited the ability to draw causal relationships. In addition, we examined only two important risk factors for substance use. Although research has shown their potential significance, several other factors can play a role in the development of substance use. Reviews of research on the development of drug use have reported that drug use of peers and friends is a major risk factor for drug use (Belcher & Shinitzky, 1998; Copans & Kinney, 1996; Hawkins, Catalano, & Miller, 1992). Other risk factors for substance use that were unaccounted for included trauma history, psychopathology, social support, and biological factors (Stone, Becker, Huber, & Catalano, 2012; Swendsen et al., 2010).

The findings for the group differences should be carefully examined. The Drug Use Screening Inventory has traditionally been used as a qualitative measure of the frequency and type of drug use. For statistical purposes, there needed to be a measure of total drug use. The qualitative data was formatted quantitatively for the purposes of the present study. The issue lies in the reliability of the measure in a quantitative form. Further, the grouping of participants into major and minor drug use categories may create an artificial distinction, because research has illustrated high rates of polysubstance use in this population (Connor, Gullo, White, & Kelly, 2014). Also, although we used fairly standard measures, we were unable to perform reliability statistics for the study sample. Although we have no reason to believe that the measures used were not reliable in the sample studied, we cannot verify this as being the case.

The nonrandom sampling of participants posed a generalizability issue for the study. Although we had a fairly large sample of participants, the nonrandomization recruitment of participants weakened the significance of the results. The homogeneity of the sample also lessened the generalizability of the findings. With most of the sample ethnically European American and attending the same college, perhaps the results would prove different with a more heterogeneous sample. In addition, in an effort to maintain privacy and confidentiality, our survey methodology limited our ability to be certain that the participants who received the packets were the same as the person who completed the forms.

The vast majority of etiology research has concentrated on testing main effects of models of drug use. A far smaller number of studies have examined interactions between predictors of substance use (Brook, Whiteman, Balka, Win, & Gursen, 1997; Brook, Whiteman, Gordon, & Cohen, 1986, 1989; Cooper, Peirce, & Tidwell, 1995; Curran, White, & Hansell, 1997). Universities across the United States strive to provide their students with tools to make healthy and constructive decisions. Many colleges also provide educational programs, counseling, and substance abuse services to their students. To better assist their students, it is important for colleges and universities to be aware of the risk factors and precipitants of substance use. Knowledge of these concepts will hopefully better inform prevention programs and treatment methods. Future research is needed to determine the precise nature of these relationships as well as to improve treatment approaches.

References


COPYRIGHT 2015 BY PSI CHI, THE INTERNATIONAL HONOR SOCIETY IN PSYCHOLOGY (VOL. 20, NO. 1/ISSN 2164-8204)


Online Journal Submission Process

All Psi Chi undergraduates, graduates, and faculty* are invited to submit their research to the *Psi Chi Journal of Psychological Research* through the new web based manuscript submission, tracking, and peer review software solution. Better than email submissions used in the past, this software allows users to create personal accounts to make the submission process more efficient.

Four items are required for all submissions:

1. **Cover Letter**  
   Include primary author's education status, manuscript originality statement, IRB approval

2. **Sponsoring Statement**  
   Undergraduate first authors only

3. **Cover Page**  
   Author names, school affiliation, and any author note

4. **Masked Manuscript**  
   MS Word with all personal information removed

Simply register an account, then click Submit Manuscript:

1. **Upload Files**  
   Files can be removed, replaced, or reorganized

2. **Enter Manuscript Information**  
   E.g., title, abstract, authors, keywords, etc.

3. **Review Manuscript Material**  
   Summary of all information/files submitted

4. **Submit Manuscript**  
   Receive email confirmation

New software benefits:

- Allows users to track their manuscripts' progress
- Inserts multiple files including cover letters, manuscripts, and figures
- Permits users to prioritize files and coauthors
- Checks for mistakes in the submission process and points out any errors
- Streamlines the process for authors and reviewers

* Psi Chi member ID number required

Register an account: http://pcj.msubmit.net/cgi-bin/main.plex
Tutorial videos: http://www.ejpress.com/demos