ABOUT PSI CHI
Psi Chi is the International Honor Society in Psychology, founded in 1929. Its mission: “recognizing and promoting excellence in the science and application of psychology.” (Note. Our new mission statement is available at http://www.psichi.org/?page=purpose) Membership is open to undergraduates, graduate students, faculty, and alumni 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 1,130 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 membership provides two major opportunities. The first of these is academic recognition to all 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®, EBSCO®, and Crossref® databases. In 2016, the Journal also became open access (i.e., free online to all readers and authors) to broaden the dissemination of research across the psychological science community.

JOURNAL INFORMATION
The Psi Chi Journal of Psychological Research (ISSN 2325-7342) is published quarterly in one volume per year by Psi Chi, Inc., 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; psichijournal@psichi.org.

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.

Advertisements that appear in Psi Chi journal do not represent endorsement by Psi Chi of the advertiser or the product. Psi Chi does not endorse or is responsible for the contents of third-party promotions. Learn about advertising with Psi Chi at http://www.psichi.org/?page=Advertise

Permission must be obtained from Psi Chi to reprint or adapt a table or figure; to reprint quotations exceeding the limits of fair use from one source, and/or to reprint any portion of poetry, prose, or song lyrics. All persons wishing to utilize any of the above materials must write to the publisher to request nonexclusive world rights in all languages to use copyrighted material in the present article and in future print and nonprint editions. All persons wishing to utilize any of the above materials are responsible for obtaining proper permission from copyright owners and are liable for any and all licensing fees required. All persons wishing to utilize any of the above materials must include copies of all permissions and credit lines with the article submission.
154 Conquering APA Style: Advice From APA Style Experts
Jennifer L. Hughes, Agnes Scott College; Debi Brannan, Western Oregon University; Bradley Cannon, Psi Chi Central Office; Abigail A. Camden, Agnes Scott College; and Amber M. Anthenien, University of Houston

163 The Workplace Interpersonal Conflict Scale: An Alternative in Conflict Assessment
Robert R. Wright, Brigham Young University – Idaho; Ashley E. Nixon, Willamette University; Zachary B. Peterson, Brigham Young University – Idaho; Sharon V. Thompson and Ryan Olson, Oregon Health & Science University; Scott Martin and Devon Marrott, Brigham Young University – Idaho

181 Personality Traits, Early Maladaptive Schemas, and Severity of Nonsuicidal Self-Injury
Sarah D. Arthurs and Josephine C. H. Tan
Lakehead University

193 Not Just Right Experiences in Obsessive-Compulsive and Related Disorders
Jeremy P. Schmidt and Sara M. Stasik-O’Brien
Knox College

206 Childhood Abuse and Adult Interpersonal Functioning: Analysis of Self-Defeating Beliefs and Binge Drinking as Mediators
Jacob E. Allen, Georgia State University
Daniel T. Rogers and Sharon M. Pearcey, Kennesaw State University

220 The Effects of Mood on Writing Apprehension, Writing Self-Efficacy, and Writing Performance
Jared Vanhille, Bradley Gregory*, and Grant Corser
Southern Utah University

231 The Interaction Effect of Facial and Vocal Attraction on Overall Perceived Attractiveness
Sheila Bonnough and Erin Moore
Stetson University

242 Rape Acknowledgment Status and Recency Since Rape as Correlates of College Women’s Body Shame
Carolyn Paige Merwin and Suzanne L. Osman
Salisbury University

*Faculty mentor
Conquering APA Style: Advice From APA Style Experts

Jennifer L. Hughes, Agnes Scott College; Debi Brannan, Western Oregon University; Bradley Cannon, Psi Chi Central Office; Abigail A. Camden, Agnes Scott College; Amber M. Anthenien, University of Houston

ABSTRACT. Learning and teaching APA style is often a challenge because of the detailed rules for the writing style. Resources that provide accurate information about APA style that are easy for learners to consume can be useful tools for students and instructors. The goal of this article is to provide information to help writers become more competent and comfortable with APA style. The first section of this article contains lists of common APA style mistakes that are made when papers are submitted for publication. Editors look for very specific APA formatting and style, and this section will cover these issues. Moreover, the second section includes a list of APA style rules frequently encountered by an APA style tutor. The third section has information to help students navigate writing assignments in a research methods class. The fourth section addresses additional APA style rules that many writers do not know about. Finally, this information can serve as a guide for writers to use when writing APA style papers.

American Psychological Association (APA) style was developed to be a rigorous standard for scientific communication (APA, 2010). These guidelines help authors to structure their papers, write more clearly and concisely, know the mechanics of style, display their results, credit their sources, and provide references (APA, 2010). Learning and teaching APA style can be a challenge because writers are not often interested in learning the intricate details of the writing style or they are intimidated by all the details involved in learning APA style. However, an accurate knowledge of APA style is an important tool for writers, instructors teaching APA style, and reviewers (McDonald, 2011).

We have written this article with the hope of helping writers learn more about APA style. In an attempt to engage students while teaching APA style, instructors have used checklists (Franz & Spitzer, 2006), templates (Franz & Spitzer, 2006; Stahl, 1987), games (Hughes, 2017), sample manuscripts riddled with APA mistakes (Smith & Eggleston, 2001), peer review of APA style in papers (Mandernach, Zafonte, & Taylor, 2016), and online tutorials (Mages, & Garson, 2010). Franz and Spitzer (2006) found that students do better with learning APA style if the material is presented in many formats. Mandernach et al. (2016) argued that direct instruction, including requiring repeated practice and teaching meaning behind the theory of APA style, as well as a combination of the other methods listed above, is the optimal way to learn APA style. We hope this article will serve as an additional resource for those learning and teaching APA style.

This article covers common APA style mistakes. We present four lists of issues, which are seen by the Psi Chi Journal of Psychological Research Managing Editor, an APA Tutor, the Editor of the journal, a graduate student who teaches APA style, and an Associate Editor for the journal. Mandernach et al. (2016) demonstrated that writers using APA style had the most difficulty with documentation (i.e., citations, quoting, and references). These are repeated themes in the advice we give, and we aim to highlight those areas as critically important for learners and teachers alike.

Part 1: Common APA Style Mistakes When Submitting a Paper for Publication

By Bradley Cannon, Writer/Journal Managing Editor, Psi Chi Journal of Psychological Research
Mastering APA Style can seem like a lot to take in, especially for students, many of whom were taught only MLA format in their prior writing courses, if any format at all. Short of reading the APA Publication Manual from cover to cover, what rules should one learn first? Having conducted more than 700 APA Style reviews in my role as the Psi Chi Journal’s APA Style reviewer, I would like to encourage authors to start by considering the following 10 mistakes. Based on my experience, these mistakes occur often, and if they are not caught early, they also tend to be more challenging and time-consuming to correct later. Although many students learning APA format are not yet writing papers for publication, it is a good practice to treat all writing projects as if they might eventually be published. In the classroom, mastering APA Style will make (a) future classes less tedious, (b) material easier to consume, (c) increase confidence and self-efficacy, and (d) improve overall grades. Learning proper APA Style before (or at least early in) the review process will also help (a) prevent tedious revisions for the author, (b) increase reviewers’ comprehension of the article, and/or (c) minimize publication delays. The good news is this: These mistakes are easy to avoid with only a little awareness and practice.

1. Missing References and Other Materials
Before submitting a manuscript, it is crucial to make sure that all materials are openly available to reviewers that would be helpful in analyzing the article. This includes references and matching citations, tables and figures, participant demographics (i.e., age, sex and/or gender, race/ethnicity, and any other relevant demographics), and other data and facts necessary for a reader to understand the study, its validity, and the implications. For instance, authors should manually scroll through their article to make sure that all citations have a corresponding reference in the references section, and vice versa. As another example, authors should double-check that they included all tables and figures that are mentioned in the manuscript body, as well as a copy of materials such as the brand-new scale that they created. Editors and reviewers cannot assess and correct information that is not available to them, so any of these missing items are almost certain to cause a manuscript to go through an additional round of reviews.

2. Incorrect Verb Tense
As a general rule, authors should use the past tense (e.g., had, have) throughout an empirical research manuscript. For example, “For our first hypothesis, we expected participants would . . .” “Bandura (1976) found . . .” or “Researchers have found . . .” Only use the present tense to discuss the implications of the results and to present conclusions (e.g., “Based on our results, people may be more likely to . . .”; APA, 2010, section 3.06). Inappropriate use of the present tense can mislead readers about who did what in a sentence, or lead readers to believe that a hypothesis or finding is more understood and thoroughly tested than it really is. Learn more about verb tense in APA (2010) section 3.18.

3. No Permission to Reprint/Adapt
The repercussions of committing plagiarism can be extremely costly, and yet authors sometimes overlook this. Remember: It is the responsibility of the author to obtain written permission from the original publisher to reprint or adapt anything previously published that is beyond fair use (APA, 2010, section 6.10). This includes an image, for example, taken from Google as well as a previously published table. (Yes, even if you were the original author of that table.) Obtaining permission from the publisher can sometimes take time, so this process should be started as early as possible, preferably before conducting the research. Failing to receive permission to reprint materials could force an author to make significant revisions to an article to ensure that the material is sufficiently described without use of the material itself. To determine whether permission is needed and what steps should be taken, see APA (2010) sections 6.10 and 8.04 on “Permission to Reprint” and “Obtaining Written Permission” from the copyright owner. APA (2010) section 2.12 shows how to properly cite that permission has been granted.

4. Interchangeable Use of the Terms Sex and Gender
Gender identity can change over time (Westbrook & Saperstein, 2015) and is defined “as a person’s deeply-felt, inherent sense of being a boy, a man, or male; a girl, a woman, or female; or an alternative gender (e.g., genderqueer, gender nonconforming, boygirl, ladyboi)” (APA, Divisions 16 and 44, 2015, p. 20). On the other hand, sex refers to the biological sex (i.e., male, female). In other words, if participants were unable to select options other than man and woman, then the term sex should be used, not gender. Inconsistent or inaccurate use of these terms
APA Style Editorial | Hughes, Brannan, Cannon, Camden, and Anthenien

...can mislead reviewers and is sometimes challenging to correct without the author’s assistance. See APA (2010) section 3.11 to learn more, or Hughes, Camden, and Yangchen (2016) for specific examples of how to word demographic questions about sex, gender, and many other demographic questions.

5. Missing or Noncongruent Races/Ethnicities

Many people do not realize that race and ethnicity mean different things to different people. Race tends to refer to skin color or physical appearance (e.g., Black, White), whereas ethnicity tends to refer to location (e.g., African American, European American). Authors are encouraged to use congruent language throughout (e.g., Blacks and Whites or African Americans and European Americans) as opposed to noncongruent language (e.g., Blacks and European Americans). Also, it is best to refer to people by who they are (e.g., Blacks) than who they are not (e.g., non-Whites). Again, see Hughes, Camden, and Yangchen (2016) and APA (2010) section 3.14 to learn more.

6. Capitalizing Scales, Subscales, Variables, and Groups

Incorrect or inconsistent capitalization can also sometimes make it difficult for readers to decipher an author’s meaning. For example, Fearful the subscale or fearful the emotion are not necessarily the same thing. Scales (e.g., Sensation-Seeking Scale) and subscales (e.g., Sensation-Seeking Disinhibition scale) should always be capitalized as shown, whereas variables (e.g., happiness) and conditions and groups (e.g., information and no-information conditions) should be lowercased. An exception to this rule occurs if variables appear with multiplication signs (e.g., Happiness × Hunger). See APA (2010) sections 4.18–4.20 for more examples.

7. Writing in the Passive Voice

Writing in the passive voice can be awkward to read and generally wastes space. Moreover, many authors do not know that each page often costs some journals money to print and/or lay out for publication. Therefore, authors should watch out for and generally replace passive verbs and phrases (e.g., “was/were” and “There was/were”). For example, change “There were three participants who we eliminated from further analyses” to “We eliminated three participants from further analyses.” See APA (2010) section 3.18 to learn more.

8. Not Including Digital Object Identifiers (DOIs)

Most major academic journals now use DOI numbers, which function as hyperlinks to specific articles, issues, or publications. For example, https://doi.org/10.24839/2325-7342.JN22.3.154 is the DOI for this invited editorial. As opposed to URLs, DOIs are permanent and can be updated by the publisher so that the same DOI hyperlink will lead to the latest version of an article, even if the article’s URL changes. DOI numbers should be placed at the end of all journal references in clickable hyperlink form as shown above, not by listing the doi as doi:10.24839/2325-7342.JN22.3.154 (McAdoo, 2017). DOI numbers can be found online for most articles or by using the “Search Metadata” tool at https://doi.crossref.org/guestquery. DOIs were first introduced in 2000, but many journals have created DOIs for their past articles going back far before that year. If a journal article does not have an assigned DOI number and the article was found online, state “Retrieved from [insert URL]” instead (APA, 2010, section 7.01).

9. Misuse of Abbreviations

APA (2010) states that abbreviations are occasionally helpful for communicating complex scientific terms. However, the manual also explains that abbreviations should generally be avoided if the terms may be unfamiliar to readers (APA, 2010, p. 106, section 4.22). Furthermore, APA (2010) section 4.22 requires authors to avoid abbreviations that are used fewer than three times, and authors should consider avoiding abbreviations that occur more often than three times, too (Hales, Kipling, & Rector, 2017). If an abbreviation is deemed to be useful, it should be written out on the first use and introduced in parentheses; all further uses of that term in the manuscript should be written in abbreviation form only. As an exception, a few abbreviations are accepted as established word entries such as IQ and AIDS, and do not need to be introduced; to see if an abbreviation is an established term, authors can check Merriam-Webster’s Collegiate Dictionary online (APA, 2010, section 4.24).

10. Unmasked Manuscripts

When a journal requests a masked manuscript submission, this means that all identifying information such as author and school affiliation should be removed from the file. Doing so protects author confidentiality, which is why the Psi Chi Journal and many other journals require all submissions to...
be submitted in this way. Identifying information should be carefully removed from the first page and also from all other pages throughout the manuscript (e.g., data was collected at CSU). Students often fail to fully mask their manuscripts, which causes their work to be returned to them to correct this before the peer-review process can begin. Authors should also review submission guidelines for any other specific instructions for the title page or elsewhere throughout the submission; journals do sometimes have specialized requirements outside of the APA Publication Manual in order to best fit each journal’s unique needs.

Part 2: Common Writing Issues That Tutors Address

By Abigail A. Camden, APA Style Tutor

Even if a paper has strong writing, errors in APA Style can detract not only from its clarity, but also its physical appearance and perceived quality. Therefore, tutoring students to craft psychology papers under the basic parameters of APA Style is a necessity. The following are 10 APA Style errors that I often see as an APA Style tutor:

1. Paraphrased Material Must Be Cited

The most critical APA error that I see in many student papers is not citing paraphrased material. Section 6.01 of the APA (2010) Publication Manual notes that this delves into the dangers of plagiarism: "Whether paraphrasing, quoting an author directly, or describing an idea that influenced your work, you must credit that source" (p. 170). Additionally, though page numbers or paragraph numbers are not required for paraphrased material, section 6.04 (p. 171) of the manual recommends providing them, particularly for material derived from text that is intricate or extensive.

2. Formatting of Running Heads

A running head should be a shortened version of a paper’s title, with the latter conveying essential information about the paper’s main topic (APA, 2010, section 2.01). As an APA Style tutor, common running head errors are: incorrect typeface or font size, repeating the words “Running head” on pages after the first, capitalizing “head” on the first page, using more than 50 characters, and not capitalizing the running head (i.e., correct: “Running head: YOUR TITLE HERE,” APA, 2010, section 8.03). Because of automatic formatting of headers in word-processing systems, authors must check the APA Style of their running heads (particularly the typeface and font size). For an example of a correct running head, see the first two pages of the sample paper on p. 41 of the APA (2010) Publication Manual.

3. Incorrect Headings and Levels of Heading

Students often make errors when adding headings to their manuscripts (e.g., incorrect levels of heading, bolding headings that should not be). The basic headings within an APA Style manuscript are: Abstract, Method (singular!), Results, Discussion, and References, all of which are centered, and all of which are bolded except Abstract and References (APA, 2010). Authors should not include a heading for the introduction, unless requested by the journal (section 3.03). In tutoring, I find that students are often hesitant to add additional subheadings to their manuscripts; however, subheadings (particularly within the literature review) add “signposts,” clarity, and structure to writing. This said, APA's levels of heading must be followed, including having at least two subsections with a given section in order for a subheading to be warranted (much like an outline; APA, 2010; see p. 62 for a useful table showing subheading formatting).

4. Formatting of the Text

Incorrect APA formatting of the text can reduce a manuscript’s readability and professionalism (APA, 2010, p. 228). When formatting a manuscript, it is recommended in section 8.03 that authors use serif fonts—with Times New Roman being the preferred font—except in tables and figures, in which sans-serif fonts are acceptable. Double spacing should also be used throughout, margins should be at least 1 in., and paragraphs should be indented at ½ inch (APA, 2010, section 8.03). Two spaces should be used following the period at the end of a sentence (APA, 2010, section 4.01). (Note: Psi Chi Journal of Psychological Research prefers one space when submitting for publication.) However, many students whom I tutor find it unnatural to use two spaces. As a remedy, after completing a manuscript, authors can use their word processor’s “find and replace function” (Ctrl + F) to search for one space (“") that should be replaced with two spaces (“ “). After using this technique, authors must edit the spaces after initials (e.g., in references and the byline) and in-text punctuation (such as e.g., i.e.) to be one space (APA, 2010, section 4.01). For more comments on formatting, see “Preparing the Manuscript for Submission” in section 8.03 (pp. 228–230) of the Publication Manual (APA, 2010).
5. Improper Use of “et al.”
Students are sometimes challenged by the proper use of “et al.” The general rules for the use of et al. within in-text citations are as follows: for in-text citations of three to five authors, writers should initially list the names of all authors, and subsequently use et al.; for six or more authors, et al. should be used immediately (APA, 2010, p. 175, 177; also see p. 177 for a comprehensive chart). However, when shortening a citation with the same first author and the same year, using et al. can result in separate works appearing identical. In this circumstance, one should list enough authors prior to et al. to differentiate them (APA, 2010, section 6.12). Authors should ensure that they do not use the full in-text citation (i.e., without et al.) more than once; this can be easily accomplished using Ctrl + F to search for an author’s last name throughout the document. In addition to its use in parenthetical citations, et al. should also be used within a sentence instead of “and colleagues” (APA, 2010, section 4.26).

6. Incorrect References
Rather than following samples provided in section 7.01 (APA, 2010), some authors are tempted to use online quick-fixes for creating references, or copy and paste references from PsycINFO or Google Scholar without checking the accuracy of the APA Style. This can result in incorrect references. Frequent errors I see in tutoring are imprecise use of words such as since (pp. 83–84) or like (p. 68), problems with word choice or colloquialisms (p. 68), and a lack of transition words (p. 65). Namely, APA (2010) notes that—to increase the clarity of scientific writing—since is best used as an indicator of time (i.e., rather than meaning because). Similarly, like should not be used to mean such as but rather for conveying analogies or similarities. Colloquialisms

7. Italicize Most Statistical Coefficients
Another common APA error is not italicizing statistical coefficients when reporting results. In short, most statistical symbols used for reporting results of statistical analyses are italicized (e.g., $N$, $df$), whereas Greek letters (e.g., $\alpha$), subscripts, and superscripts are not (APA, 2010; see section 4.45 for an extensive chart). Relatedly, when mentioned within the sentence rather than within commas, statistical symbols should be written in nonitalicized words instead of symbols (i.e., “mean,” not $M$).

8. Effective Abstracts
Students often wrangle with crafting effective abstracts, with errors such as not including all aspects of the paper; writing too much or too little (i.e., it should be 150 to 250 words), or not including keywords. APA (2010) notes that the abstract can be viewed as the most important paragraph in the manuscript. It should include a terse overview of all aspects of the research including the purpose, methods, results (including significance levels, confidence intervals, or effect sizes), and the study’s applications and implications. As a tutor, one useful technique I use to help students create abstracts is the “structured abstract” (see the APA Style Blog; Stefanie, 2014) to ensure that all elements of the abstract are present. After writing a structured abstract, authors can then remove the headings if they are not required for the journal or professor to whom the paper is being submitted. Ultimately, abstracts are tools that provide readers a brief glimpse of one’s research. Thus, also using carefully chosen keywords at the end of an abstract can further enhance its effectiveness.

9. Common Errors With Parentheses
In APA Style, sets of parentheses should never be beside each other, but should instead be combined with a semicolon (e.g., parentheses offering an example and parentheses with an in-text citation; APA, 2010, section 4.09, p. 94). Students are also often challenged by multiple citations within one set of parentheses. In this instance, citations should be listed in alphabetical order, separated by semicolons (APA, 2010, section 6.16). Further, work by the same authors (in the same order) should be cited by listing the name(s) followed by the years of publication in ascending order, separated by commas. For more detail, see APA, 2010, section 6.16.

10. Word Choice and Flow of Writing
One challenge for all writers is increasing the clarity and effectiveness of their written communication. APA’s (2010) sections 3.07 through 3.11 offer guidance on writing style. Some common errors that I see in tutoring are imprecise use of words such as since (pp. 83–84) or like (p. 68), problems with word choice or colloquialisms (p. 68), and a lack of transition words (p. 65). Namely, APA (2010) notes that—to increase the clarity of scientific writing—since is best used as an indicator of time (i.e., rather than meaning because). Similarly, like should not be used to mean such as but rather for conveying analogies or similarities. Colloquialisms
and ambiguities should also be avoided (e.g., a lot; see section 3.09). To further increase flow and precision in writing, APA (2010) suggests the use of transition words to link thoughts and provide continuity. Online sources for transition words (e.g., Campbell, Buckhoff, & Dowell, 1997) that organize the words by type of transition (e.g., concession, similarity) can be useful in helping students learn to add transitions to their writing. For more direction on improving the clarity and precision of one’s writing, see sections 3.07 through 3.11 (APA, 2010).

A major barrier to students mastering APA Style is not obtaining or using the APA (2010) Publication Manual. This is a loss to learning because the manual provides irreplaceable resources that explain all aspects of APA Style, in addition to providing useful content such as sample papers. However, if students or authors are not able obtain the manual, the APA Style Blog (http://blog-apa-style.org/) constitutes a comprehensive and reliable resource maintained by APA itself. The blog is wonderful—filled with examples of references, headings, phrasing (you name it!). Many of the exact resources from the APA manual are available on this website including various explanatory charts and sample papers.

Part 3: Writing Tips and Tricks to Help in a Research Methods Class

By Debi Brannan, Editor, Psi Chi Journal of Psychological Research, and Amber Anthenien, Graduate Research Assistant

When students enter a research methods class, often it is their first experience with attempting to read, write, and understand scientific language and materials. Mastering these skills is critical during an introductory research methods course because, as students move to higher-level psychological courses, professors often assume that the students can communicate and think like a researcher/professional. We often see the following APA writing errors in our method courses. The following include tips for writing a better APA style paper.

1. Use of “As Cited In . . .”
Authors should cite material using “as cited in” only when they are unable to obtain the original articles (APA, 2010, section 6.17). This should be done sparingly. Authors should check whether their articles are available using Google Scholar or PsycINFO, and cite them directly. It is possible that once the authors read the article and full results they may interpret the study findings or contributions differently than other researchers. Note: If authors use “as cited in,” only the material they read is listed in the reference section (APA, 2010, section 6.17). See page 178 for an example of how to cite secondary sources in text.

2. Direct Quotations
Authors should use direct quotations sparingly, and they should only use them when they are crucial to making their point (Cash, 2009). When authors use quotations, they must cite the source and list a page number or numbers (APA, 2010, section 6.03). Quotes with fewer than 40 words should appear in the text, but quotes that are 40 words or longer should be displayed in a freestanding block about a half inch from the left margin. The quote should be indented, double-spaced, and appear without quotation marks. Additionally, in-text citations for block quotations should be placed outside the period, as opposed to within the period typically used in writing.

3. Use of Contractions
Another common mistake is for authors to use contractions. In scientific writing, authors should spell words out completely (Lee, 2015).

4. Semantics Matter
When writing in APA format, authors should not write proven. Instead, they should use words such as revealed, indicated, or suggested. As much as researchers want to prove something to society, scientists can never say with 100% accuracy that they are correct because there will always be situations in psychological research that can change the outcome of the investigation (e.g., age, culture). Moreover, all psychological research requires replication. It is also important to remember that researchers never have complete control over all the confounding and extraneous variables threatening the validity of their study. One other common mistake is to discuss “the study” as if it were a person (APA, 2010, section 3.09). For example, “The study examined the influence of mood on music” should be changed to “The researchers examined the influence of mood on music.”

5. Citing of Dissertations and Theses
Authors should avoid citing dissertations or theses. If they are good enough for publication, they will have a corresponding article that can be cited. The reason to use published papers is because they go through a rigorous peer-review process that provides some assurance in the researcher’s findings.
6. Citing Measures
If authors use a measure or measures, they need to be sure to cite them (APA, 2010, section 7.08). Researchers worked hard to create the measures; they deserve the credit. Moreover, others reading the paper may want to consider the measures used, and proper citations will allow them to find the measures. Authors should also give evidence for the reliability and validity for the measures to establish that they used strong measures (APA, 2010, section 2.06).

7. Paragraph Length
One sentence paragraphs come across as abrupt for readers, and paragraphs that are lengthy (i.e., longer than one double-spaced page) can lose the reader’s attention (APA, 2010, section 3.08). The first sentence of each paragraph should give the reader a brief overview of the information that will be covered in that paragraph. The body of the paragraph will provide supporting evidence for that statement. And the last sentence of the paragraph should relate that information back to the hypothesis or set up the following paragraph.

8. Seriation Throughout the Paper
APA (2010) states that, to aid the reader, authors should use seriation. Authors should make sure that items within a series are “syntactically and conceptually parallel” in papers (APA, 2010, p. 63, section 3.04). One trick that many authors use is to alphabetize the items they are using in their papers and always present those items in that order. Remember that one or two sentences do not make up a paragraph.

9. Use of Which and That
Many authors do not know when to use which and that correctly. Authors should use which for information that is parenthetical in nature, and they should set off the material with a comma (APA, 2010, section 3.22). For information that is essential to the meaning of the sentence, authors should use that. For example, “Floors that do not have carpet are cold to stand on” as compared to “I stood on the cold floor, which did not have carpet.”

10. Use of Numbers
We often see authors use numbers incorrectly. The general rule is that authors should use numerals for all numbers 10 and above (APA, 2010, section 4.31). However, using numbers can be more complicated than that general rule. See Figure 1 for additional guidelines.

Part 4: APA Style Rules Many Writers Do Not Know
By Jennifer L. Hughes, Associate Editor, Psi Chi Journal of Psychological Research
I frequently see the following APA rules violated when I review papers for possible publication. When I talk with writers about these errors, they often tell me that they have never heard about these rules. The following list includes my top 10 errors that many writers do not know about. I also added an explanation for how to avoid committing each error. Readers can test themselves to see how many of these rules they know about.

1. Emphasizing Terms
APA (2010) advises authors to use quotation marks for words or phrases that are slang, ironic comments, or invented or coined expressions (section 4.07). Authors should do so the first time the phrase is used, but not after that. In contrast, authors should italicize technical or key terms for emphasis the first time they are used (section 4.07).

2. Presentation of Lists of a Series of Items
APA (2010) dictates how authors should present a series of items (section 3.04). The manual states that authors using a series within a sentence or paragraph should use letters, not numbers, in parentheses such as: (a) first, (b) second, and (c) third (section 3.04). They should not use: (1), (2), (3); 1, 2, 3; or a), b), c). They should use commas to separate three or more items that do not have internal commas. If there are three or more elements with internal commas, they should use semicolons.

---

Figure 1

| Authors should use numerals for numbers that represent “time, dates, ages, scores and points on a scale, exact sums of money, and numerals as numerals” (APA, 2010, p. 112, section 4.31). As mentioned earlier in this paper, authors should use numerals for statistics, but they also should use numbers for percentages, decimals, ratios, and percentiles/quartiles (APA, 2010, section 4.31). When authors are using the plurals of numbers they should add s or es alone, without an apostrophe (APA, 2010, section 4.38). A common mistake involves the use of decades and authors should write them like this: 1970s. Authors should use commas between groups of three or more digits (APA, 2010, section 4.37). An example would be 7,300. |
If authors are using separate paragraphs in a series, they should identify those by using Arabic numerals followed by a period. These numerals should not be enclosed in or followed by parentheses. Finally, if authors are using separate sentences in a series, they also should use Arabic numerals and follow them with periods. The first word should be capitalized and the sentence should end with correct punctuation. However, it should be noted that the use of “numbered lists” may signal an unwanted ordinal position among the items. If authors do not wish this to happen, they can use bullets instead.

3. Spacing of Numbers in Text
Authors should treat numbers as they would treat words (APA, 2010, see section 4.46). For example, spaces should go before and after numbers in the text, so that they are easier to read (e.g., not 1+2 but 1 + 2).

4. When to Use a Zero Before Decimals
Authors typically do not know when to use a zero before decimal points. In section 4.35, APA (2010) states authors should “Use a zero before the decimal point with numbers that are less than 1 when the statistic can exceed 1” (p. 113). A common example of this is Cohen’s d and it would be listed as 0.60. APA (2010) also states that authors should not “. . . use a zero before a decimal fraction when the statistic cannot be greater than 1 (e.g., correlations, proportions, and levels of statistical significance)” (p. 113).

5. Punctuation and Quotation Marks
Authors should close quotation marks after periods and commas (APA, 2010, section 4.08). For example, “. . . like this.” Other punctuation marks are placed within quotations only when they are part of the quoted material.

6. Using En Dashes Between Page Numbers in References
Authors often use the terms hyphen and dash interchangeably (McAdoo, 2010). However, dashes are different from hyphens. APA (2010) specifies two types of dashes in section 4.13: em dashes (i.e., longer than a hyphen and used to set off an element) and en dashes (i.e., longer and thinner than a hyphen, but shorter than an em dash; they are used between items of equal weight in a compound adjective, page ranges, and other types of ranges). Most authors do not know to use an en dash between the page range for books and articles listed in their reference sections. McAdoo (2010) describes how to create an en dash: “On a PC, hold the Control key and type the minus sign (specifically, the one on the numeric keypad to the right; this shortcut will not work with the one at the top of the keyboard)” and “On a Mac, hold the Option key and type the minus sign (specifically, the one on the top of the keyboard)” (para. 7). Here is an example hyphen (-), em dash (—), and en dash (–).

7. Periods Do Not Follow DOIs or URLs in References
Authors typically do not know when to use a zero before decimal points. In section 4.35, APA (2010) states authors should “Use a zero before the decimal point with numbers that are less than 1 when the statistic can exceed 1” (p. 113). A common example of this is Cohen’s d and it would be listed as 0.60. APA (2010) also states that authors should not “. . . use a zero before a decimal fraction when the statistic cannot be greater than 1 (e.g., correlations, proportions, and levels of statistical significance)” (p. 113).

8. Article Titles That End With an Exclamation Point or Question Mark
When writing references, for article titles that already contain punctuation such as an exclamation point or question mark, the author should keep the original punctuation and not add a period (Lee, 2011). The author should not have two punctuation marks after the title.

9. Listing Editors of Books in References
Writers typically know how to list authors’ and editors’ last names and initials in references. An example would be: Hughes, J. L., & Camden, A. A. (APA, 2010, section 6.27). However, for editors, initials appear before surnames (e.g., J. L. Hughes & A. A. Camden). A common mistake is that writers put a comma between two editors’ names. There should not be comma between the two editors’ names, but a comma is used if the book has three or more editors. Note: As listed above, the initials of the authors and editors should have a space between them, and many writers violate this rule (APA 2010, section 4.01). A good guideline to use is that all elements of a reference should have one space between them (APA, 2010, section 6.25).

10. Listing Publishers in References
Authors should list publishers in as brief a form as possible (APA, 2010, section 6.30). For example, authors should remove the following: The, Publishers, Co., and Inc. However, the following should be kept: Books and Press.

In conclusion, we believe it is important for academics as writers, instructors, and reviewers to continue to learn about APA style. Students also need a solid foundation for APA style. This article
presented common APA mistakes with the hope that both academics and students will learn from the errors.

References
Hughes, J. L. (2017, March). Teaching APA style by making it into a game. In J. L. Hughes, Keeping students’ interest: Class activities for undergraduate psychology courses; Panel presentation given at the annual meeting of the Southeastern Psychological Association, Atlanta, GA.

Author Note. Jennifer L. Hughes, Department of Psychology, Agnes Scott College and Associate Editor for the Psi Chi Journal of Psychological Research; Debi Brannan, Western Oregon University and Editor for the Psi Chi Journal of Psychological Research; Bradley Cannon, Psi Chi Central Office, Writer/Manager Editing Editor for the Psi Chi Journal of Psychological Research; Abigail A. Camden, Agnes Scott College and APA Style Tutor; Amber M. Anthenien, University of Houston and Graduate Research Assistant.
Correspondence concerning this article should be addressed to Jennifer Hughes, Department of Psychology, 141 E. College Ave., Agnes Scott College, Decatur, GA 30030. E-mail: jhughes@agensscott.edu
The Workplace Interpersonal Conflict Scale: An Alternative in Conflict Assessment

Robert R. Wright, Brigham Young University–Idaho; Ashley E. Nixon, Willamette University; Zachary B. Peterson, Brigham Young University–Idaho; Sharon V. Thompson and Ryan Olson, Oregon Health & Science University; Scott Martin and Devon Marrott, Brigham Young University–Idaho

ABSTRACT. Interpersonal conflict is a prevalent workplace problem associated with numerous adverse consequences for both employees (e.g., depression, negative affect) and organizations (e.g., turnover, reduced productivity). However, many currently available self-report measures suffer from multiple methodological challenges that substantially hinder accurate and comprehensive measurement including the lack of specific elements of conflict identified by the literature subsequent to measure development, and a lack of rigorous empirical examination of psychometric properties. Thus, there is a need for a valid, psychometrically sound scale that can briefly capture perceptions of workplace interpersonal conflict in a contemporary work environment. The Workplace Interpersonal Conflict Scale (WICS) was developed to examine the frequency of conflict characteristics identified from a prior qualitative study of interpersonal conflict themes. In the current examination, we conducted an in-depth analysis of the psychometric properties of the WICS across 3 occupational samples including a cross-sectional study in health care (home care workers), a pre/post study in service (food service workers), and a large, diverse online sample (via MTurk). Results supported the 6-item WICS as a valid measure of workplace interpersonal conflict related to many other important workplace variables in the domains of work, health, and safety. Practical implications are presented and discussed. The WICS offers a promising measure that can be used in assessment, remediation, and prevention of noxious interpersonal cultures within the workplace.

Interpersonal conflicts at work contribute to numerous detrimental health and safety outcomes that impact the well-being of employees and organizations, and are often upsetting, troublesome, and frequent work stressors (Bolger, DeLongis, Kessler, & Schilling, 1989; Williams, 2003; Wright, Mohr, Sinclair, & Yang, 2015). Interpersonal conflicts are pervasive, with employees reporting occurrences on, at minimum, 50% of work days (Hahn, 2000). Estimates suggest that more than 26% of the general worker population is negatively impacted by conflict (de Raeve, Jansen, van den Brandt, Vasse, & Kant, 2009). Furthermore, workplace interpersonal conflicts are frequently precursors to severe social problems such as workplace aggression and violence (Barling, Dupre, & Kelloway, 2009; Pearson, Andersson, & Wegner, 2001), as well as costly employee outcomes including turnover (Frone, 2000), burnout (Fujisawa, Tsukishima, Tsutsumi, Kawakami, & Kishi, 2003; Harvey, Blouin, & Stout, 2006), and depression and withdrawal (Inoue & Kawakami, 2010; Spector & Jex, 1998). Previous research has identified interpersonal conflict to be one of the largest reducible organizational costs and the single most important workplace stressor...
Workplace Interpersonal Conflict Scale | Wright, Nixon, Peterson, Thompson, Olson, Martin, and Marrott

for organizations to address (Dana, 1999; Keenan & Newton, 1985).

Numerous measures of workplace interpersonal conflict and related social stressor constructs have been developed. Although related constructs are prolific in the literature (e.g., workplace violence and aggression, workplace incivility, workplace bullying, social undermining; cf. Hirschcovis, 2011; Nixon & Spector, 2015), due to interpersonal conflict’s near ubiquitous nature within the workplace, this article focuses exclusively on the measurement of perceptions of interpersonal conflict from the perspective of the individual.

Many definitions and conceptualizations of conflict have been proposed in the literature (see Barki & Hartwick, 2004, and Weingart, Behfar, Bendersky, Todorova, & Jehn, 2015, for reviews), demonstrating some disagreement regarding what constitutes interpersonal conflict and how conflicts are expressed. As such, interpersonal conflict is distinguished from other related constructs in Barki and Hartwick’s (2004) comprehensive conceptual definition as “a dynamic process that occurs between interdependent parties as they experience negative emotional reactions to perceived disagreements and interference with the attainment of their goals” (Barki & Hartwick, 2004, p. 234). Additional clarity is provided by Weingart and colleagues’ (2015) recent assertion that conflicts involve people with opposing viewpoints advocating for different outcomes (Weingart et al., 2015), though conflict expression may vary in frequency and intensity.

Even within the confines of these definitions, measures of interpersonal conflict are considerably heterogeneous. These differences largely arise from dissimilar epistemological and methodological approaches to measure development. For instance, several measures reflect the fact that workplace interpersonal conflict can encompass a wide range of behaviors from mild disagreements between individuals to more intense behaviors such as sabotage or even verbal and physical intimidation (Spector & Jex, 1998). Moreover, each measure that is currently available within the literature has not been evaluated under the same level of scrutiny, making it difficult to identify a measure with acceptable psychometric properties.

There are three main limitations of existing measures of workplace interpersonal conflict including current conceptualization, measure development, and evaluation. These limitations are discussed below and presented in Table 1. As such, the overall purposes of the present evaluation are to

(a) articulate the need for a valid, psychometrically sound self-report scale that can briefly capture perceptions of workplace interpersonal conflict in a contemporary work environment from the perspective of the individual and (b) examine the psychometric suitability of the 7-item Workplace Interpersonal Conflict Scale (WICS) within three separate contexts.

Measurement Issue of Current Conceptualization

Many currently available self-report measures of workplace interpersonal conflict were developed many years ago (see Table 1). This is a concern because the understanding of the complexities of interpersonal conflict in the workplace has broadened the scope of what aspects should be considered part of conflict including additional behaviors, cognitions, and negative emotions. For example, Schieman and Reid (2008) argued that interpersonal conflict should include perceptions of injustice, goal impediments, incompetence, and being the target of another person’s verbal or physical antagonism. Yet, within most available workplace interpersonal conflict measures, perceptions of these cognitions, emotions, and behaviors are not included. Moreover, many early self-report measures of conflict within organizational contexts have struggled regarding appropriate conceptual measurement of cognitions, behaviors, and negative affect (Barki & Hartwick, 2004), which are inherently perceptual in nature. Whereas these measures may query conflict behaviors, it should be acknowledged that any self-report measure of conflict is inherently perceptual, where even a report on behaviors observed is according to the person observing and perceiving such behaviors. Therefore, measures created many years earlier may lack comprehensive assessment of conflict and claim to be measures of behavior when, in fact, this ignores the importance of perception in conflict measurement.

As an example, consider the most widely used and extensively psychometrically evaluated measure of workplace interpersonal conflict, the Interpersonal Conflict at Work Scale (ICAWS; Spector & Jex, 1998). Although this scale has been well-validated (Bruk-Lee & Spector, 2006; Frone, 2000; Spector & Jex, 1998), without the inclusion of all current aspects of interpersonal conflict, it lacks an appropriate and contemporary workplace outlook. Primarily, it does not specifically address all the properties of conflict (Barki & Hartwick, 2004), but rather assesses the frequency of conflict based
on perceptions of the largely behavioral indicators of arguments, yelling, actions of rudeness, and negative interpersonal interactions. Further, the ICAWS does not capture more recently identified components of conflict such as perceptions of injustice, goal impediments, or incompetence (Schieman & Reid, 2008). Consequently, this measure of conflict is likely unable to capture the construct of workplace interpersonal conflict as it is currently defined and understood.

**Measurement Issue of Development**

A second major issue with previous measures of workplace interpersonal conflict reflects the process of measure development including empirical derivation, selecting an appropriate epistemological approach, and testing among multiple populations (see Table 1). As a sound methodological practice, quantitative measures can be developed based on qualitative findings including those obtained from focus groups, content analysis, or behavioral observations, where the strengths of both approaches enrich and reinforce the other (Morgan, 2014). However, epistemological approaches and assumptions can further influence measure construction and development. For example, a sharp contrast in methodological approach has emerged within the conflict literature, where many studies have examined conflict in work groups based on the focus of the conflict or “conflict types” (e.g., relationship, task content, task process, nontask organizational, status; Behfar, Mannix, Peterson, &

<table>
<thead>
<tr>
<th>TABLE 1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Currently Existing Measures of Workplace Interpersonal Conflict in the Literature</strong></td>
</tr>
<tr>
<td>Measure Reference</td>
</tr>
<tr>
<td>Amason (1996)</td>
</tr>
<tr>
<td>Barki &amp; Hartwick (2001)</td>
</tr>
<tr>
<td>Interpersonal Conflict in Organizations Scale (Bruk-Lee, 2006)</td>
</tr>
<tr>
<td>Cox Conflict Scale Cox (1997) Friedman et al. (2000)</td>
</tr>
<tr>
<td>Dierdorff &amp; Ellington (2008)</td>
</tr>
<tr>
<td>Intragroup Conflict Scale (ICS) Jehn (1995) Jehn &amp; Chatman (2000) Jehn et al. (1997)</td>
</tr>
<tr>
<td>ICS Modification Scales Behfar et al. (2010) on two types of process conflict Bendersky &amp; Hays (2012) on status conflict</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Schieman &amp; Reid (2008)</td>
</tr>
<tr>
<td>Interpersonal Conflict at Work Scale (ICAWS) Spector &amp; Jex, 1998</td>
</tr>
</tbody>
</table>

*Note: Empirical refers to the whether the measure is based off of themes in the literature with an empirical derivation of the measure. Properties refers to whether the scale centers on properties of conflict. Sample refers to being examined within multiple appropriate occupational samples. Psychometric Evaluation refers to the report of psychometric properties beyond internal consistency or factor analysis (i.e., EFA, CFA) of the scale including discriminant, convergent, predictive, concurrent, and incremental validity.*
Workplace Interpersonal Conflict Scale

Wright, Nixon, Peterson, Thompson, Olson, Martin, and Marrott

Trochim, 2010; Bendersky & Hays, 2012; Bruk-Lee, Nixon, & Spector, 2013; Jehn, 1995; Weingart et al., 2015) and other studies have assessed conflict by the properties that define it (e.g., disagreement, interference, negative emotion; Barki & Hartwick, 2001; Cox, 1997; Friedman, Tidd, Currall, & Tsai, 2000; Schieman & Reid, 2008; Spector & Jex, 1998).

Although it has been well-documented that both of these approaches are important in conceptualizing interpersonal conflict (cf. Barki & Hartwick, 2004), it is important to note that the Intragroup Conflict Scale (ICS; Jehn, 1995), including two subsequent scales introduced by Behfar et al. (2010) on process conflict and Bendersky and Hays (2012) on status conflict, expressly considers the focus of conflict within existing work groups. Although valuable in the workplace where productivity is increasingly reliant upon successful collaborations, a measure of interpersonal conflict developed within a group setting may inadvertently confound group processes with interpersonal processes. For example, the ICS includes an item that reads, “How much friction is there among members in your work unit?” By asking about the entire work unit (group) rather than perceptions of dyadic interactions (that may or may not be within a larger group setting), the measurement of conflict may be confounded with other group or social processes. One’s perceptions may differ dramatically when there are multiple observers of conflict, or participants in conflict, as opposed to the potential anonymity perceived by an employee when engaged in a one-on-one interpersonal conflict. Thus, the ICS adopts a fundamentally sociological perspective of conflict, which has a limited ability to examine specific aspects of conflict as a psychological work phenomenon.

Finally, when a measure is not inspected within different theoretically appropriate groups, the adequacy of the content and generalizability of that measure is uncertain. Barki and Hartwick (2001) and Cox (1997) developed their measures of interpersonal conflict (see Table 1) within only one occupational context: information system organizations and nursing staff, respectively. Although their samples included participants with differing job descriptions, it was used in only one occupational setting, making it difficult to know whether the validity is acceptable and appropriate in other work contexts. Thus, a measure of workplace interpersonal conflict should be empirically derived, examine specific properties of conflict, and be examined within multiple occupational settings.

Measurement Issue of Evaluation

The third major concern with measures of interpersonal conflict is the substantial variance in the psychometric evaluation they undergo (see Table 1). Evidence for reliability and construct validity including convergent and discriminant validity, as well as criterion-related validity (e.g., concurrent, predictive, incremental), of a measure’s scores is essential for adequate measure evaluation (Crocker & Algina, 1986). Although it has become common practice to examine reliability (i.e., internal consistency) of a measure’s data (e.g., Cronbach’s alpha statistic) and, to a lesser extent, conduct exploratory and confirmatory factor analysis (e.g., exploratory and confirmatory factor analysis; EFA and CFA), few studies on workplace interpersonal conflict conduct additional psychometric evaluation to examine other forms of validity.

Indeed, of the currently available measures of workplace interpersonal conflict, only three have explicitly examined validity beyond internal consistency and factor analysis. First, regarding modified versions of the ICS (Jehn, 1995), Jehn and Chatman (2000) investigated discriminant validity of scores between the scales (task conflict, process conflict, relationship conflict), and both Behfar et al. (2010) and Bendersky and Hays (2012) examined process conflict and status conflict, respectively, relative to convergent, concurrent, and discriminant validity of the scores. Second, Spector and Jex (1998) evaluated the relationship between the ICAWS and several dispositional variables (e.g., negative affectivity, autonomy) and job strain variables (e.g., depression, job satisfaction, physical symptoms) and concluded that the observed relationships provided adequate support for convergent and discriminant validity. Third, in a dissertation on the subject, Bruk-Lee (2006) reported evidence of data reliability and concurrent, convergent, and discriminant validity of the scores for each subscale of the Interpersonal Conflict in Organizational Scale, a 63-item scale examining both the focus and properties of conflict. However, these are exceptions because even Schieman and Reid’s (2008) measure of workplace interpersonal conflict, one of the most recently developed self-report published measures, has not been examined beyond internal consistency. Given the great discrepancy in evaluation methods for available conflict measures, there is a need for a brief, comprehensive, and perceptual measure of workplace interpersonal conflict that has been rigorously appraised to provide an acceptable measure of this potent workplace stressor.
A New Measure
To address this need, following Hinkin’s (1998) general recommendations, we constructed a new scale called the Workplace Interpersonal Conflict Scale (WICS; see Appendix). In doing so, all three of the critical issues of concern for measures enumerated above were addressed because the WICS was (a) created as a self-report perceptual measure based on a current conceptualization and definition of workplace interpersonal conflict, (b) developed using rigorous empirical methods including multiple worker samples, and (c) evaluated across multiple types of validity. In the analysis of the WICS, our focus was on important organizational variables including health and safety outcomes, thus substantiating beneficial uses of this measure in future workplace studies.

Scale Development of the WICS
The foundation for the content of the WICS originated from a qualitative study (Wright, Mohr, & Sinclair, 2014) investigating common themes in 150 descriptions of conflicts experienced at work among 144 registered nurses, using several independent raters including experts in the nursing field to verify theme validity and interrater reliability. This qualitative study was selected for several reasons including a focus on conflicts with lateral sources (e.g., peers, colleagues) rather than hierarchical sources (e.g., supervisors, subordinates), the empirical qualitative methods employed (i.e., content analysis with multiple raters), and the straightforward application of the themes within the conflict conceptualization (cognitions, behaviors, affect). Five reliable themes (κ > .40; Landis & Koch, 1977) were identified (i.e., feeling unfairly treated, feelings of dislike/rudeness, others’ incompetence, work disagreement) and developed as nonmutually exclusive thematic codes, which allowed for a richer and deeper representation of each conflict description.

Using this qualitative study as a guide (Wright et al., 2014), a total of seven items were developed for the WICS with six of these items patterned after the five reliable themes. Given the intense negative nature of the feeling unfairly treated theme, two items were created to represent it and one item was created to represent each of the remaining themes. The seventh item incorporated another behavior (yelling) that was not represented in the qualitative study’s themes due to good face validity (i.e., yelling; Spector & Jex, 1998). Thus, using Barki and Hartwick’s (2004) definition of conflict for a conceptual guide and Wright et al. (2014) as a guide for specific content, we constructed the WICS to examine the perception of the frequency of important theoretical and practical instances of workplace interpersonal conflict.

Scale Evaluation of the WICS
We gathered validation evidence for the WICS from three studies conducted in unique occupational settings. In each study, construct validity was examined by estimating internal consistency (i.e., Cronbach’s alpha), interitem correlations, factor structure (EFA Principal Components and Direct Oblimin rotation for potential multiple factors, CFA) and model fit indexes among the WICS scores. A variety of fit indexes were selected due to the unique limitations of each (Hu & Bentler, 1999; Kline, 2005) including absolute fit (i.e., chi-square, Standardized Root Mean Square Residual or SRMR), relative fit (i.e., Normal Fit Index or NFI, Incremental Fit Index or IFI, Tucker-Lewis Index or TLI), and noncentrality fit (i.e., Comparative Fit Index or CFI, Root Mean Square Error of Approximation or RMSEA). In their seminal publication, Hu and Bentler (1999) provided a thorough review of each model fit index, highlighting some limitations and recommending cutoffs for determining adequate model fit. Based on these recommendations, the cutoff values of < .08 for SRMR; ≥ .95 for NFI, IFI, TLI, and CFI (though many have used .90 as the cutoff; see Marsh & Hau, 1996); and ≤ .06 for RMSEA were adopted.

To assess criterion validity (both concurrent and discriminant validity), we examined bivariate correlations between the WICS scores and theoretically related variables, with the expectation that the WICS would be positively related to negatively valenced constructs (e.g., negative affect) and negatively related to positively valenced constructs (e.g., positive affect). For these analyses, several work variables important to productivity and health were selected (e.g., work engagement, turnover intentions, counterproductive work behavior, affect, depressive symptoms, perceived stress, safety compliance). Convergent and incremental validity were assessed by comparing correlation estimates between the scores of the WICS, ICAWS, and ICS. Finally, using linear regression, some initial evidence for predictive validity was also examined.

Study 1
The sample for Study 1 came from a large, cluster randomized controlled trial called COMPASS (COMmunity of Practice and Safety Support),...
conducted to evaluate the efficacy of a peer-led, scripted health and safety curriculum intervention for home care workers (Olson et al., 2015; Olson et al., 2016). These home care workers were independent contractors who provided one-on-one in-home services for older and disabled individuals who required assistance with tasks of daily living such as housecleaning, bathing, and dressing. The primary purpose of this study was to evaluate the intervention’s effectiveness in improving and promoting home care worker health and safety.

Method
Participants. Home care workers \((n = 148)\) were recruited from two metropolitan locations in the Pacific Northwest. Eligibility criteria included employment by at least one consumer-employer who paid for services through personal funds or who were receiving home care services through the publicly funded Oregon Department of Human Services Aging and Disabilities, and residing or working in one of the two study areas to facilitate study participation. The sample was predominantly women \((92.2\%)\), with an average age of 51.7 years \((SD = 13.2; \text{Median} = 53.0)\). A third of the sample held a high school diploma \((33.6\%)\), many had a vocational or technical certificate \((24.5\%)\), and many had either some college education \((16.1\%)\) or earned a college degree \((21.7\%)\). Participants reported an average of 7.44 years tenure in home care work \((SD = 7.61; \text{Median} = 4.5)\) and a mean of 24.00 weekly work hours \((SD = 17.1; \text{Median} = 22.0)\).

Measures. Workplace interpersonal conflict was assessed using the 7-item WICS constructed for this study. Participants indicated the frequency of each of the seven items over the past month on a 5-point Likert-type frequency scale \(1 = \text{never}, 5 = \text{very often}\); see Appendix for entire scale). Affect was captured using an 8-item measure of mood combining four items from Larsen and Diener (1993) and four items from Watson, Clark, and Tellegen (1988), where participants indicated on a 5-point Likert-type scale \(1 = \text{not at all}, 5 = \text{extremely}\) how much a certain mood descriptor described their mood over the past month along positive \(\text{(i.e., happy, alert, enthusiastic, relaxed}; \alpha = .83)\) and negative \(\text{(i.e., sad, irritable, bored, nervous}; \alpha = .68)\) dimensions (Wright, Broadbent, Graves, & Gibson, 2016). Loneliness was assessed using the 3-item measure developed by Hughes, Waite, Hawkley, & Cacioppo (2008), over the past month using a 5-point Likert-type frequency scale \(1 = \text{never}, 5 = \text{very often}; \alpha = .88\).

Acute depressive symptoms over the past seven days were examined using the 5-item CES-D (Bohannon, Maljanian, & Goethe, 2003) on a 4-point Likert-type frequency scale \(1 = \text{rarely or none of the time [less than 1 day]}, 4 = \text{most or all of the time [5–7 days]}\); \(\alpha = .83\). Finally, life stress over the past three months was captured using seven items from the Perceived Stress Scale (Cohen, Kamarck & Mermelstein, 1983) on a 5-point Likert-type frequency scale \(1 = \text{never}, 5 = \text{very often}; \alpha = .84\). These included the shortened 5-item scale plus two items with good face validity for the population.

Two questions from Buxton et al. (2009) with minor adaptations assessed the frequency of sugary snacks and drinks over the past month using a 10-point scale \(1 = \text{never}, 10 = 5 \text{ or more times per day}\). The mean of these items was used to estimate the frequency of added sugar consumption. Because this construct was comprised of two items, internal consistency estimates are not meaningful, though we did verify a significant positive relationship \((r = .30, p < .001)\) between these variables.

Procedure. We obtained institutional review board approval at Oregon Health & Science University (#5473). Study recruitment and data

<table>
<thead>
<tr>
<th>Variable</th>
<th>M (SD)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Workplace Interpersonal Conflict</td>
<td>1.76 (.88)</td>
<td>.92</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Positive Affect</td>
<td>3.46 (.82)</td>
<td>- .29*</td>
<td>.83</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Negative Affect</td>
<td>2.18 (.79)</td>
<td>.29</td>
<td>- .48*</td>
<td>.68</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Depressive Symptoms</td>
<td>0.75 (.70)</td>
<td>.23</td>
<td>- .51*</td>
<td>.70*</td>
<td>.83</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Loneliness</td>
<td>2.49 (1.13)</td>
<td>.25</td>
<td>- .60*</td>
<td>.66*</td>
<td>.54*</td>
<td>.88</td>
<td></td>
</tr>
<tr>
<td>6. Perceived Stress</td>
<td>1.42 (.73)</td>
<td>.27</td>
<td>- .67*</td>
<td>.59*</td>
<td>.55*</td>
<td>.61*</td>
<td>.84*</td>
</tr>
<tr>
<td>9. Unhealthy Sugar Diet</td>
<td>3.09 (1.37)</td>
<td>.04</td>
<td>- .21*</td>
<td>.13</td>
<td>.19*</td>
<td>.19*</td>
<td>.15</td>
</tr>
</tbody>
</table>

Note: Internal consistency (Cronbach’s alpha) estimates, where appropriate, are provided on the diagonal. * \(p < .05\); ** \(p < .01\).
collection procedures for the COMPASS trial have been described at length previously (Olson et al., 2015). In the current analyses, we focused on a limited range of survey measures collected from home care workers at the baseline time point. Surveys were collected from participants in person by researchers and inspected for completeness to minimize missing data.

Results

Internal consistency and construct validity. Descriptive results including means, standard deviations, and correlations for the WICS and other study variables are displayed in Table 2. No demographic variables were statistically related to the WICS scores. It was observed that, when Item 7 (“been given unclear directions by others at work”) was not included, every measure of internal consistency and construct validity improved or remained unchanged. There was also a theoretical concern about this item, receiving unclear directions, as being more related to role ambiguity than conflict. Given these concerns, we proceeded to analyze the 6-item WICS, which is consistent with Hinkin’s (1998) recommendation to retain four to six items in a scale. Interitem correlations ranged from 0.49 to 0.78 ($M = 0.66$, $SD = 0.08$), corrected item-total correlation was above the recommended cutoff of .40 (Gliem & Gliem, 2003), and Cronbach’s alpha was .92, suggesting acceptable internal consistency.

Using Principal Component Analysis and Direct Oblimin rotation, the EFA yielded one factor (no rotation resulted) because only one Eigenvalue was above the convention of 1.0 (4.30 with the next highest at .60), and visual inspection of the scree plot supported this conclusion. Each item strongly loaded on the factor, ranging from 0.80 to 0.91 ($M = 0.85$, $SD = 0.04$). The CFA produced similar results with standardized item loadings ranging from 0.74 to 0.90 ($M = 0.81$, $SD = 0.06$). Although model fit was not acceptable using chi-square analysis, $\chi^2(9) = 60.50$, $p < .05$, the SRMR indicated acceptable model fit at .048 and relative fit index values approached acceptable model fit ($\text{NFI} = 0.91$, $\text{IFI} = 0.92$, $\text{TLI} = 0.81$). Although the TLI was below the 0.95 recommendation and the RMSEA was .20, which was well above the cutoff of .06, these indexes are sensitive to small sample size and can over-reject models that other indexes would indicate are otherwise acceptable (Hu & Bentler, 1998; Kline, 1999). Finally, the CFI was 0.92, suggesting the model was nearing adequate fit. Taken together, these estimates suggest some evidence for construct validity of the 6-item WICS scores.

Criterion validity of the WICS scores. Regarding concurrent validity of the WICS scores (part of criterion validity), as expected, the WICS was negatively and significantly ($p < .01$) related to positive affect, but positively and significantly ($ps < .01$) related to negative affect, depressive symptoms, loneliness, and perceived life stress (see Table 2). For discriminant validity, the relationship between the WICS scores and the unhealthy diet construct was, as expected, nonsignificant ($p > .05$). Thus, these results provided some initial support for criterion validity of the WICS scores.

Discussion

Collectively, these results suggest that the WICS measures a singular perceptual construct of interpersonal conflict and provides some support for construct and criterion-related validity of the WICS scores. As anticipated, the WICS was positively related to negatively valenced constructs and negatively related to positively valenced constructs within a sample of home care workers, providing support for the utility of this measure within the health care context. However, some validity issues remain to be addressed. First and foremost, the WICS should be evaluated within other worker populations to determine generalizability of this measure, particularly given the characteristics of the sample in Study 1 (e.g., age, sex). Second, although the results of Study 1 provide some evidence of construct and criterion-related validity, a prospective study design would establish further confidence in criterion validity (e.g., predictive validity). Relatedly, incremental validity should be examined where the WICS is compared to another validated measure (e.g., ICAWS). Finally, additional work and safety variables should be examined to determine whether the WICS can be useful in practical assessments of work including variables like turnover intentions and safety adherence. Study 2 was conducted to address these concerns.

Study 2

Study 2 involved an investigation of workplace interpersonal conflict within the food service industry including food service and grocery retail. Two specific food service organizations located in the Pacific Northwest and the Intermountain West expressed concern over their respective employee turnover rates and agreed to participate. The data were used to examine the validity of the WICS scores in Study 2.
Method

Participants. Participants were recruited by direct solicitation from organization managers and researchers to participate in a prospective (pre/post) study assessing working conditions using a two-time survey. Interested employees were sent an e-mail invitation and a link to an online survey (Qualtrics, Provo, UT). A total of 112 participants participated, 69 from the food service franchises and 43 from the grocery retail store. Participants at the food service organization were offered nonmonetary incentives for their participation (i.e., points used to get store-level rewards), and participants at the retail grocery store were included in a random drawing for $25 gift cards. There was almost an equivalent number of men as women (n = 52 men, n = 60 women), and most were European American (93%) with an average age of 22 years (SD = 3.17) and an average of 12.04 years of education (SD = 3.56). Nearly half (49%) of the sample had been working for less than a year, and 41% had been working between one and five years. Of the total 112 original study participants at the first assessment (baseline), only 38 provided viable responses at the second assessment (follow-up). Demographics of these 38 participants showed strong similarities with the demographics of the 112 participants in sex (53% women), ethnicity (92% European American), age (M = 21.98, SD = 3.19), and education (M = 12.31, SD = 3.80).

Measures. Workplace interpersonal conflict was assessed using the 7-item WICS measure and the IGAW (Spector & Jex, 1998). The ICWA is a 4-item measure of the frequency of conflict behaviors over the past month on a 5-point Likert-type frequency scale (1 = never, 5 = very often; α = .79). Similar to Study 1, the same measures, time-point anchors, and scales were used to assess positive (α = .68) and negative affect (α = .71; Larsen & Diener, 1992; Watson et al., 1988) and depressive symptoms (α = .71; Bohannon et al., 2003). Work engagement during the past month was assessed using the 9-item Work Engagement Scale developed by Schaufeli, Bakker, and Salanova (2006), on a 5-point Likert-type frequency scale (1 = never, 5 = very often; α = .93). Turnover intention was captured using the 3-item measure developed by Hom, Griffeth, and Sellaro (1984), where participants indicated their intentions to leave their current organization on a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree; α = .91). Job search behavior in the past month was examined using the four items developed by Kopelman, Rovenpor, and Millsap (1992) on a 5-point Likert-type frequency scale (1 = never, 5 = nearly every day; α = .79). Safety compliance was assessed using Griffin and Neal’s (2000) 3-item measure on a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree; α = .94). Constructs were assessed in both surveys with the exception of work engagement (baseline only), and depressive symptoms and safety compliance (follow-up only).

Procedure. We first obtained institutional review board approval from Brigham Young University–Idaho (F2014-0001). Participation in this pre/post study involved the completion of an online questionnaire, which took approximately 10 minutes. Participants were contacted six weeks following the administration of the baseline survey for a follow-up questionnaire. A total of 47 participants completed the follow-up questionnaire for a retention rate of 42%. Nine responses were omitted from the dataset due to a technical issue (i.e., online data were incomplete when downloaded), leaving 38 total viable responses for analysis. Given the suboptimal retention rate, we investigated differences between those who completed the follow-up measure and those who did not, and found only one significant difference among all study variables (p = .04); those who completed the follow-up had self-reported greater levels of negative affect (M = 2.26, SD = 0.78) than those who did not (M = 1.94, SD = 0.75). No other significant differences emerged on any other study variable between the two groups, suggesting a lack of systematic bias in responses.

Results

Internal consistency and construct validity. Descriptive results including means, standard deviations, and correlations for the baseline and follow-up surveys are presented in Table 3. Internal consistency and construct validity were assessed using only baseline data. No demographic variables were statistically related to the WICS scores. Similar to Study 1, when Item 7 of the WICS was not included, every measure of internal consistency and construct validity either did not change or improved and, as such, the 6-item WICS was analyzed. Intertemporal correlations of the WICS ranged from 0.26 to 0.73 (M = 0.54, SD = 0.14), the corrected item-total correlation was above the recommended cutoff of .40 (Gliem & Gliem, 2003), and Cronbach’s alpha was .88, suggesting acceptable internal consistency. Using Principal Component Analysis and Direct Oblimin rotation, the EFA yielded one factor (no rotation resulted) because only one Eigenvalue
was above the convention of 1.0 (3.76 with the next highest at .81), and visual inspection of the scree plot supported this conclusion. Each item loaded highly on the factor, ranging from 0.56 to 0.86 ($M = 0.78$, $SD = 0.11$), suggesting that all items were acceptable in the assessment of a single factor. The CFA produced similar results with standardized item loadings ranging from 0.45 to 0.85 ($M = 0.74$, $SD = 0.15$).

Model fit was not acceptable using chi-square analysis, $\chi^2(9) = 28.60$, $p < .05$, but several indexes suggested adequate or approaching adequate model fit (SRMR = .047; NFI = 0.92; IFI = 0.94; TLI = .86; CFI = 0.94; RMSEA = .12). Although both the TLI and RMSEA continued to be suboptimal, it is important to note that all model fit indexes improved from the estimates in Study 1, suggesting that Study 1 might have provided initial low estimates. As predicted, the WICS was strongly positive and significantly related to the ICAWS at baseline, suggesting convergent validity of the WICS scores (see Table 3). These results collectively support the WICS as an appropriate and valid measure.

**Criterion validity of the WICS scores.** For criterion validity, concurrent, predictive, and incremental validity of the WICS scores were evaluated. Concurrent validity was demonstrated using only data from the baseline with negative and significant ($ps < .01$) relationships to positive affect and work engagement as well as positive and significant ($ps < .01$) relationships to negative affect, organizational turnover intentions, and job search behavior. To establish predictive validity, Ordinary Least Squares (OLS) Linear Regression models were evaluated where each outcome measure at follow-up was regressed individually (univariate) on the WICS scores at baseline because temporal primacy is necessary to establish a true predictive relationship. As expected, the WICS scores negatively and significantly ($ps < .01$) predicted positive affect and safety compliance as well as positively and significantly ($ps < .01$) predicted negative affect, turnover intentions for organization, job search behavior, and depressive symptoms (see Table 4). These results suggest that the WICS is useful in the prediction of theoretically related workplace constructs.

Initial modest support for incremental validity was observed; the WICS scores were more strongly associated (bivariate correlations) than the ICAWS with every other study variable (in the expected direction), ranging from an increase of 0.03 to 0.12 ($M = 0.07$, $SD = 0.04$) in favor of the WICS, and this overall mean difference was statistically significant, $t(5) = 4.29$, $p = .008$. The difference in percent variance explained ($R^2$) at baseline, ranging from 0.02 to 0.09 ($M = 0.06$, $SD = 0.03$), was also significantly in favor of the WICS, $t(5) = 4.90$, $p = .005$. To establish additional incremental validity, we conducted regression analyses to compare the ICAWS scores with the WICS scores (see Table 4). Of the six outcomes examined, the WICS had a significant ($ps < .05$) and stronger predictive relationship (in

---

### TABLE 3

<table>
<thead>
<tr>
<th>Variable</th>
<th>$M$ (SD)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Workplace Interpersonal Conflict (WICS)</td>
<td>1.83 (.78)</td>
<td>.88</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Interpersonal Conflict at Work (ICAWS)</td>
<td>1.41 (.64)</td>
<td>.77</td>
<td>.79</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Positive Affect</td>
<td>3.77 (.65)</td>
<td>-.47</td>
<td>-.39</td>
<td>.68</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Negative Affect</td>
<td>2.04 (.78)</td>
<td>.51</td>
<td>.42</td>
<td>-.54</td>
<td>.71</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Work Engagement</td>
<td>3.48 (.80)</td>
<td>-.39</td>
<td>-.29</td>
<td>-.59</td>
<td>-.61</td>
<td>.93</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Turnover Intentions</td>
<td>2.83 (1.19)</td>
<td>.35</td>
<td>.30</td>
<td>-.40</td>
<td>.37</td>
<td>-.63</td>
<td>.91</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Job Search Behavior</td>
<td>1.63 (0.63)</td>
<td>.43</td>
<td>.40</td>
<td>-.40</td>
<td>.45</td>
<td>-.39</td>
<td>.58</td>
<td>.79</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Positive Affect</td>
<td>3.58 (.76)</td>
<td>-.54</td>
<td>-.54</td>
<td>.63</td>
<td>-.53</td>
<td>.56</td>
<td>-.62</td>
<td>-.31</td>
<td>.77</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Negative Affect</td>
<td>2.21 (.79)</td>
<td>.43</td>
<td>.40</td>
<td>-.43</td>
<td>.53</td>
<td>-.58</td>
<td>.61</td>
<td>.25</td>
<td>-.74</td>
<td>.70</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Turnover Intentions + Organization</td>
<td>2.91 (1.29)</td>
<td>.67</td>
<td>.57</td>
<td>-.50</td>
<td>.47</td>
<td>-.62</td>
<td>.86</td>
<td>.70</td>
<td>-.48</td>
<td>.57</td>
<td>.92</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Job Search Behavior</td>
<td>1.86 (1.07)</td>
<td>.59</td>
<td>.66</td>
<td>-.49</td>
<td>.66</td>
<td>-.38</td>
<td>.61</td>
<td>.77</td>
<td>-.57</td>
<td>.42</td>
<td>.69</td>
<td>.91</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Safety Compliance</td>
<td>4.50 (0.62)</td>
<td>-.40</td>
<td>-.22</td>
<td>-.59</td>
<td>-.46</td>
<td>-.46</td>
<td>-.32</td>
<td>.52</td>
<td>-.25</td>
<td>-.31</td>
<td>-.33</td>
<td>.94</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. Depressive Symptoms</td>
<td>1.48 (0.51)</td>
<td>.49</td>
<td>.54</td>
<td>-.59</td>
<td>.50</td>
<td>-.46</td>
<td>.49</td>
<td>.38</td>
<td>-.69</td>
<td>.51</td>
<td>.39</td>
<td>.55</td>
<td>-.25</td>
<td>.71</td>
</tr>
</tbody>
</table>

Note: Variables 1–7 are from the baseline assessment and variables 8–13 are from the follow-up assessment. Internal consistency (Cronbach’s alpha) estimates are provided on the diagonal. $p < .05$, $*p < .01$. 

---
the expected direction) than the ICAWS for four: positive affect, negative affect, turnover intentions, and safety compliance. The overall mean difference (where positive values indicate stronger predictions by the WICS over the ICAWS) in the standardized regression coefficients \((M = 0.03, SD = 0.09)\) and the overall mean difference in the percent variance explained \((M = 0.02, SD = 0.09)\) demonstrated that WICS scores had some improved predictive ability over the ICAWS, though these differences were not significant overall, \(t(5) = .82, p = .45\), and, \(t(5) = .54, p = .61\), respectively. However, the WICS was statistically more strongly predictive of safety compliance \((r = -.40; 95\% \text{ CI} = -.23, -.55)\) than the ICAWS \((r = -.22)\). Taken together, these results provide additional support for criterion-related validity of the WICS scores and tentatively suggest that the WICS may offer a better alternative to assessing workplace interpersonal conflict.

Discussion

Overall, the results from Study 2 are consistent with the results from Study 1, adding support for the WICS. First, the WICS was related, in the expected direction, to work health and safety variables including work engagement, turnover intentions, job search behavior, and safety compliance. Second, convergent validity was observed by the strong correlation between the WICS and ICAWS scores, a well-validated measure of workplace interpersonal conflict (Spector & Jex, 1998). Third, regarding predictive validity, the WICS was a significant predictor of many important outcomes including organizational turnover intentions, safety compliance, and acute depressive symptomology, which suggests the potential usefulness of the WICS measure in organizational, safety, and health domains. Fourth, the WICS was more strongly associated with outcome variables of interest than the ICAWS, suggesting a potential practical benefit to using the WICS, particularly regarding safety outcomes.

Study 2 provided additional contributions to the establishment of the validity of the WICS scores, lending further confidence in the usefulness of the measure. Although the worker sample, prospective study design, and ability to compare the WICS to the ICAWS addressed many limitations of Study 1, some problems remained. For instance, the poor retention rate at follow-up might have biased results toward those experiencing more negative mood in our analyses of predictive and incremental validity. However, it is possible that this smaller sample size could have inflated Type II error, potentially making the results observed even more noteworthy. Finally, only one validated measure was compared to the WICS, limiting convergent and incremental validity. Therefore, Study 3 was conducted to evaluate the WICS within a larger and more diverse sample, to further investigate the relationship of the WICS relative to both the ICAWS and the ICS, and to determine whether the WICS may offer a better alternative in workplace assessments of the psychosocial context.

Study 3

Study 3 encompassed a large, online sample that was gathered using the online workforce platform Amazon Mechanical Turk (MTurk). MTurk allows interested parties to pay an on-demand workforce to complete tasks using a computer and Internet connection. Data collected were used to assess the validity of the WICS scores.
Method

Participants. Study 3 participants were recruited through MTurk by posting a short description and anonymous link to an online survey (Qualtrics, Provo, UT) as a task to complete for a one-time payment of $1.25. A total of 646 participants completed the survey, though only 632 met inclusion criteria. In the description of the study posted on MTurk, eligibility included employment outside of their involvement with MTurk and specified that self-employed individuals were not eligible. Average age of the sample was 34.52 years (SD = 11.78), the number of men and women participants were nearly equivalent (55% men), and average education was 14.47 years (SD = 3.23). Most of the sample was European American (77%), with African American (1%), Asian (8%), Native Hawaiian/Pacific Islander (0.2%), Hispanic/Latino(a) (6%), multiple race (2%), and other (0.5%) represented in the sample. Average tenure at the participants’ current jobs was 5.81 years (SD = 6.06).

Moreover, 78% of participants reported being employed full time (n = 495), 40% (n = 252) reported supervising other employees, and the average number of hours worked per week was 38.41 hours (SD = 10.54). Participants reported a wide range of occupations such as “data services programmer” and “brick mason laborer”, with “teacher” reported most frequently (27%), followed by “programmer” and “brick mason laborer” (13.9%), “sales” (10.7%), “data services manager” (8.8%), and “teaching” (3.9%). The majority of participants were “teacher” reported most frequently (27%), followed by “programmer” and “brick mason laborer” (13.9%), “sales” (10.7%), “data services manager” (8.8%), and “teaching” (3.9%).

In terms of industry, the most commonly reported were sales (n = 87, 14.20%), healthcare (n = 68, 11.10%), technology (n = 66, 10.80%), education (n = 56, 9.13%), and food services (n = 43, 7.01%).

Measures. Similar to Study 2, workplace interpersonal conflict over the past month was assessed using both the WICS and the ICAWS (a = .88; Spector & Jex, 1998), and also included the 8-item ICS (a = .90; Jehn, 1995), with all three measures using the same 5-point Likert-type scale (1 = never, 5 = very often). Utilizing the same measures, time-point anchors, and scales as Study 1 and Study 2, we measured positive (a = .78) and negative (a = .78) affect (Larsen & Diener, 1992; Watson et al., 1988), depressive symptoms (a = .82; Bohannon et al., 2003), loneliness (a = .91; Hughes et al., 2008), work engagement (a = .95; Schaufeli et al., 2006), organizational turnover intentions (a = .93; Hom et al., 1984), job search behavior (a = .86; Kopelman et al., 1992), perceived life stress (a = .90; Cohen et al., 1983), safety compliance (a = .94; Griffin & Neal, 2000), and unhealthy sugar diet (Buxton et al., 2009). Spector and Jex’s (1998) 18-item Physical Symptoms Inventory was included, and participants indicated whether they had experienced any of the listed physical symptoms (e.g., headache, chest pain) in the past 30 days. Finally, counterproductive work behavior during the past 30 days was assessed using the 10-item Counterproductive Work Behavior Checklist (a = .88; Spector, Bauer, & Fox, 2010), a 5-point Likert-type frequency scale (1 = never, 5 = everyday).

Procedure. We first obtained institutional review board approval from Brigham Young University–Idaho (S2016-F008). Participation in the study consisted of completing the online survey in English, whereupon participants received a random numerical code and entered this into the MTurk website to receive payment. As an indicator of data fidelity, completed surveys that took less than 150 seconds were excluded because we observed a natural break in the data and reasoned that 2.5 minutes was insufficient time to accurately complete the survey. As such, 14 responses were excluded from analysis. After these 14 entries were dropped, the average survey completion time was 8.72 minutes (SD = 4.91).

Results

Internal consistency and construct validity. Descriptive results including means, standard deviations, and correlations between the WICS and other variables are located in Table 5. Among the demographic variables, age (r = -.08, p = .044) and years of education (r = -.17, p < .001) were statistically related to the WICS scores. Similar to Study 1 and Study 2, when Item 7 was not included, every measure of internal consistency and construct validity either did not change or improved and, as such, only the 6-item WICS was analyzed. Interitem correlations of the WICS ranged from 0.51 to 0.73 (M = 0.62, SD = 0.07), corrected item-total correlation was above the recommended cutoff of .40 (Gliem & Gliem, 2003), and Cronbach’s alpha was 0.91, collectively indicating acceptable internal consistency of the scale. Using Principal Component Analysis and Direct Oblimin rotation, the EFA yielded one factor (no rotation resulted) because only one Eigenvalue was above the convention of 1.0 (4.11 with the next highest at .56), and visual inspection of the scree plot supported this conclusion. The EFA yielded only one factor with each item loading highly on the factor, ranging from 0.76 to 0.88. (M = 0.83, SD = 0.04), and the CFA produced similar results with standardized item loadings ranging from 0.69 to 0.86 (M = 0.79, SD = 0.05).
Model fit was not acceptable using chi-square analysis, $\chi^2(9) = 103.60, p < .05$. However, nearly every other model fit index exceeded recommended cutoffs (SRMR = .084; NFI = .96; IFI = .96; TLI = .93; CFI = .96; RMSEA = .13). It should be noted that TLI is much closer to the recommended 0.95 cutoff (Hu & Bentler, 1999) and that RMSEA, though still not less than .06, is the only index that was not satisfied; and the remaining five indices suggested acceptable model fit. Likely due to the larger and more diverse sample in Study 3, each model fit index improved over Study 2, suggesting a better fit and less sample bias. Respecting convergent validity (see Table 5), the correlation between the WICS and the ICAWS scores were strongly positive and significant ($p < .001$) as well as the correlation between the WICS and ICS scores ($p < .001$). For discriminant validity, the relationship between the WICS scores and unhealthy diet was nonsignificant, as expected ($p > .05$).

Criterion validity of the WICS scores. Concurrent validity of the WICS scores was established by observing negative and significant ($ps < .01$) relationships with positive affect, safety compliance, and work engagement. Additionally, positive and significant ($ps < .01$) relationships with negative affect, depressive symptoms, loneliness, perceived stress, turnover intentions, job search behavior, physical symptoms, and counterproductive work behavior emerged. Regarding incremental validity (see Table 6), the WICS scores were more strongly associated than the ICAWS scores with every variable (in the expected direction) except counterproductive work behavior and safety compliance, ranging from an additional 0.02 (physical symptoms) to 0.13 (turnover intentions) increase ($M = 0.06, SD = 0.05$) in favor of the WICS, $t(10) = 3.98, p = .003$. Percent variance explained ($R^2$) was also greater for the WICS compared to the ICAWS, ranging from .01 to .10 ($M = 0.04, SD = 0.04$), which was also significant, $t(10) = 3.32, p = .008$. Compared to the ICS scores, the WICS scores were more strongly and significantly associated with

---

**TABLE 5**

<table>
<thead>
<tr>
<th>Variable</th>
<th>$M (SD)$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Workplace Interpersonal Conflict Scale (WICS)</td>
<td>1.90 (0.82)</td>
<td>.91</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Interpersonal Conflict at Work (ICAWS)</td>
<td>1.64 (0.79)</td>
<td>.87**</td>
<td>.88</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Intragroup Conflict Scale (ICS)</td>
<td>2.21 (0.81)</td>
<td>.77**</td>
<td>.73**</td>
<td>.90</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Positive Affect</td>
<td>3.38 (0.85)</td>
<td>-33**</td>
<td>-25**</td>
<td>-22**</td>
<td>.78</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Negative Affect</td>
<td>2.00 (0.80)</td>
<td>.50**</td>
<td>.44**</td>
<td>.40**</td>
<td>-51**</td>
<td>.78</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Depressive Symptoms</td>
<td>7.95 (3.10)</td>
<td>.41**</td>
<td>.36**</td>
<td>.36**</td>
<td>-.48**</td>
<td>.60**</td>
<td>.82</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Loneliness</td>
<td>2.28 (1.03)</td>
<td>.43**</td>
<td>.35**</td>
<td>.33**</td>
<td>-.46**</td>
<td>.58**</td>
<td>.64**</td>
<td>.91</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Turnover Intentions</td>
<td>2.67 (1.33)</td>
<td>.47**</td>
<td>.34**</td>
<td>.36**</td>
<td>-.44**</td>
<td>.44**</td>
<td>.36**</td>
<td>.37**</td>
<td>.93</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Job Search Behavior</td>
<td>1.79 (0.94)</td>
<td>.49**</td>
<td>.45**</td>
<td>.39**</td>
<td>-.28**</td>
<td>.42**</td>
<td>.38**</td>
<td>.32**</td>
<td>.71**</td>
<td>.86</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Work Engagement</td>
<td>3.33 (0.92)</td>
<td>-.34**</td>
<td>-.22**</td>
<td>-.21**</td>
<td>.72**</td>
<td>-.48**</td>
<td>-.36**</td>
<td>-.38**</td>
<td>-.56**</td>
<td>-.37**</td>
<td>.95</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Safety Compliance</td>
<td>4.22 (0.75)</td>
<td>-.24**</td>
<td>-.24**</td>
<td>-.19**</td>
<td>.16**</td>
<td>-.19**</td>
<td>-.11**</td>
<td>-.21**</td>
<td>-.09**</td>
<td>-.14**</td>
<td>.17**</td>
<td>.94</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Physical Symptoms</td>
<td>4.77 (3.81)</td>
<td>.33**</td>
<td>.31**</td>
<td>.33**</td>
<td>-.26**</td>
<td>.44**</td>
<td>.61**</td>
<td>.39**</td>
<td>.27**</td>
<td>.29**</td>
<td>-.19**</td>
<td>-.05</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>13. Counterproductive Work Behaviors</td>
<td>1.40 (0.54)</td>
<td>.54**</td>
<td>.58**</td>
<td>.49**</td>
<td>-.20**</td>
<td>.46**</td>
<td>.33**</td>
<td>.32**</td>
<td>.29**</td>
<td>.44**</td>
<td>-.21**</td>
<td>-.35**</td>
<td>.31**</td>
<td>.88</td>
</tr>
<tr>
<td>14. Unhealthy Sugar Diet</td>
<td>0.73 (0.83)</td>
<td>.05</td>
<td>.07</td>
<td>.07</td>
<td>-.08</td>
<td>.16**</td>
<td>.18**</td>
<td>.10**</td>
<td>-.03</td>
<td>.03</td>
<td>-.04</td>
<td>-.07</td>
<td>.22**</td>
<td>.15**</td>
</tr>
</tbody>
</table>

Note: Internal consistency (Cronbach’s alpha) estimates, where appropriate, are provided on the diagonal. † $p < .05$. ‡ $p < .01$.
every variable (except physical symptoms; \( M = 0.08, \ SD = 0.04; \ t(10) = 6.63, \ p < .001 \)). Similarly, the percent variance explained \( (R^2) \) was significantly greater for the WICS scores compared to the ICS scores, \( t(10) = 6.63, \ p < .001 \). Collectively, these results provide additional validity evidence and support the WICS as an appropriate measure of interpersonal conflict.

**Discussion**

Overall, the results from Study 3 were consistent with the results from Study 1 and Study 2, adding additional evidence to support the WICS as a valid measure in regard to both construct and criterion-related validity. First, the large sample size and occupational diversity of the sample boosts confidence in the generalization of these findings. Second, the WICS was related, in the expected direction, to additional work health and safety variables including counterproductive work behaviors, safety compliance, and physical health symptoms. Third, the WICS scores demonstrated strong convergent validity and additional incremental validity when compared to the ICWAS (Spector & Jex, 1998) and ICS scores (Jehn, 1997). The online sample made fidelity of the data very difficult to determine, though obvious cases of response set bias were eliminated by deleting out surveys with too little time spent on completion. Previous research has also suggested that data collected via MTurk are as reliable as other methods (Buhrmester, Kwang, & Gosling, 2011). Another limitation of the present study was the limited support for discriminant validity due to the examination of only one relationship. However, though when compared to other forms of construct and criterion validity such as incremental validity, discriminant validity may not be as critical when establishing the utility of a new measure. In conclusion, these results collectively suggest that the WICS may provide an improved assessment of workplace interpersonal conflict in organizational contexts.

**General Discussion**

Although interpersonal conflict has been studied extensively in the workplace, the validation evidence indicating adequate psychometric properties of available self-report measures has been lacking. In this present examination, three studies were conducted that evaluated a newly constructed brief self-report perceptual scale, the WICS, based on a previous qualitative study (Wright et al., 2014). In Study 1, we investigated the construct validity of the WICS scores within a group of home care workers including important health outcomes, whereby some support for construct and criterion

---

**TABLE 6**

<table>
<thead>
<tr>
<th>Outcome Variable</th>
<th>WICS</th>
<th>ICWAS</th>
<th>ICS</th>
<th>Difference Scores of Absolute Values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( r )</td>
<td>( R^2 )</td>
<td>( r )</td>
<td>( R^2 )</td>
</tr>
<tr>
<td>Positive Affect</td>
<td>-.33</td>
<td>.11</td>
<td>-.25</td>
<td>.06</td>
</tr>
<tr>
<td>Negative Affect</td>
<td>.50</td>
<td>.25</td>
<td>.44</td>
<td>.19</td>
</tr>
<tr>
<td>Depressive Symptoms</td>
<td>.41</td>
<td>.17</td>
<td>.36</td>
<td>.13</td>
</tr>
<tr>
<td>Loneliness</td>
<td>.43</td>
<td>.19</td>
<td>.35</td>
<td>.12</td>
</tr>
<tr>
<td>Perceived Stress</td>
<td>.48</td>
<td>.23</td>
<td>.40</td>
<td>.16</td>
</tr>
<tr>
<td>Turnover Intentions</td>
<td>.41</td>
<td>.22</td>
<td>.34</td>
<td>.12</td>
</tr>
<tr>
<td>Job Search Behavior</td>
<td>.49</td>
<td>.24</td>
<td>.45</td>
<td>.20</td>
</tr>
<tr>
<td>Work Engagement</td>
<td>-.34</td>
<td>.11</td>
<td>-.22</td>
<td>.05</td>
</tr>
<tr>
<td>Safety Compliance</td>
<td>-.24</td>
<td>.06</td>
<td>-.24</td>
<td>.06</td>
</tr>
<tr>
<td>Physical Symptoms</td>
<td>.33</td>
<td>.11</td>
<td>.31</td>
<td>.10</td>
</tr>
<tr>
<td>CWB</td>
<td>.54</td>
<td>.29</td>
<td>.58</td>
<td>.34</td>
</tr>
<tr>
<td>M (SD)</td>
<td>(.04)</td>
<td>.06(.04)</td>
<td>(.04)</td>
<td>.06(.04)</td>
</tr>
</tbody>
</table>

*Note: All relationships reported above were statistically significant at the \( p < .01 \) level, and biserial correlations \( (r) \) are reported. CWB is Counterproductive Work Behavior. Each relationship represented above resulted from biserial correlations and in the case of \( R^2 \), the correlation coefficient was squared. Positive values in the last four difference columns on the right represent stronger relationships for the Workplace Interpersonal Conflict Scale (WICS) over the Interpersonal Conflict at Work Scale (ICWAS) or Intragroup Conflict Scale (ICS).*
validity was obtained. In Study 2, a prospective two-time point study, the WICS was examined among food industry employees and included additional important work variables (e.g., turnover intentions, safety compliance) and a comparison of the WICS to the validated ICAWS (Spector & Jex, 1998) in order to provide some incremental validity. In Study 3, the WICS scores were examined within a large and diverse MTurk sample and compared to the ICAWS (Spector & Jex, 1998) and the ICS (Jehn, 1995), providing further evidence for construct and criterion validity. Six items of the WICS were strongly correlated (interitem correlations) across all three studies (\( M = 0.61, SD = 0.08 \)) and were well-suited for the purpose of assessing workplace interpersonal conflict. The specific content of the WICS came from themes identified in a qualitative study (Wright et al., 2014), and the WICS includes elements not previously considered in most prior measures such as perceived injustices, goal impediments, and being unfairly blamed. Moreover, because the WICS included components that many other measures do not, it seems likely that the WICS more comprehensively captured conflict in the workplace, allowing for the identification of unique constellations of cognitions, emotions, and behaviors in perceptions of conflict. However, caution should be exercised in the interpretation of specific items because the conflict behaviors represented in the WICS may or may not, by themselves, suggest interpersonal conflict. For instance, if there is frequent yelling at the workplace, this could mean conflict is present or other conditions may necessitate the yelling (e.g., manager has a deadline for a subordinate to meet).

Given our results, however, confidence in the WICS’ ability to assess workplace interpersonal conflict seems robust. Removal of one of the original items of the 7-item WICS (“been given unclear directions by others at work”) due to statistical (low interitem correlations, low factor loadings) and theoretical issues (item wording may suggest that this is more in-line with role ambiguity) consistently improved construct validity and was consistent with Hinkin’s (1998) recommendation that a scale have four to six items measuring one construct. Moreover, despite the observation that RMSEA was consistently poor, multiple fit indexes provided evidence that the WICS had adequate and appropriate model fit, suggesting a psychometrically sound measure. Indeed, a critical reason for having multiple fit indexes is to increase the confidence in the evaluation of a measure so that the judgment of the appropriateness of the scale is not influenced too strongly by the limitations of any one index (Hu & Bentler, 1999). Thus, RMSEA failing to provide support for the WICS measure, although certainly not ideal, should not discredit it altogether because all other indexes supported it. Additionally, our comparison results between the WICS, the ICAWS (Spector & Jex, 1998), and the ICS (Jehn, 1995), demonstrated substantive and statistically significant improvements in correlations with and predictions of important organizational, health, and safety outcomes. In sum, these results provide evidence of the consistency, construct validity, and criterion validity of the WICS scores for employee well-being, organization efficiency, and employee safety.

Third, the WICS seems useful within multiple occupational populations. Indeed, the WICS may be used as a diagnostic tool in assessments of the workplace in terms of several important workplace outcomes (e.g., turnover intentions, work engagement), worker health outcomes (e.g., depressive symptoms, perceived stress), and to safety compliance, which makes it particularly useful to organizations interested in a broad spectrum of outcomes. For example, an organization seeking to assess employee turnover risk may use the WICS as a proxy measure to predict future turnover. Going a step further, the same organization may seek to reduce or eliminate workplace interpersonal conflict assessed through the WICS in efforts to prevent employee turnover and improve worker productivity. Finally, although further validation is needed, the results from Study 3 suggest that the WICS seems sensitive enough to capture some subtle differences between lateral and hierarchical conflicts, which can lead to very different outcomes (Frone, 2000). Thus, the WICS may be an appropriate diagnostic tool for hierarchical conflicts such as conflicts between managers and subordinates, or for comparisons between lateral and hierarchical conflicts.
<table>
<thead>
<tr>
<th>Psychometric Property</th>
<th>Study 1</th>
<th>Study 2</th>
<th>Study 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Scale Reliability (Internal Consistency)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cronbach’s Alpha</td>
<td>.92</td>
<td>.88</td>
<td>.91</td>
</tr>
<tr>
<td><strong>Construct Validity</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interitem Correlation M (SD)</td>
<td>.66 (0.08)</td>
<td>.54 (0.14)</td>
<td>.62 (0.07)</td>
</tr>
<tr>
<td>Range of Factor Loadings for EFA</td>
<td>.80-.91</td>
<td>.56-.86</td>
<td>.76-.88</td>
</tr>
<tr>
<td>Model Fit Indexes for CFA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chi-Square (df)</td>
<td>p &gt; .05</td>
<td>60.50(9), p &lt; .05</td>
<td>28.6(9), p &lt; .05</td>
</tr>
<tr>
<td>Standardized Root Mean Square Residual (SRMR)</td>
<td>&lt; .08</td>
<td>.048</td>
<td>.047</td>
</tr>
<tr>
<td>Normed Fit Index (NFI)</td>
<td>.91</td>
<td>.92</td>
<td>.92</td>
</tr>
<tr>
<td>Incremental Fit Index (IFI)</td>
<td>.95</td>
<td>.92</td>
<td>.94</td>
</tr>
<tr>
<td>Tucker-Lewis Index (TLI)</td>
<td>.95</td>
<td>.81</td>
<td>.86</td>
</tr>
<tr>
<td>Comparative Fit Index (CFI)</td>
<td>.95</td>
<td>.92</td>
<td>.94</td>
</tr>
<tr>
<td>Root Mean Square Error of Approximation (RMSEA)</td>
<td>&lt; .06</td>
<td>.20, 90% CI [.15, .24]</td>
<td>.12, 90% CI [.07, .17]</td>
</tr>
<tr>
<td>Range of Factor Loadings for CFA</td>
<td>.74-.90</td>
<td>.45-.85</td>
<td>.69-.86</td>
</tr>
<tr>
<td><strong>Concurrent Validity (bivariate correlations)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive Affect</td>
<td>-.29**</td>
<td>-.47**</td>
<td>-.33**</td>
</tr>
<tr>
<td>Negative Affect</td>
<td>-.29**</td>
<td>-.51**</td>
<td>-.50**</td>
</tr>
<tr>
<td>Depressive Symptoms</td>
<td>.23**</td>
<td>--</td>
<td>.41**</td>
</tr>
<tr>
<td>Loneliness</td>
<td>.25**</td>
<td>--</td>
<td>.43**</td>
</tr>
<tr>
<td>Perceived Stress</td>
<td>.27**</td>
<td>--</td>
<td>.48**</td>
</tr>
<tr>
<td>Turnover Intentions – Organization</td>
<td>--</td>
<td>.35**</td>
<td>.47**</td>
</tr>
<tr>
<td>Job Search Behavior</td>
<td>--</td>
<td>.43**</td>
<td>.49**</td>
</tr>
<tr>
<td>Work Engagement</td>
<td>--</td>
<td>-.39**</td>
<td>-.34**</td>
</tr>
<tr>
<td>Safety Compliance</td>
<td>--</td>
<td>--</td>
<td>-.24**</td>
</tr>
<tr>
<td>Physical Symptoms</td>
<td>--</td>
<td>--</td>
<td>.33**</td>
</tr>
<tr>
<td>Counterproductive Work Behaviors</td>
<td>--</td>
<td>--</td>
<td>.54**</td>
</tr>
<tr>
<td><strong>Discriminant Validity (bivariate correlations)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unhealthy Sugar Diet</td>
<td>.04</td>
<td>--</td>
<td>.05</td>
</tr>
<tr>
<td><strong>Predictive Validity (OLS linear regression)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive Affect</td>
<td>--</td>
<td>-.54**</td>
<td>--</td>
</tr>
<tr>
<td>Negative Affect</td>
<td>--</td>
<td>.43**</td>
<td>--</td>
</tr>
<tr>
<td>Turnover Intentions – Organization</td>
<td>--</td>
<td>.67**</td>
<td>--</td>
</tr>
<tr>
<td>Job Search Behavior</td>
<td>--</td>
<td>.59**</td>
<td>--</td>
</tr>
<tr>
<td>Safety Compliance</td>
<td>--</td>
<td>-.40**</td>
<td>--</td>
</tr>
<tr>
<td>Depressive Symptoms</td>
<td>--</td>
<td>.49**</td>
<td>--</td>
</tr>
<tr>
<td><strong>Convergent Validity (bivariate correlations)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interpersonal Conflict at Work Scale (ICAWS)</td>
<td>--</td>
<td>.77**</td>
<td>.87**</td>
</tr>
<tr>
<td>Intragroup Conflict Scale (ICS)</td>
<td>--</td>
<td>--</td>
<td>.77**</td>
</tr>
<tr>
<td><strong>Incremental Validity</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlation Strength Difference: WICS &amp; ICAWS</td>
<td>--</td>
<td>.07 (.04)**</td>
<td>.06 (.04)**</td>
</tr>
<tr>
<td>Correlation Strength Difference: WICS &amp; ICS</td>
<td>--</td>
<td>.06 (.03)**</td>
<td>.04 (.04)**</td>
</tr>
<tr>
<td>Correlation Strength Difference: ICAWS &amp; ICS</td>
<td>--</td>
<td>--</td>
<td>.08 (.04)**</td>
</tr>
</tbody>
</table>

Note. EFA is Exploratory Factor Analysis and CFA is Confirmatory Factor Analysis. Model cutoff criteria were taken from Hu and Bentler (1999). For predictive validity, OLS linear regression was conducted where each outcome (collected in the follow-up assessment) was regressed on the Workplace Interpersonal Conflict Scale (WICS) in univariate regression models, and standardized betas (β) are reported. Regarding incremental validity, positive values represent when the strength of the relationship of the WICS and the outcome variable was stronger (absolute value) than the ICAWS or the ICS. All WICS analyses reflect only the first six items of the WICS. *p < .05, **p < .01.
Workplace Interpersonal Conflict Scale

Limitations and Future Directions

Our current examination of the WICS has a few limitations. First, all outcomes were not included in all three studies, making it difficult to be certain whether the WICS would have had similar relationships in each context. However, as a way to monitor this issue, positive and negative affect were included across all studies, which, as expected, yielded similar general patterns of results. It should be noted that the Study 2 relationships between affect and conflict were generally stronger, likely due to the temporary nature of food service jobs and the younger ages of the participants. More specifically, conflict may have a stronger impact on mood for temporary jobs and for younger employees. All forms of validity were unable to be examined in each study, making it necessary to infer that similar results would have been observed. Again, given the general consistency of the results across the studies for reliability and construct validity, this inference is likely justified, and the fact that all major types of validity were examined in at least one study is a substantial improvement in the literature (Bendersky & Hays, 2012; Jehn, 1995; Schieman & Reid, 2008; Spector & Jex, 1998). Third, our studies relied exclusively on self-reported data, which can be affected by subjective biases, though when studying occupational stress processes (including perceptions), self-report measures are typically appropriate (Spector, 2006). Finally, in efforts to produce a brief scale, only seven items were originally drafted. Thus, it is difficult to assess whether additional items would have altered the results. However, these items are consistent across the literature (i.e., Schieman & Reid, 2008; Spector & Jex, 1998) and one item was dropped, which reduced this concern.

Notwithstanding these limitations, the present examination made several important contributions to future research. First, the examination addressed the critical issues of (a) using a current conceptualization of workplace interpersonal conflict; (b) empirical derivation, assessment of properties of conflict, and using multiple occupational samples; and (c) rigorous psychometric evaluation. As such, the WICS offers many advantages beyond other currently available measures of workplace interpersonal conflict. Future research could administer the WICS within a wider variety of occupations than those examined here including high stress occupations (e.g., firefighters, physicians, midlevel managers), during demanding conditions (e.g., job insecurity, low decision latitude, organizational change), or among managers/supervisors for hierarchical forms of conflict.

Research should investigate the relationship between the WICS and other more objective variables such as job performance, turnover, absenteeism, and clinical diagnoses of depression. Another domain to explore could be the assessment of conflict from a dyadic perspective (e.g., supervisor-subordinate, coworker-coworker) or in other contexts beyond the workplace (e.g., marriage and family counseling). As identified by Weingart et al. (2015), it is important to consider how conflict can be assessed not only according to frequency, as done here, but in terms of intensity of conflict. Finally, other factors such as potential positive outcomes of conflict (e.g., task conflict), organizational culture, cross-cultural interactions, personality, and gender (Wright, Bates, & Ferguson, 2007) should be considered in future assessments.

In conclusion, this study represents an important step to address a major limitation in the assessment of workplace interpersonal conflict. Indeed, research in this area has often relied upon scales with limited psychometric support based on outdated construct definitions. Yet, the WICS offers a valid way of addressing this concern with a brief self-report perceptual scale of conflict interactions with others in the workplace. As such, this research can benefit future researchers and practitioners in the overall quest to promote organizational productivity, employee health, and safety within organizations.

References


https://doi.org/10.1037/0022–3514.99.2.145
https://doi.org/10.1177/1745691110393980
https://doi.org/10.2105/AJPH.2008.158501
https://doi.org/10.1136/oem.2007.037655
https://doi.org/10.1002/job.688
https://doi.org/10.1080/1076-8998.199540118
https://doi.org/10.1108/0030507002316999
https://doi.org/10.1002/job.493060206
https://doi.org/10.1006/jvbe.1992.1001
Wright, Nixon, Peterson, Thompson, Olson, Martin, and Marrott | Workplace Interpersonal Conflict Scale
Workplace Interpersonal Conflict Scale | Wright, Nixon, Peterson, Thompson, Olson, Martin, and Marrott


Author note. Robert R. Wright, Department of Psychology, Brigham Young University–Idaho; Ashley E. Nixon, Atkinson Graduate School of Management, Willamette University; Zachary B. Peterson, Department of Psychology, Brigham Young University–Idaho; Sharon V. Thompson, Oregon Institute of Occupational Health Sciences, Oregon Health & Science University; Ryan Olson, Oregon Institute of Occupational Health Sciences, Oregon Health & Science University; Scott Martin, Department of Psychology, Brigham Young University–Idaho; Devon Marrott, Department of Psychology, Brigham Young University–Idaho.

This research was supported by funding from the National Institute of Occupational Safety and Health (NIOSH) as part of the Oregon Healthy Workforce Center (U19OH010154), a NIOSH Center of Excellence in Total Worker Health™ and by internal funding from Brigham Young University – Idaho for student- and faculty-directed research. Special thanks to Psi Chi journal reviewers for their support.

Correspondence concerning this article should be addressed to Robert R. Wright, Department of Psychology, Brigham Young University–Idaho, Rexburg, ID 83460. E-mail: wrighthro@byui.edu

APPENDIX

The Workplace Interpersonal Conflict Scale (WICS)

For the next set of questions, please show how often each happen by circling one response next to each question. In the past 30 days, how often have you:

<table>
<thead>
<tr>
<th></th>
<th>Never</th>
<th>Almost Never</th>
<th>Sometimes</th>
<th>Often</th>
<th>Very Often</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Felt like you were treated unfairly by others at work?</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>2. Had a disagreement with others over the work you do?</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>3. Been shown a lack of respect or felt underappreciated by others at work?</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>4. Been treated with hostility or rude behavior by others at work?</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>5. Had others yell at you at work?</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>6. Been blamed or criticized for something that was not your fault by others at work?</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>7. Been given unclear directions by others at work?</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

Note. Results from the studies reported here support using only the first six items and not including Item 7.
Personality Traits, Early Maladaptive Schemas, and Severity of Nonsuicidal Self-Injury

Sarah D. Arthurs and Josephine C. H. Tan* 
Lakehead University

ABSTRACT. Nonsuicidal self-injury (NSSI) is used to regulate emotions and communicate distress. Research has examined its prevalence, forms, and functions, but little is known about its severity and the personality and cognitive correlates. This study examined personality traits and early maladaptive schemas in a Canadian nonclinical sample (N = 156, age M = 25.23 years, SD = 8.14 years) consisting of 3 equal-sized groups (n = 52; 43 women, 9 men), matched on sex and age, and that varied on NSSI severity: high NSSI, low NSSI, and control. Participants completed the Deliberate Self-Harm Inventory, Big Five Inventory, and Early Maladaptive Schema Questionnaire – Short Form. Results showed that high and low NSSI groups scored significantly higher than the control on neuroticism, and significantly lower on extraversion, agreeableness, and conscientiousness (Cohen’s d ranges between .50 and 1.53; p ranges from < .01 to < .001). High NSSI group was significantly less agreeable than Low NSSI group (Cohen’s d = .37; p = .04). Both high and low NSSI groups scored significantly higher on all early maladaptive schemas (Cohen’s d ranges between .63 and 1.71; p ranges from < .01 to < .001), except for self-sacrifice beliefs. High NSSI group was significantly higher than Low NSSI (Cohen’s d = .60; p = .005) and control (Cohen’s d = .69; p < .001) groups on beliefs related to unrelenting standards. Given the link between perfectionism and suicide, the findings support the importance of focusing on specific maladaptive beliefs in NSSI with implications for interventions.

Nonsuicidal self-injury (NSSI) involves self-inflicted damage to bodily tissues without suicidal intent, and is often used as a coping mechanism for distress (Nock, 2009). Lifetime prevalence rates range from 13% to 23.2% for nonclinical samples (Jacobson & Gould, 2007), and 17% to 35% for college samples (Whitlock, Eckenrode, & Silverman, 2006). Among clinical samples, prevalence rates are as high as 60% (DiClemente, Ponton, & Hartley, 1991), and repetition rates exceed 40% (Kripalani, Badanapuram, Gash, & Morris, 2007). Higher rates of self-harm incidents and the number of methods used (e.g., cutting, sticking with pins, burning, punching) are reflective of more severe levels of NSSI (MacLaren & Best, 2010; Whitlock, Muehlenkamp, & Eckenrode, 2008). Persons who self-harm repeatedly often require injuries of increasing severity in order to experience the desired effects, predisposing these individuals to...
Personality, EMS, and NSSI | Arthurs and Tan

NSSI and Personality
The dominant model of personality involves five traits: neuroticism, extraversion, agreeableness, conscientiousness, and openness to experience (McCrae & John, 1992). Research has shown that NSSI in students is related to high neuroticism and openness to experience, and to low agreeableness and conscientiousness (Allroggen et al., 2014; Brown, 2009; MacLaren & Best, 2010). Among patients with eating disorders, NSSI is linked to low extraversion (Claes, Vandereycken, & Vertommen, 2004). Furthermore, research using a college sample demonstrated that the number of NSSI methods is positively associated with openness to experience and negatively related to conscientiousness (Robertson, Miskey, Mitchell, & Nelson-Gray, 2013).

NSSI and EMS
Early maladaptive schemas are negative beliefs about the self and world that develop when a child with a predisposition for emotional lability is exposed to negative social environments. EMS represent patterns of rigid unconditional assumptions that are often formed in emotion, difficult to unlearn, and result in dysfunctional stress-coping (Young et al., 2003). According to the schema theory, there are five EMS domains and 15 associated core beliefs. Individuals with Disconnection and Rejection beliefs expect that security, nurturance/love, and acceptance will not be consistently provided by others; this domain includes emotional deprivation, abandonment/instability, mistrust/abuse, social alienation, and defectiveness/shame beliefs. The Impaired Autonomy domain describes an inability to separate oneself from significant others, as well as beliefs that one is incompetent and incapable of functioning independently; beliefs under this domain include failure to achieve, dependence/independence, incompetence, vulnerability to harm, and enmeshment. The Other-Directedness domain describes an excessive focus on the needs and desires of others, often at the expense of one’s own health and well-being; these include subjugation and self-sacrifice beliefs. Individuals with Emotional Inhibition beliefs excessively inhibit spontaneous emotion and communication, or might exhibit unreasonably high standards for themselves; these include emotional inhibition and unrelenting standards beliefs. Finally, the Impaired Limits domain describes poor self-control and low frustration tolerance, as well as beliefs that one is superior and should not be required to follow the same social rules as others; beliefs under this domain include entitlement and insufficient self-control.

NSSI is widely viewed as a means of emotion regulation (Nock & Prinstein, 2004). Research has suggested that negative affect often precedes self-harm behaviours, self-injury is associated with a temporary reduction in aversive emotions, and NSSI acts are performed to reduce negative emotional states (see review by Klonsky, 2007). NSSI is also highly associated with psychological disorders that are marked by emotion dysregulation such as borderline personality disorder, post-traumatic stress disorder, and eating disorders (Claes, Klonsky, Muehlenkamp, Kuppens, & Vandereycken, 2010; Linehan, 1993), giving further credence to the view that self-harm acts can serve as a method of emotion regulation.

Nock (2009) proposed a model delineating the pathway through which self-harm develops and is maintained as a preferred method of stress coping. The author suggested that distal risk factors (e.g., childhood maltreatment, genetic predispositions) lead to individual (e.g., high aversive emotions) and interpersonal (e.g., poor communication skills) vulnerabilities, which then mediate stress responses (e.g., over-arousal) that are attenuated by self-injury. According to this model, individuals with genetic predispositions for heightened emotional and cognitive reactivity, and who experience highly aversive emotions (e.g., anger, dysphoria, shame) and cognitions (e.g., negative thoughts marked by rumination and self-criticism), are at increased risk for repetitive NSSI. The quality of one’s emotional and cognitive content underlie an individual’s expression of personality (Costa & Widiger, 2002) and early maladaptive schemas (EMS; Young, Klosko, & Weishaar, 2003), both of which take shape in early childhood and influence one’s intra- and interpersonal experiences across the lifespan. Scant research has been carried out to investigate these psychological constructs as predictors of NSSI severity.

NSSI and Personality
The dominant model of personality involves five traits: neuroticism, extraversion, agreeableness, conscientiousness, and openness to experience (McCrae & John, 1992). Research has shown that NSSI in students is related to high neuroticism and openness to experience, and to low agreeableness and conscientiousness (Allroggen et al., 2014; Brown, 2009; MacLaren & Best, 2010). Among patients with eating disorders, NSSI is linked to low extraversion (Claes, Vandereycken, & Vertommen, 2004). Furthermore, research using a college sample demonstrated that the number of NSSI methods is positively associated with openness to experience and negatively related to conscientiousness (Robertson, Miskey, Mitchell, & Nelson-Gray, 2013).

NSSI and EMS
Early maladaptive schemas are negative beliefs about the self and world that develop when a child with a predisposition for emotional lability is exposed to negative social environments. EMS represent patterns of rigid unconditional assumptions that are often formed in emotion, difficult to unlearn, and result in dysfunctional stress-coping (Young et al., 2003). According to the schema theory, there are five EMS domains and 15 associated core beliefs. Individuals with Disconnection and Rejection beliefs expect that security, nurturance/love, and acceptance will not be consistently provided by others; this domain includes emotional deprivation, abandonment/instability, mistrust/abuse, social alienation, and defectiveness/shame beliefs. The Impaired Autonomy domain describes an inability to separate oneself from significant others, as well as beliefs that one is incompetent and incapable of functioning independently; beliefs under this domain include failure to achieve, dependence/independence, incompetence, vulnerability to harm, and enmeshment. The Other-Directedness domain describes an excessive focus on the needs and desires of others, often at the expense of one’s own health and well-being; these include subjugation and self-sacrifice beliefs. Individuals with Emotional Inhibition beliefs excessively inhibit spontaneous emotion and communication, or might exhibit unreasonably high standards for themselves; these include emotional inhibition and unrelenting standards beliefs. Finally, the Impaired Limits domain describes poor self-control and low frustration tolerance, as well as beliefs that one is superior and should not be required to follow the same social rules as others; beliefs under this domain include entitlement and insufficient self-control.

NSSI is widely viewed as a means of emotion regulation (Nock & Prinstein, 2004). Research has suggested that negative affect often precedes self-harm behaviours, self-injury is associated with a temporary reduction in aversive emotions, and NSSI acts are performed to reduce negative emotional states (see review by Klonsky, 2007). NSSI is also highly associated with psychological disorders that are marked by emotion dysregulation such as borderline personality disorder, post-traumatic stress disorder, and eating disorders (Claes, Klonsky, Muehlenkamp, Kuppens, & Vandereycken, 2010; Linehan, 1993), giving further credence to the view that self-harm acts can serve as a method of emotion regulation.

Nock (2009) proposed a model delineating the pathway through which self-harm develops and is maintained as a preferred method of stress coping. The author suggested that distal risk factors (e.g., childhood maltreatment, genetic predispositions) lead to individual (e.g., high aversive emotions) and interpersonal (e.g., poor communication skills) vulnerabilities, which then mediate stress responses (e.g., over-arousal) that are attenuated by self-injury. According to this model, individuals with genetic predispositions for heightened emotional and cognitive reactivity, and who experience highly aversive emotions (e.g., anger, dysphoria, shame) and cognitions (e.g., negative thoughts marked by rumination and self-criticism), are at increased risk for repetitive NSSI. The quality of one’s emotional and cognitive content underlie an individual’s expression of personality (Costa & Widiger, 2002) and early maladaptive schemas (EMS; Young, Klosko, & Weishaar, 2003), both of which take shape in early childhood and influence one’s intra- and interpersonal experiences across the lifespan. Scant research has been carried out to investigate these psychological constructs as predictors of NSSI severity.
Research has suggested a link between EMS and NSSI through aversive childhood experiences. For example, EMS have been found to be associated with negative parenting and childhood maltreatment, and to mediate the relation between perceived quality of parenting and psychiatric symptoms (Dale, Power, Kane, Stewart, & Murray, 2010; Wright, Crawford, & Del Castillo, 2009). In a similar vein, NSSI has been found to be associated with higher rates of childhood trauma (Barker, Arsenault, Brendgen, Fontaine, & Maughan, 2008; Fliege, Lee, Grimm, & Klapp, 2009; Gratz, 2003) and poorer family relationships (Di Pierro, Samo, Perego, Gallucci, & Madeddu, 2012; Kaess et al., 2013), with NSSI severity specifically related to negative interactions with childhood caretakers (Di Pierro et al., 2012). This research has suggested that early negative messages about children’s autonomy, worth, and role in the world affect their ability to regulate intrapsychic experiences in adulthood; the more pervasive the beliefs that are developed in response to these interactions, the more likely/often an individual will encounter stimuli that trigger the need for emotion-coping strategies (e.g., self-harm).

Research on Young et al’s (2003) EMS model has suggested that the maladaptive beliefs subsumed under the Disconnection and Rejection domain are most detrimental to one’s well-being. Individuals with such beliefs are likely to experience a persistent inability to escape the source of their distress (e.g., themselves) and/or to expect social exchanges to result in further psychological pain. Disconnection and rejection beliefs are endorsed to a greater degree in clinical populations such as borderline personality disorder and post-traumatic stress disorder (Ball & Cecero, 2001; Cockram, Drummond, & Lee, 2010; Ford & Gómez, 2015) that report disassociative experiences and self-injurious behaviours.

**Present Study**

Overall, limited extant research has suggested that certain personality traits and maladaptive schemas are associated with NSSI behaviour. Most research has been carried out with clinical samples, and less is known about personality and EMS in nonclinical individuals who engage in NSSI. Moreover, there is even less information when it comes to NSSI severity. Examining NSSI correlates among nonclinical samples can help isolate predispositional risk factors (e.g., negative emotionality) from general expressions of psychopathology (e.g., depression), aiding the identification of those generally at higher risk for self-directed violence and informing the urgency/nature of preventative strategies.

The present study compared the personality traits and EMS beliefs in three age- and sex-matched nonclinical groups of varying NSSI severity: high NSSI, low NSSI, and no NSSI control. It was predicted that, compared to the control, both the high and low NSSI groups would show greater neuroticism and openness to experience, and lower agreeableness, conscientiousness, and extraversion (Hypothesis 1), as well as endorse EMS beliefs more strongly (Hypothesis 2). It was also predicted that the high NSSI group would have stronger EMS beliefs that are associated with the Disconnection and Rejection domain than the low NSSI group (Hypothesis 3).

**Method**

**Participants**

The initial sample consisted of 344 individuals recruited from postsecondary educational institutions and the general community in a northern Canadian city. Of these, 142 met the classification criteria for high NSSI, 56 for low NSSI, and 146 for control. After matching on age and sex across groups, the sample was reduced to 156 participants (age \( M = 25.23 \) years, \( SD = 8.14 \) years) for the present study. Most participants were unmarried (90.38%) and White (87.82%). A small number were Aboriginal (6.41%) or a combination of other ethnicities (3.85%). Most had completed high school (56.41%); 24.35% had completed trade school/junior college; and 16.02% had completed university. Table 1 displays the descriptive characteristics by group.

**Group Classification Guidelines**

Using the same classification criteria that was employed by MacLaren and Best (2010), the 156 participants were divided into the high NSSI, low NSSI, or control group of equal size (n = 52; 43 women and 9 men per group) and matched on sex and age. The high NSSI group endorsed 10 or more NSSI incidents and/or three or more methods, whereas the low NSSI endorsed one to nine incidents and/or one to two methods. It is noted that five participants did not provide frequency estimates, and stated that the number of previous NSSI incidents was too high to count; they were placed in the high NSSI group. Information on self-harm history was obtained from the participants’ responses on the Deliberate Self-Harm Inventory (Gratz, 2001).
The group classification criteria guidelines are informed by research and are clinically meaningful: between 52.6% and 96.5% of borderline personality disorder patients reported 10 or more NSSI episodes, and between 38.1% and 76.7% of borderline personality disorder patients reported using three or more methods (Zanarini et al., 2006). Also, MacLaren and Best (2010) found more severe psychopathology in the high NSSI group than in the low NSSI group.

**Procedure**

The appropriate research ethics approval was received from the Lakehead University Research Ethics Board prior to commencement of the study. Participants were recruited through community fliers and online bulletins, which invited individuals to complete a confidential and anonymous survey.

Table 1

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>High NSSI</th>
<th>Low NSSI</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (M (SD) years)</td>
<td>24.65 (7.52)</td>
<td>24.79 (8.00)</td>
<td>24.35 (6.42)</td>
</tr>
<tr>
<td>Sex (%)</td>
<td>9 (17.31)</td>
<td>9 (17.31)</td>
<td>9 (17.31)</td>
</tr>
<tr>
<td>Man</td>
<td>43 (82.69)</td>
<td>43 (82.69)</td>
<td>43 (82.69)</td>
</tr>
<tr>
<td>Ethnicity (%)</td>
<td>4 (7.69)</td>
<td>5 (9.62)</td>
<td>6 (11.54)</td>
</tr>
<tr>
<td>Marital status (%)</td>
<td>13 (74.19)</td>
<td>13 (74.19)</td>
<td>13 (74.19)</td>
</tr>
<tr>
<td>Highest education received (%)</td>
<td>0 (0.00)</td>
<td>0 (0.00)</td>
<td>0 (0.00)</td>
</tr>
<tr>
<td>College or trade school</td>
<td>14 (26.92)</td>
<td>13 (25.00)</td>
<td>11 (21.15)</td>
</tr>
<tr>
<td>Undergraduate degree</td>
<td>5 (9.62)</td>
<td>11 (21.15)</td>
<td>9 (17.31)</td>
</tr>
<tr>
<td>Graduate school</td>
<td>0 (0.00)</td>
<td>2 (3.85)</td>
<td>1 (1.92)</td>
</tr>
</tbody>
</table>

Note. Within-group n = 52, Total sample N = 156. *Self-identified as more than one ethnicity.

TABLE 1

Sample Characteristics by Group

The group classification criteria guidelines are informed by research and are clinically meaningful: between 52.6% and 96.5% of borderline personality disorder patients reported 10 or more NSSI episodes, and between 38.1% and 76.7% of borderline personality disorder patients reported using three or more methods (Zanarini et al., 2006). Also, MacLaren and Best (2010) found more severe psychopathology in the high NSSI group than in the low NSSI group.

The recruitment information also noted that the investigators were seeking both individuals who hurt themselves and those who do not. The survey was an online research questionnaire consisting of the Deliberate Self-Harm Inventory, the Early Maladaptive Schema Questionnaire-Short Form (EMSQ-SF). At the end of the study, participants were provided with a list of mental health resources in the community and, in the event of mental health concerns, had access to the second author who is a registered clinical psychologist.

**Measures**

**Deliberate Self-Harm Inventory** (DSHI; Gratz, 2001). The DSHI measures the practice and severity of NSSI. It lists 16 acts of self-harm plus an open-ended option for acts not included in the list. For each act, participants indicate whether they have ever engaged in that method, the age at which the act was first undertaken, the number of times participants have used that method of self-harm, the most recent time it was employed, the number of years participants have engaged in that act, and whether hospitalization or medical attention was ever required as a result of the act. For the purpose of the present study, the DSHI was used to identify the total number of methods used and the total frequency of self-harm incidents in the participants’ lifetime. This measure has high internal consistency (α = .82) and adequate test-retest reliability over a 2 to 4-week period (r = .68, p < .001), as well as good convergent validity with other measures of self-harm and positively predicts such psychopathological symptoms as stress, depression, anxiety, and hostility (Fliede et al., 2006).

**Big Five Inventory** (BFI; John, Donahue, & Kentle, 1991). The 44-item BFI measures personality traits of neuroticism, extraversion, agreeableness, conscientiousness, and openness to experience. Each item is rated on a 5-point Likert scale from 1 (disagree strongly) to 5 (agree strongly). Ratings on items that load on each of the five personality factors are added and averaged to obtain a mean score on each factor. Higher scores on each factor indicate greater endorsement of a particular trait. The BFI has high internal consistency, with coefficient alphas ranging between .79 and .87 for the five factors (John, Naumann, & Soto, 2008), an average test-retest stability of .90 (Hampson & Goldberg, 2006), and good convergent validity with...
other measures of personality (John et al., 2008).

Early Maladaptive Schema Questionnaire-Short Form (EMSQ-SF; Young & Brown, 2005). The EMSQ-SF consists of 75 items that load on five EMS domains and 15 associated core beliefs. Each item loads on a 6-point Likert-type scale from 1 (completely untrue) to 6 (describes me perfectly). Scores for items that load on a particular belief are added and averaged to yield a mean score, with higher scores indicating stronger endorsement of that belief. The EMSQ-SF has excellent internal consistency (Cronbach’s α have been reported to range between .76 and .93 for the 15 subscales of the EMSQ-SF) and good external validity (Welburn, Coristine, Dagg, Pontefract, & Jordan, 2002), and has been established as a good predictor of different psychopathological symptoms and disorders (Ball & Cecero, 2001; Cockram et al., 2010; Ford & Gómez, 2015; Welburn et al., 2002).

Initial Data Analysis

Prior to analysis, the data was screened for within-group univariate and multivariate outliers to determine whether a participant’s response on one or more dependent variables would exert undue influence on subsequent analyses. Univariate outliers were defined as cases with z scores greater than ±3, and multivariate outliers as cases with significant Mahalanobis distance. No univariate or multivariate outliers were found.

Normality was examined by transforming the skewness and kurtosis values associated with the distribution of each of the dependent variables into z scores, using a conservative alpha level of .001 to determine significance. The assumption of linearity was also investigated by examining the bivariate plots associated with the BFI factors and EMSQ-SF core beliefs. Violations of normality and linearity were found only for the EMSQ-SF subscales. Problems with kurtosis, positive skewness, and linearity were found primarily within the control group, becoming less evident within the low NSSI and least within the high NSSI group. Transformation of the EMSQ-SF data was not carried out for the following reasons. It was not unreasonable to expect that the groups would tend toward lower scores of the EMSQ-SF (a scale which reflects underlying psychopathology) given the nonclinical nature of the sample. Further, transforming data can limit or obfuscate the interpretation of findings, and given that cell sizes in the present study were equal and >20, and no multivariate outliers were identified, the analyses were considered robust to violations of the assumptions of normality (Tabachnick & Fidell, 2015).

Description of NSSI Behaviour

The high NSSI group (M = 155.44, SD = 416.81, median = 47) reported significantly more self-harm incidents than the low NSSI group (M = 2.81, SD = 2.24, median = 2), t(46) = 2.51, p = .016; d = .52. (It should be noted that five participants in the high NSSI group did not provide frequency estimates, stating that the number of previous NSSI incidents was too high to count; subsequently, these participants were excluded from frequency analyses.) The high NSSI group also used more self-harm methods (M = 4.21, SD = 1.85) than the low NSSI group (M = 1.35, SD = 0.48), t(57) = 10.81, p < .001; d = 2.12. Compared to the low NSSI group, the high NSSI group started self-harming at a younger age (M = 11.54, SD = 3.76 versus M = 14.41, SD = 3.41), t(97) = 4.02, p < .001; d = .80, engaged in self-harm over a longer period of time (M = 8.93 years, SD = 6.85 versus M = 0.90 years, SD = 1.55), t(54) = 8.08, p < .001; d = 1.62, and were more likely to have harmed themselves within the previous year, χ²(1, N = 104) = 20.07, p < .001 (57.69% versus 15.38%). Compared to the low NSSI group, the high NSSI group was also significantly more likely to endorse most of the NSSI methods, with the largest differences demonstrated for cutting, χ²(1, N = 104) = 20.54, p < .001 (76.92% versus 32.69%), sticking the body with pins or sharp objects, χ²(1, N = 104) = 22.78, p < .001 (55.76% versus 11.54%), scratching, χ²(1, N = 104) = 18.09, p < .001 (57.69% versus 17.31%), and preventing wounds from healing, χ²(1, N = 104) = 15.82, p < .001 (30.77 % versus 1.92%).

Main Analyses

Table 2 shows the Cronbach’s α for the BFI and EMSQ-SF belief subscales, all of which exceeded .70 internal consistency indices and ranged from adequate to excellent. The exception was the EMSQ-SF entitlement belief subscale (α = .59) that showed low internal consistency. Table 2 also provides within-group means and standard deviations, and pairwise group means comparisons with effect size (Cohen’s d). The results are explicated below.

Personality Traits. A one-way Multivariate Analysis of Variance (MANOVA) with group as the independent variable and the five BFI personality traits as dependent variables showed a significant omnibus effect, Wilk’s λ = .67, F(10, 298) = 6.27,
Personality, EMS, and NSSI

Conscientiousness. The high NSSI group was also significantly less Agreeable than the low NSSI group (see Table 2).

EMS. A one-way MANOVA with group as the independent variable and the five EMSQ-SF domains as the dependent variables showed a significant omnibus effect, Wilk’s $\lambda = .66$, $F(10, 298) = 7.02$, $p < .001$, $\eta^2 = 0.19$. Separate follow-up ANOVAs were performed on each of the 15 EMSQ-SF subscales and interpreted at $\alpha = .003$ so that the overall error rate was kept at .045. Significant group differences were obtained for 14 of the 15 EMSQ-SF factors including Emotional Deprivation, $F(2, 153) = 10.58$, $p < .001$, $\eta^2 = 0.12$, Abandonment,
Discussion

Personality

Significant differences in personality traits were found between individuals who reported a history of NSSI and those who did not. As predicted in Hypothesis 1, both high and low NSSI groups reported greater neuroticism and less agreeableness, conscientiousness, and extraversion than the control group. However, no group differences were observed for openness to experience. Previous works have identified similar personality profiles in NSSI samples (Brown, 2009; MacLaren & Best, 2010), which have been linked to an increased openness to experience. The one prediction that was not supported in Hypothesis 1 was that an NSSI history would be linked to increased openness to experience. Several studies that employed a similar sample (e.g., students from a North American university) have reported a relation between NSSI and Openness to Experience (Brown, 2009; MacLaren & Best, 2010; Robertson et al., 2013); however, there were several methodological and sample differences that might have contributed to the lack of significant findings observed in the present study. Previous works used a group format to administer questionnaires, whereas the present study collected responses anonymously online. The increased ease and accessibility of participation could have attracted a broader demographic of students who were generally more receptive to and interested in new experiences, but who otherwise would not have been able to attend an in-person survey. Indeed, the present study obtained mean scores on Openness that were moderately high across the three groups; perhaps the study format encouraged a self-selection bias for individuals who are higher in openness to experience.

As well, sex and age have been found to modulate expressions of openness to experience (Costa, Terracciano, & McCrae, 2001; McCrae et al., 1999) and endorsement of self-injurious behaviour (Hawton & Harriss, 2008). The mean age of participants in the present study was approximately 5 years older than those reported in previous works, and other studies did not match groups on age or sex. It is possible that, in using a sample with a higher mean age and controlling for the effects of sex and age on personality expression, any variance that was explained by the Openness trait was eliminated.

The findings in the present study also showed that individuals in the high NSSI group were significantly less agreeable than those in the low NSSI group. Previous investigations have shown that more severe NSSI behaviours are associated with a greater number of adverse life events (Liu et al., 2014), and that the experience of more uncontrollable life stressors is related to decreased expressions of agreeableness over an 8-year period (Löckenhoff, Terracciano, Patriciu, Eaton, & Costa, 2009). It is possible that the lower agreeableness scores within the high NSSI group reflect a history likely increase one’s risk for stress exposure while simultaneously decreasing the ability to cope with stressors, which is congruent with the view that NSSI represents a maladaptive form of emotion regulation in response to stress (Nock & Prinstein, 2004).

Post-hoc LSD tests were carried out on the 14 significant EMSQ-SF subscales. The high and low NSSI groups scored higher than the control on each EMSQ-SF subscale, except for Unrelenting Standards beliefs. Compared to the low NSSI and control groups, the high NSSI group scored higher on Unrelenting Standards. The high NSSI group also scored significantly higher than the low NSSI group on Social Alienation, Defectiveness/Shame, Vulnerability to Harm, and Enmeshment (see Table 2).
Personality, EMS, and NSSI

important predictor of NSSI severity and suicidal
might be an ideation (O’Connor et al., 2007). It is possible that
has been found to predict increased suicidal
NSSI and high socially prescribed perfectionism
More importantly, the combination of repetitive
2009; O’Connor, Rasmussen, & Hawton, 2010).
beliefs, when compared to the control. The robust presence of EMS in the NSSI groups is
not surprising given that EMS has been linked to psychological difficulties (Dale et al., 2010; Wright et al., 2009). The factors that contribute to the
development of EMS include childhood trauma or dysfunctional family environments (Dale et al., 2010; Wright et al., 2009), and these in turn have been linked to NSSI (Barker et al., 2008; Fliege et al., 2009; Kaess et al., 2013). It is not clear whether there is a link between NSSI and adverse childhood experiences in the present study because the latter was not assessed.

In support of Hypothesis 3, it was found that the high NSSI group scored higher than the low
NSSI group on EMS beliefs related to the Disconnection and Rejection domain (social alienation, deftiveness/shame). The high NSSI group also scored higher on beliefs related to Impaired Autonomy (vulnerability to harm, enmeshment) and Emotional Inhibition (unrelenting standards). These beliefs are significantly related to psychiatric symptoms such as depression and anxiety that are highly associated with self-injury (Klonsky, Oltmanns, & Turkheimer, 2003; Nock & Kessler, 2006; Welburn et al., 2002). Furthermore, beliefs relating to social alienation, deftiveness/shame, and vulnerability to harm all positively predict risk of suicidal behaviour (Dale et al., 2010), which is highly related to repetitive NSSI (Anestis, Khazem, & Law, 2015).

Interestingly, it was also found that the high NSSI group could be differentiated from both the low NSSI and control groups on the basis of its stronger endorsement of unrelenting standards beliefs. Considerable research has linked perfectionism to suicide (see Flett, Hewitt, & Heisel, 2014) and to NSSI (Hoff & Meuhlenkamp, 2009; O’Connor, Rasmussen, & Hawton, 2010). More importantly, the combination of repetitive NSSI and high socially prescribed perfectionism has been found to predict increased suicidal ideation (O’Connor et al., 2007). It is possible that EMS unrelenting standards beliefs might be an important predictor of NSSI severity and suicidal

of more negative life events. However, this cannot be confirmed because negative life events were not assessed in the present sample.

EMS

In line with Hypothesis 2, both the high and low NSSI groups indicated greater endorsement of all EMS beliefs, except for self-sacrifice and unrelenting standards, when compared to the control. The present study showed that individuals who engaged in NSSI exhibited a personality profile that is often associated with psychological difficulties (i.e., higher levels of neuroticism and lower levels of extraversion, agreeableness, and conscientiousness), and generally endorsed more severe levels of EMS beliefs compared to those who do not engage in self-harm behaviours. As well, individuals with more severe NSSI were less agreeable than those with less severe NSSI, and they also more strongly endorsed a number of EMS beliefs within the domains of Disconnection and Rejection, Impaired Autonomy, and Emotional Inhibition. These findings support Nock’s (2009) model of the development and maintenance of NSSI, whereby highly aversive cognitions and emotions, poor social skills and perceived support, and increased exposure to perceived stressors, are suggested to increase one’s risk for repetitive self-harm. Indeed, the personality and EMS profiles linked to NSSI severity in the present study predispose individuals to increased interpersonal conflict, social isolation, and psychological distress, also increasing the frequency with which these persons would need to engage stress-coping—an association that was corroborated by the finding that those in the high NSSI group reported self-harming over a longer span of time and were more likely to have self-injured in the previous year. Finally, the high NSSI group endorsed a greater degree of unrelenting standards compared to the low and no NSSI groups.

Limitations and Future Directions

The present findings have to be viewed with several caveats in mind. There were more women than men in the study. Sex differences in both NSSI methods and motivations have been reported. Men are more likely to self-harm under the influence of drugs or alcohol (Madge et al., 2008), utilize more public and violent NSSI methods (e.g., breaking bones, self-harming while in the presence of peers), endorse more motivations related to influencing others and testing one’s will/strength, and are at a greater risk for eventually committing suicide. Compared to men, women report more incidents
of cutting and ingesting pills, endorse more despair and self-punitive motivations, and are more likely to self-harm while alone (Hawton & Harriss, 2008; Laye-Gindhu & Schonert-Reichl, 2005; Taylor, 2003; Zahl & Hawton, 2004). It has been proposed that sex differences in self-injurious behaviour are related to differences in sex-role socialization, whereby men are encouraged to seek autonomy, explore new experiences, and display their social dominance, while women are often socialized to suppress negative emotions and depend on others for security (Brody, 2000; Gilbert, 1987; Laye-Gindhu & Schonert-Reichl, 2005). These socialization factors are also thought to be related to the externalizing (e.g., anger and antisocial behaviour) versus internalizing (e.g., anxiety and depression) symptoms that are associated with sex differences in the expressions of psychopathology (Rosenfield, 2000). Given the significantly higher number of women in the present study, it is possible that the personality and EMS profiles that were linked to NSSI severity more aptly describe traits predictive of self-harm history and severity in women, particularly in relation to EMS—a construct whose theoretical foundation is based on the nature of childhood interactions with caregivers (Young et al., 2003).

An algorithm relying on the number of incidents and the number of methods was used to establish severity of NSSI. Whether or not employing another system to assess NSSI severity might produce different results remains to be seen. It could be valuable to compare individuals who have self-injured once or twice with chronic self-harmers; perhaps those who briefly experimented with NSSI and then abandoned its use are dispositionally different from those who use it over several years or on a daily basis. It is also noted in the present study that those in the high NSSI group were more likely to have engaged in self-harm within the previous year. This raises the question of whether recency of self-harm acts might act as a confound, such that recency instead of severity of self-injury might explain the group differences in personality and EMS. Future research could examine how NSSI recency relates to personality and cognitive schemas. Investigating resilience factors (e.g., behavioural and emotional regulation strategies, personality traits, cognitive schemas and beliefs, availability of support resources) that differentiate occasional from chronic self-harmers, or that predict cessation of self-harming behaviours, might also be helpful to guide clinical services.

The observation that unrelenting standards were linked to more severe NSSI warrants replication, with particular attention to whether these beliefs are related to increased risk for suicidal behaviour among persons who exhibit severe NSSI. Despite increasing research interest in the link between NSSI and suicidal behaviour (see Hamza, Stewart, Willoughby, 2012), more needs to be done to identify the pathways and associated risk factors. More work could also be carried out to learn about the role of openness to experience in self-harm behaviours and ideation.

Finally, the internal consistency for the EMSQ-SF entitlement subscale was low (Cronbach’s α = .59), which limits the validity of the findings associated with that subscale. The items on the entitlement factor are broadly characterized by one of two concepts: (a) the unwillingness to be constrained by or accept “no” as an answer from other people (e.g., “I hate to be constrained or kept from doing what I want”), and (b) the belief that one is special and not subject to the same rules as others (e.g., “I feel that what I have to offer is of greater value than the contributions of others”). It is not known how well either or both of the constructs characterize the different groups or the overall sample in the present study.

**Implications**

The results point to the importance of examining EMS beliefs in treatments for NSSI. Evidence supports the use of behavioural and cognitive interventions such as dialectical behaviour therapy (DBT) and cognitive therapy to reduce self-injurious behaviours (Stanley, Fineran, & Brodsky, 2014) and improve emotion regulation (Koerner, 2012). These approaches address thought content, among other intrapersonal experiences, to identify maladaptive internal processes such as negative thoughts and judgmental/rigid beliefs that otherwise heighten and prolong exposure to aversive emotions. DBT utilizes mindfulness skills to help individuals observe and articulate their experiences, aversive or otherwise, in a nonjudgmental manner. DBT also teaches persons to engage in dialectics so that they might synthesize seeming contradictions to adopt a more fluid and holistic understanding of personal experiences (e.g., accepting oneself as one is while simultaneously recognizing the need for positive change; Linehan, 1993). DBT has been found to effectively reduce NSSI frequency and suicidal behaviours (Turner, Austin, & Chapman, 2014), which might (in part) be related to a reduction in unrelenting standards (linked in the present study...
to more severe NSSI) and perfectionism (linked previously to suicidal thoughts and behaviours; Baumeister, 1990; Chatard & Selimbegović, 2011; Tang et al., 2013)—constructs that are marked by judgmental self-statements, unrealistic expectations for oneself, and rigid/unidimensional beliefs.

Findings from the present study suggest that maladaptive beliefs related to a sense of being disconnected from others, defectiveness and shame, vulnerability to harm, and/or enmeshment could also be the focus of cognitive interventions in individuals who struggle with self-harm. Generally assessing an individual’s EMS could help both the clinician and the client identify situations in which a particular maladaptive belief might be triggered (e.g., defectiveness and shame beliefs might be activated when in a performance). The relation between such core beliefs and emotional responses, as well as how to better regulate negative emotions when core schemas are triggered by schema-relevant stimuli. The benefits of Young’s (2003) schema-focused therapy in reducing self-injurious behaviours have yet to be studied, though the results of the present study suggest that such investigations would have merit.

Understanding how schema and personality profiles might predispose an individual to self-directed harm would also inform preventative strategies in nonclinical settings. Informed individuals, whether counsellors, teachers, or parents, might better recognize persons at higher risk for NSSI—persons who not only express negative emotionality and low social support, but who hold specific negative beliefs that might be triggered during schema-relevant stressors. Recognizing periods of increased vulnerability for self-harm could motivate and direct efforts to intervene, preventing at risk individuals from trying NSSI to cope with stress. Continuing to explore the nuances of NSSI behaviour can elucidate underlying processes that trigger and reinforce this dysfunctional coping strategies, as well as contribute to the development of effective interventions with persons who struggle with the urge to hurt themselves.

References


http://dx.doi.org/10.1016/j.jpsychores.2005.10.006

http://dx.doi.org/10.1016/j.jpsychores.2008.10.013

http://dx.doi.org/10.1176/appi.ajp.160.8.1501


http://dx.doi.org/10.1371/journal.pone.0110907

http://dx.doi.org/10.1192/bjp.191.5.458b

http://dx.doi.org/10.1007/s10964-005-7262-2


http://dx.doi.org/10.1080/138111118.2013.824835

http://dx.doi.org/10.1002/jts.20385

http://dx.doi.org/10.1155/2011/935871

http://dx.doi.org/10.1016/j.paid.2010.05.019

http://dx.doi.org/10.1111/j.1469-7610.2008.01879.x

http://dx.doi.org/10.1037/0012-1649.35.2.466

http://dx.doi.org/10.1111/j.1467-6494.1992.tb00970.x

http://dx.doi.org/10.1111/j.1467-1877.2009.01613.x

http://dx.doi.org/10.1016/j.psychres.2006.05.010

http://dx.doi.org/10.1037/0021-843X.115.3.616

http://dx.doi.org/10.1037/0022-006X.72.5.885

http://dx.doi.org/10.1016/j.brat.2009.09.008

perfectionism and positive future thinking. *Behaviour Research and Therapy, 45*, 1543–1555.  
http://dx.doi.org/10.1016/j.brat.2006.11.006

http://dx.doi.org/10.1080/13811118.2013.748410


http://dx.doi.org/10.1068/j.1743-6861.2003.00042.x


http://dx.doi.org/10.1542/peds.2005-2543


http://dx.doi.org/10.1016/j.chiabu.2008.12.007


http://dx.doi.org/10.1192/bjp.185.1.70

http://dx.doi.org/10.1521/pedi.2006.20.1.9

**Author Note.** Sarah D. Arthurs, Department of Psychology, Lakehead University; Josephine C. H. Tan, Department of Psychology, Lakehead University.

Data for this article was retrieved from a larger database for an MA thesis project carried out by the first author under the supervision of the second author. Special thanks to all Psi Chi Journal reviewers for their support.

Correspondence concerning this article should be addressed to Sarah D. Arthurs, Department of Psychology, Lakehead University, Thunder Bay, Ontario P7B 5E1, Canada. E-mail: sarthurs@lakeheadu.ca
Obsessive-compulsive disorder (OCD) is characterized in the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5; American Psychiatric Association, 2013) by persistent, unwanted thoughts (obsessions) and/or repetitive behaviors (compulsions) performed to diminish feelings of anxiety or distress. Although the specific content of obsessions and compulsions can vary across individuals, several replicable symptom dimensions have been identified including cleaning (e.g., contamination, obsessions with cleaning compulsions), symmetry (e.g., obsessions about symmetry with counting, repeating, and ordering compulsions), and taboo thoughts (e.g., aggressive or sexual obsessions and compulsions relating to acts viewed as “forbidden” or “wrong”; Mataix-Cols, Rosario-Campos, & Leckman, 2005). The disorder was reclassified as an obsessive-compulsive and related disorder (OCRD) in the newest edition of the DSM, alongside hoarding disorder, body dysmorphic disorder, trichotillomania (hair-pulling disorder), and excoriatio (skin-picking disorder).

Research efforts examining the underlying mechanisms of OCD have revealed two distinct motivational domains for compulsions: harm avoidance and incompleteness (Summerfeldt, 2004). Compulsions motivated by harm avoidance are typically enacted upon to reduce anxiety associated with safety concerns or feared consequences of anticipated and potentially harmful events (Salkovskis, 1991). For example, individuals with OCD may check their stoves multiple times in order to prevent the start of a fire or wash their hands repeatedly to ensure they do not contract physical illness.

However, compulsions driven by a sense of incompleteness are intended to prevent or alleviate uncomfortable internal sensations associated with one’s immediate environment feeling flawed or imperfect in some way (Rasmussen & Eisen, 1992). This “not just right experience” (NJRE), which is defined in the context of this article as a discomfort experienced when there is a discrepancy between one’s desired and current sensory state, can be experienced across a range of sensory phenomena including sight (“this doesn’t look right”), touch...
Not Just Right Experiences in fact be a transdiagnostic risk factor underlying NJREs (Fergus, 2014). This suggests that NJREs may act as a motivational core dimension of OCD symptoms remained significant predictors of regression analyses, both generalized anxiety and anxiety, social anxiety, and depressive symptoms. In correlations between NJREs and OCD, generalized as well as Tourette’s syndrome (Neal & Cavanna, 2010), previous work has also demonstrated symptomatology (Ghisi, Chiri, Marchetti, Sanavio, shown to correlate highly with obsessive-compulsive the self-reported experience of an NJRE has been that could be seen as a part of depersonalization, derealization, impaired psychological mindedness, and a range of psychiatric disorders. Although the self-reported experience of an NJRE has been shown to correlate highly with obsessive-compulsive symptomatology (Ghisi, Chiri, Marchetti, Sanavio, & Sica, 2010), previous work has also demonstrated its associations with a range of emotional disorders as well as Tourette’s syndrome (Neal & Cavanna, 2013). Fergus (2014) recently reported significant correlations between NJREs and OCD, generalized anxiety, social anxiety, and depressive symptoms. In regression analyses, both generalized anxiety and OCD symptoms remained significant predictors of NJREs (Fergus, 2014). This suggests that NJREs may in fact be a transdiagnostic risk factor underlying several psychiatric disorders. The transdiagnostic conceptualization and treatment of emotional disorders, which has gained increased attention in the empirical literature (Mansell, Harvey, Watkins, & Shafran, 2008), suggests more similarities than distinctions between emotional disorders in regard to risk factors, and calls for future research to enhance understanding of their shared aspects. Focusing on processes shared across emotional disorders has the potential to lead to development of efficient treatments for multiple disorders simultaneously.

NJREs and Obsessive-Compulsive and Related Disorders
As indicated earlier, NJREs represent a mismatch between perception of an actual, current state and a desired one, with the desired state being one of certainty and perfection. Moreover, it has been noted that NJREs may characterize emotional disorders beyond OCD (Fergus, 2014). However, previous research has not examined the association between NJREs and other disorders in the OCRD category. NJREs have been posited to represent a motivational domain underlying OCD symptoms, and have also been preliminarily shown to be transdiagnostic in nature. If the NJRE sensation is indeed a shared risk factor, examining its potential role in other disorders with which OCD shares a diagnostic category may increase understanding of the shared features of these disorders.

Although the base of literature examining NJREs and OCRDs is presently lacking, there are some points that may indicate the crossdiagnostic presence of NJREs beyond OCD. For example, NJREs may act as a motivational core dimension of specific facets of hoarding disorder such as acquiring (e.g., “I don’t think I would feel ‘just right’ or ‘complete’ without this object”) and saving behavior (e.g., “I don’t feel ‘right’/this space would be ‘incomplete’ by throwing this object away”; Ecker, Kupfer, & Gönner, 2014; Frost, Steketee, Tolin, Sinopoli, & Ruby, 2014). Additionally, individuals diagnosed with body dysmorphic disorder have reported greater levels of perfectionism than control samples (Hartmann, Thomas, Greenberg, Matheny, & Wilhelm, 2014) and have demonstrated appearance-related symmetrical concerns (Hart & Phillips, 2013; Reese, McNally, & Wilhelm, 2010) akin to OCD-related symmetry symptomology. A recent study found that individuals diagnosed with body dysmorphic disorder reported more severe NJREs compared to control participants (Summers, Matheny, & Cougle, 2017). This suggests the
potential presence of an underlying transdiagnostic mechanism related to NJREs.

Furthermore, despite the lack of research regarding the immediate relationship between NJREs and psychophysiological disorders such as hair pulling or skin picking, individuals diagnosed with these disorders may experience NJREs in a similar way. Research has shown that emphasis on control and perfectionism were among the superordinate themes of cognitions/beliefs that contribute to both the onset and maintenance of trichotillomania in a clinical hair-pulling sample (Rehm, Nedeljokovic, Thomas, & Moulding, 2015). In addition, both control and perfectionism have been strongly related to NJREs such that individuals may not feel right without the presence of pulling or picking sensations.

Although little research exists linking NJREs to skin picking and hair pulling (Summerfeldt, 2004), evidence has suggested that skin picking and hair pulling have substantial similarities in a variety of etiological clinical characteristics such as symptom presentation and course of illness (Stein et al., 2010). It may be possible that individuals who display symptoms of, or are diagnosed with, skin-picking disorder may engage in such behaviors similar to those with hair-pulling disorder, in part to achieve a just-right state. We believe the similarities across OCRDs (e.g., recurrent and intrusive thoughts and compulsive, motivated action to alleviate uncomfortable sensations) suggest that a not-just-right feeling acts as a core motivator of symptoms. Therefore, the association between NJREs and OCRDs merits further investigation.

The current study aimed to address this question by providing a preliminary descriptive analysis of the relationship between NJREs and symptoms of the disorders classified in the OCRD category. Although exploratory in nature, this study represents an important first step in elucidating the role of NJREs in OCRDs; an association between NJREs and symptoms of each of the OCRDs would lend support to their coclassification, increase understanding of the shared features of the disorders in the OCRD diagnostic class, and have implications for conceptualization and treatment of these disorders.

In light of past research that has demonstrated correlations between NJREs and symptoms of several emotional disorders beyond OCD (Fergus, 2014), as well as the similarities between NJRE symptom presentation and the underlying factors of a variety of OCRDs, we hypothesized that the experience of NJREs would be associated with symptoms of the other disorders currently classified in the DSM-5 as OCRDs. Although we expected OCD symptoms to account for more variance in total NJRE scores than other OCRD symptomology, we nonetheless hypothesized that OCRD symptoms would demonstrate a significant positive association with the self-reported experience of NJREs.

Method

Participants

Three hundred thirty-three individuals recruited via Amazon’s Mechanical Turk consented to participate and completed all study measures. Fourteen individuals were excluded from data analysis due to invalid responding: three were excluded due to the attention check questions being answered incorrectly and 11 more were dropped due to responding in far less time (less than 5 minutes) than it would have taken to validly respond to the items (an estimated 15–20 minutes); thus, we report on data from 319 participants.

Participants were required to be at least 18 years old, residing in the United States, and fluent in English. The sample consisted of 163 (51%) women and 155 men (49%; 1 otherwise specified: female to male transsexual) who were on average 35.60 years old ($SD = 11.99$). Of these, 79.6% identified as European American or White, 8.5% African American or Black, 7.2% Asian, 1.6% American Indian or Alaskan Native, and 2.5% as multiracial. Two individuals (0.6%) preferred not to indicate race. Additionally, 7.8% of all participants identified their ethnicity as Hispanic. Finally, 21.9% of participants reported having a current psychiatric disorder diagnosis, and 34.8% reported ever being diagnosed with a psychiatric disorder.

Procedures

Prior to the study, approval was given by the Institutional Review Board at Knox College. Participants were recruited from Amazon Mechanical Turk, a crowdsourcing venue hosted via Amazon.com in which “workers” (i.e., study participants) can sign up to complete “HITS” (i.e., tasks) for payment. Workers read descriptions of the task including what is involved, how much time it will take, and how much compensation will be provided, prior to deciding to accept participation in a HIT. Mechanical Turk has been particularly effective at helping researchers find participants with specific psychiatric symptoms, risk factors, or rare demographic characteristics that may be crucial for research in
the social, behavioral, and clinical sciences. Furthermore, Mechanical Turk has been shown to offer an even more representative sample than other popular samples utilized in psychological research such as college students (Shapiro, Chandler, & Mueller, 2013). Mechanical Turk has also been used in previous research on NJREs and emotional disorders (Fergus, 2014).

Upon indicating consent, participants completed a battery of questionnaires assessing NJREs and symptoms of OCD, body dysmorphic disorder, hoarding disorder, hair-pulling disorder, and skin-picking disorder. Several attention check items (e.g., “Please select Option 4”) were included to assess response validity.

**Measures**

**Dimensional Obsessive-Compulsive Scale (DOCS; Abramowitz et al., 2010).** The DOCS is a 20-item self-report questionnaire that assesses four factor analytically derived OCD symptom dimensions: Contamination (e.g., “About how much time have you spent each day thinking about contamination and engaging in washing or cleaning behaviors because of contamination?”), Responsibility (e.g., “About how much time have you spent each day thinking about the possibility of harm or disasters and engaging in checking or efforts to get reassurance that such things do not or did not occur?”), Unacceptable Thoughts (e.g., “About how much time have you spent each day with unwanted unpleasant thoughts and with behavioral or mental actions to deal with them?”), and Symmetry (e.g., “About how much time have you spent each day with unwanted thoughts about symmetry, order, or balance and with behaviors intended to achieve symmetry, order, or balance?”). Each dimension is assessed by five items assessing time spent, avoidance behaviors, distress, interference, and difficulty disregarding thoughts about that dimension. The DOCS total score has demonstrated excellent internal consistency ($\alpha = .90$), convergent validity, and test-retest reliability, and all factor subscales were strongly correlated with the total score (Abramowitz et al., 2010).

In the present study, internal consistency reliability estimates ranged from .89 to .94, with the alpha coefficient of the total scale being .95. A cutoff score of 18 was found to effectively distinguish between clinical OCD patients and nonclinical adults (Abramowitz et al., 2010); in the current study, 97 participants (31% of the sample) scored 18 or more.

**Savings Inventory-Revised (SI-R; Frost, Steketee, & Grisham, 2004).** The SI-R is a 23-item self-report measure that assesses the cognitive and behavioral dimensions of hoarding. It contains three factor analytically derived subscales: Clutter (9 items; e.g., “To what extent does the clutter in your home cause you distress?”), Difficulty Discarding (7 items; e.g., “To what extent do you have difficulty throwing things away?”), and Excessive Acquisition (7 items; e.g., “How much control do you have over your urges to acquire possessions?”).

The SI-R has been validated in both clinical and nonclinical samples, and its total and subscale scores demonstrated good internal consistency, test-retest reliability, and convergent and discriminant validity (Frost et al., 2004). In the current study, the SI-R demonstrated excellent internal consistency, $\alpha = .94$. Scores of 41 and above on the SI-R have been used to identify clinically significant hoarding behavior (Steketee, Frost, Tolin, Rasmussen, & Brown, 2010). In the current sample, 59 participants (18% of the sample) scored 41 or above.

**Milwaukee Inventory for the Dimensions of Adult Skin Picking (MIDAS; Walther, Flessner, Conelea, & Woods, 2008).** The MIDAS is a 12-item self-report measure of skin-picking behavior. The MIDAS consists of two distinct subscales: Focused (6 items) and Automatic (6 items) Picking. Focused picking involves awareness of picking behaviors (e.g., “I pick my skin when I am anxious or upset”), whereas automatic picking refers to subconscious picking that takes place outside awareness (e.g., “I am not usually aware of picking my skin during a picking episode”). The MIDAS yields a distinct score for each of the two subscales; focused and automatic picking have been shown to be unique from one another (intersubscale $r = .008$, $p = .94$; Walther et al., 2008). The MIDAS Focused Scale demonstrated adequate internal consistency ($\alpha = .95$) in the present study, however the Automatic scale yielded an unacceptable reliability value ($\alpha = .22$).

**Milwaukee Inventory for Subtypes of Trichotillomania (MIST; Flessner et al., 2008).** The MIST is a 15-item self-report measure that assesses two forms of hair-pulling behavior via factor analytically derived subscales: Focused Pulling (10 items; e.g., “I intentionally start pulling my hair”) and Automatic Pulling (5 items; e.g., “I am usually not aware of pulling my hair during a pulling episode”), each scored on a 10-point Likert-type scale from 0 (not true of any of my hair pulling) to 9 (true for all of my hair pulling). Like the MIDAS, the MIST is scored
as two symptom subscales. Both scales have exhibited adequate convergent and construct validity, and demonstrated adequate internal consistency (focused $\alpha = .77$; automatic $\alpha = .78$), although they were not significantly associated with one another ($r = .01, p = .74$; Flessner et al., 2008). In the present study, the MIST demonstrated excellent internal consistency; the Focused scale resulted in an alpha value of .96, and the Automatic scale resulted in a value of .94. Due to a data administration error, the MIST was scored on a 1 to 9 score in the current sample.

Cosmetic Procedure Screening Questionnaire (COPS; Veale et al., 2012). The COPS is a 9-item self-report measure intended to identify individuals with body dysmorphic disorder. Items assess the cognitive and behavioral facets of body dysmorphic disorder such as feature-checking, preoccupation, and avoidance behaviors (e.g., “How often do you deliberately check your feature(s)?” and “How much does your feature(s) currently preoccupy you?”). ROC analysis has suggested that the COPS is a valid diagnostic test because a cutoff score of 40 and above successfully identified body dysmorphic disorder patients 88.9% of the time (Veale et al., 2012). Total scores on the COPS have been shown to have good test-retest reliability over a 12-week time period ($r = .87$) and have demonstrated sensitivity to change in response to various treatment methods (Veale et al., 2012). Convergent validity for the COPS has also been well-established and shown to indicate lower body image levels via high COPS scores. In the present study, the COPS resulted in a Cronbach’s alpha value of .89, suggesting good internal consistency. In the current sample, 33 participants (10% of the sample) scored 40 or above.

Not Just Right Experiences Questionnaire-Revised (NJRE-Q-R; Coles et al., 2003). The NJRE-Q-R is a unidimensional measure that assesses presence and characteristics of NJREs. The first 10 items constitute examples of NJREs and ask participants to indicate whether they have ever experienced each NJRE (e.g., “I have had the sensation after getting dressed that parts of my clothes such as tags, collars, pant legs, etc. didn’t feel just right”). The remaining nine items assess the characteristics of their most frequently occurring NJRE. In the present study, we were most interested in the presence/intensity of NJREs in general, thus participants completed the first 10 items of the NJRE-Q-R. These items were scored on a 7-point Likert-type scale from 1 (not intense at all) to 7 (extremely intense). The measure has been validated in both nonclinical and clinical OCD samples, and has demonstrated good test-retest ability ($r = .76$; Coles et al., 2003). The internal consistency of the NJRE-Q-R proved to be excellent in the current study, $\alpha = .94$.

Results

Given the unacceptably low reliability of the MIDAS Automatic subscale in the current sample, we removed this scale from further analysis. Means, standard deviations, ranges, and sample sizes for each remaining scale are presented in Table 1. We tested our hypothesis by first examining the bivariate correlations among the included scales. We subsequently conducted a hierarchical linear regression analysis to look at the incremental predictive power of each OCRD scale in relation to NJRE-Q-R scores. Multicollinearity statistics were examined for violations of assumptions. Tolerance values were all above the .2 recommended value (Field, 2009; values ranged from .337 to .703); variation inflation factor values did not exceed 10 (Field, 2009; values ranged from 1.41 to 2.93).

Bivariate Correlations Between Study Variables

In line with past research, 95% of individuals who took part in the study reported experiencing at least one NJRE (Coles et al., 2003; Jacoby et al., 2013). Two-tailed bivariate correlational tests for Pearson’s $r$ were run to determine the strength and direction of the relationships between total

### TABLE 1

<table>
<thead>
<tr>
<th>Scale</th>
<th>$M$ ($SD$)</th>
<th>Range</th>
<th>$n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>DOCS – Contamination</td>
<td>3.43 (3.66)</td>
<td>0 – 19</td>
<td>319</td>
</tr>
<tr>
<td>DOCS – Responsibility</td>
<td>3.50 (3.81)</td>
<td>0 – 20</td>
<td>319</td>
</tr>
<tr>
<td>DOCS – Unacceptable Thoughts</td>
<td>3.52 (3.95)</td>
<td>0 – 19</td>
<td>319</td>
</tr>
<tr>
<td>DOCS – Symmetry</td>
<td>2.97 (3.76)</td>
<td>0 – 18</td>
<td>319</td>
</tr>
<tr>
<td>DOCS Total</td>
<td>13.43 (12.37)</td>
<td>0 – 58</td>
<td>319</td>
</tr>
<tr>
<td>SI-R Total</td>
<td>27.48 (14.66)</td>
<td>9 – 78</td>
<td>319</td>
</tr>
<tr>
<td>MIDAS – Focused</td>
<td>8.40 (5.08)</td>
<td>6 – 30</td>
<td>319</td>
</tr>
<tr>
<td>MIST – Automatic</td>
<td>7.03 (5.96)</td>
<td>5 – 34</td>
<td>319</td>
</tr>
<tr>
<td>MIST – Focused</td>
<td>13.71 (10.90)</td>
<td>10 – 67</td>
<td>319</td>
</tr>
<tr>
<td>COPS Total</td>
<td>18.00 (13.69)</td>
<td>0 – 61</td>
<td>319</td>
</tr>
<tr>
<td>NJRE-Q-R Total</td>
<td>21.30 (12.67)</td>
<td>10 – 62</td>
<td>319</td>
</tr>
</tbody>
</table>

Note: DOCS = Dimensional Obsessive-Compulsive Scale; SI-R = Savings Inventory – Revised; MIDAS = Milwaukee Inventory for the Dimensions of Adult Skin Picking; MIST = Milwaukee Inventory for Subtypes of Trichotillomania; COPS = Cosmetic Procedure Screening Questionnaire; NJRE-Q-R = Not Just Right Experiences Questionnaire-Revised. *Due to a data administration error, this scale was scored on a 1–9 scale rather than 0–9.
scores on the NJRE-Q-R and scores on measures of OCRD symptoms (see Table 2). Total NJRE-Q-R scores were strongly correlated with each of the measures of OCRD symptoms ($r = .28–.71$, all $p$’s < .001). However, they were most strongly correlated with the DOCS total score ($r = .71$) and each of DOCS Contamination, Responsibility, Unacceptable Thoughts, and Symmetry ($r = .52–.68$). There was also a strong correlation between NJRE-Q-R and both the COPS ($r = .62$) and SI-R scores ($r = .50$). Finally, NJRE-Q-R scores were weakly to moderately correlated with the MIDAS ($r = .30$) and MIST ($r = .28–.37$) subscales.

To compare the magnitude of associations between the NJRE-Q-R and each of the OCRD measures, two-tailed significance testing of the difference between Pearson’s $r$ correlations was carried out using the Williams modification of the Hotelling test for two correlations involving a common variable (Kenny, 1987). Significance testing was performed by transforming each individual $r$ value into a $z$ score using the Fisher-$z$ transformation, at which point a normalized $z$ table was referenced to find the $p$ values for the difference score for each correlational pair. As expected, overall scores on the DOCS demonstrated significantly stronger correlations with NJRE-Q-R scores ($r = .71$) than the other OCRD symptom measures ($r = .28–.62$), $z = 2.43–8.68$, all $p$ values < .001.

Turning to the DOCS subscales, a similar pattern was found for DOCS Symmetry, in that the correlation between the NJRE-Q-R and DOCS Symmetry ($r = .68$) was significantly stronger than the relation between the NJRE-Q-R and the remaining DOCS subscales ($r = .52–.57$, $z = 3.08–4.04$, all $p$ values < .001), as well as greater than the relation between the NJRE-Q-R and all other OCRD symptom measures ($r = .28–.50$, $z = 3.99–7.45$, all $p$ values < .001, except for the COPS ($r = .62$, $z = 1.49$, $p = .07$). The remaining DOCS subscales were significantly more strongly correlated with NJRE-Q-R scores than was the Focused scale of the MIDAS ($r = .30$, $z = 3.90–4.47$, $p < .001$) and both the Automatic and Focused subscales of the MIST ($r = .28–.37$, $z = 2.60–4.95$, $p < .001$). Finally, both the SI-R and the COPS were also more significantly associated with NJRE-Q-R scores ($r = .50$–.62) than was any MIDAS and MIST scale ($r = .27–.36$, $z = 2.38–6.07$, $p = .008$).

### Hierarchical Linear Regression Analysis Predicting NJREs

To examine the incremental association between specific OCRD symptoms and NJREs, a 2-step hierarchical regression analysis was conducted with NJRE-Q-R scores as the dependent variable. The four DOCS subscales were entered in the first step to examine the combined effects of OCD symptomology. Total scores for the remaining OCRD symptom scales were entered in the second step to observe the predictive power of the other OCRD symptoms on the self-reported experience.

### Table 2

<table>
<thead>
<tr>
<th></th>
<th>Cont</th>
<th>Resp</th>
<th>Thoughts</th>
<th>Sym</th>
<th>DOCS</th>
<th>SI-R</th>
<th>MIDAS_F</th>
<th>MIST_F</th>
<th>MIST_A</th>
<th>COPS</th>
<th>NJRE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cont</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resp</td>
<td>.59</td>
<td></td>
<td>.48</td>
<td>.61</td>
<td>.82</td>
<td>.33</td>
<td>.16</td>
<td>.27</td>
<td>.22</td>
<td>.37</td>
<td>.57</td>
</tr>
<tr>
<td>Thoughts</td>
<td></td>
<td>.53</td>
<td>.58</td>
<td>.83</td>
<td>.36</td>
<td>.38</td>
<td>.32</td>
<td>.26</td>
<td>.27</td>
<td>.43</td>
<td>.55</td>
</tr>
<tr>
<td>Sym</td>
<td></td>
<td></td>
<td>.78</td>
<td>.38</td>
<td>.83</td>
<td>.45</td>
<td>.31</td>
<td>.24</td>
<td>.23</td>
<td>.42</td>
<td>.68</td>
</tr>
<tr>
<td>DOCS</td>
<td></td>
<td></td>
<td></td>
<td>.83</td>
<td>.38</td>
<td>.45</td>
<td>.25</td>
<td>.32</td>
<td>.28</td>
<td>.44</td>
<td>.71</td>
</tr>
<tr>
<td>SI-R</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.71</td>
</tr>
<tr>
<td>MIDAS_F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.50</td>
</tr>
<tr>
<td>MIST_F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.37</td>
</tr>
<tr>
<td>MIST_A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.28</td>
</tr>
<tr>
<td>COPS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.37</td>
</tr>
<tr>
<td>NJRE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.50</td>
</tr>
</tbody>
</table>

Note. $N = 319$. OCRD = Obsessive-Compulsive and Related Disorder; DOCS = Dimensional Obsessive-Compulsive Scale total score; Cont = DOCS Contamination subscale; Resp = DOCS Responsibility for Harm and Mistakes subscale; Thoughts = DOCS Unacceptable Thoughts subscale; Sym = DOCS Symmetry subscale; SI-R = Savings Inventory – Revised total score; MIDAS_F = Milwaukee Inventory for Adult Skin-Picking – Focused subscale; MIST_F = Milwaukee Inventory for Subtypes of Trichotillomania – Focused subscale; MIST_A = Milwaukee Inventory for Styles of Trichotillomania – Automatic subscale; COPS = Cosmetic Procedure Screening questionnaire; NJRE = Not Just Right Experiences Questionnaire – Revised.

*All $p$ values in table < .01.
of NJREs, above and beyond the effects of OCD symptom dimensions (see Table 3). At Stage 1, the initial four-predictor model using the DOCS subscales was significant \((p < .001)\) and accounted for 54% of the variance in NJRE-Q-R scores, \(F(4, 314) = 90.87, p < .001\). Contamination, \(\beta = .16, t(317) = 2.97, p = .003\), Responsibility, \(\beta = .13, t(317) = 2.46, p = .015\), Unacceptable Thoughts, \(\beta = .15, t(317) = 3.18, p = .002\), and Symmetry, \(\beta = .43, t(317) = 8.11, p < .001\), each contributed significantly to the model. It is notable that the effect of Symmetry was more than twice as strong as that of any of the other OCD symptom dimensions assessed. Introducing the remaining OCRD symptom scores in the second step resulted in a significant model \((p < .001)\) and explained an additional 11.0% of the variance in NJRE-Q-R scores; this change in \(R^2\) was significant, \(F(9, 309) = 62.95, p < .001\).

When all nine independent variables were included in the linear regression analysis, DOCS Symmetry remained the strongest predictor of NJRE-Q-R scores, with a standardized \(\beta\) value of .35. Although DOCS Contamination also remained a significant predictor of NJRE-Q-R scores \((\beta = .13, p = .007)\), neither Unacceptable Thoughts \((\beta = .06, p = .199)\) nor Responsibility \((\beta = .05, p = .291)\) remained significant. Notably, COPS \((\beta = .29, p < .001)\) and SI-R \((\beta = .14, p = .001)\) scores were the second and third strongest predictors of NJRE-Q-R scores, respectively. From the MIDAS and MIST, only the MIST-Focused subscale was marginally significant \((\beta = .11, p = .057)\).

**Discussion**

This study was, to our knowledge, the first to directly examine the potential association between NJREs and symptoms of the full range of the DSM-5 defined OCRDs in addition to the specific symptom dimensions of OCD. Therefore, a more explicit perspective of the possible relationships between OCD symptom dimensions and symptoms of the OCRDs, as well as all of their associations with not just right phenomena, was examined. Although there is some evidence of specificity between NJREs and OCD (Coles et al., 2003; Ghisi et al., 2010; Rasmussen & Eisen, 1992), the present study extended the efforts of past research that have postulated that NJREs may not be obsessive-compulsive symptom-specific (Fergus, 2014). Parallel to our hypothesis, the current results suggest that NJREs are associated with not just OCD symptoms, but with symptoms of all of the OCRDs. Specifically, although the strength of correlation between the DOCS and NJREs was strongest, we found that symptoms of each of the OCRDs also demonstrated a significant relationship with NJREs. However, it seems that symptoms of OCD, body dysmorphic disorder, and hoarding disorder are more strongly associated with NJREs than are symptoms of hair-pulling and skin-picking behavior.

**Expanding the Base of Literature Between NJREs and OCD**

We first aimed to replicate past work suggesting an association between NJREs and OCD by examining these associations at the symptom level. Specifically, we looked at the relationship between NJREs and each of the OCD symptom dimensions assessed by the DOCS. In line with our hypotheses, DOCS Symmetry, Contamination, Responsibility, and Unacceptable Thoughts each demonstrated stronger correlational values with NJREs than symptoms of any other OCRD measure, and each was a significant predictor of self-reported NJRE symptoms even when other OCD symptoms were part of the model. These results suggest that NJREs are

<table>
<thead>
<tr>
<th>TABLE 3</th>
<th>Hierarchical Regression Results by OCRD Predictor Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>OCRD Predictors</td>
<td>(t)</td>
</tr>
<tr>
<td><strong>Step 1: DOCS subscales only</strong></td>
<td></td>
</tr>
<tr>
<td>Overall Model</td>
<td>90.87</td>
</tr>
<tr>
<td>Contamination</td>
<td>2.97</td>
</tr>
<tr>
<td>Responsibility</td>
<td>2.46</td>
</tr>
<tr>
<td>Unacceptable Thoughts</td>
<td>3.18</td>
</tr>
<tr>
<td>Symmetry</td>
<td>8.11</td>
</tr>
<tr>
<td><strong>Step 2: DOCS subscales + remaining OCRD measures</strong></td>
<td></td>
</tr>
<tr>
<td>Overall Model</td>
<td>62.95</td>
</tr>
<tr>
<td>Contamination</td>
<td>2.70</td>
</tr>
<tr>
<td>Responsibility</td>
<td>1.06</td>
</tr>
<tr>
<td>Unacceptable Thoughts</td>
<td>1.29</td>
</tr>
<tr>
<td>Symmetry</td>
<td>7.25</td>
</tr>
<tr>
<td>SI-R</td>
<td>3.37</td>
</tr>
<tr>
<td>COPS</td>
<td>7.01</td>
</tr>
<tr>
<td>MIST – Focused</td>
<td>0.38</td>
</tr>
<tr>
<td>MIST – Automatic</td>
<td>1.91</td>
</tr>
<tr>
<td>MIST – Automatic</td>
<td>-0.82</td>
</tr>
</tbody>
</table>

Note: \(N = 319\). The dependent variable for both regressions was the overall score on the Not Just Right Experience Questionnaire – Revised (NJRE-Q-R). DOCS = Dimensional Obsessive Compulsive Scale total score; Contamination = DOCS Contamination subscale; Responsibility = DOCS Responsibility subscale; Unacceptable Thoughts = DOCS Unacceptable Thoughts subscale; Symmetry = DOCS Symmetry subscale; SI-R = Savings Inventory – Revised total score; MIDAS = Milwaukee Inventory for Dimensions of Adult Skin-Picking; MIST = Milwaukee Inventory for Subtypes of Trichotillomania; COPS = Cosmetic Procedure Screening questionnaire; NJRE = Not Just Right Experiences Questionnaire – Revised.
associated with a variety of symptomology relative to OCD, concurrent with past literature (Cougle, Fitch, Jacobson, & Lee, 2013; Fergus, 2014; Taylor et al., 2014; Wahl et al., 2008).

Despite the strong associations found between each of the four OCD symptom dimensions and NJREs, we also examined whether certain OCD symptoms shared stronger relations with NREs than others. Indeed, our results showed that, once other OCRD symptoms were taken into account, only Symmetry and Contamination remained significant predictors of NJREs, which is consistent with some findings that suggest that not all OCD symptom dimensions are equally associated with NJREs (Ecker et al., 2014). In line with the proposition that different motivational domains may underlie OCD symptom dimensions (Summerfeldt, 2004), Ecker and Gönner (2008) reported that symmetry symptoms were uniquely associated with incompleteness whereas obsessional thoughts were uniquely associated with harm avoidance. Thus, the present findings support the notion that obsessive thoughts and checking compulsions may not represent responses to NJREs. It is also important to note that the DOCS Symmetry subscale contains some items assessing the extent to which participants have felt that things are not symmetrical or just right. Although the DOCS measures overt OCD symptoms, and NJREs were being assessed as a broad dimension posited to underlie the development or maintenance of psychopathology symptoms, this shared item content may also account for the strong association between Symmetry and NJRE-Q-R scores.

Another noteworthy finding is related to Symmetry. When all predictors were included in the model, Symmetry accounted for 46.7% of the variance in NJRE-Q-R scores on its own. This is interesting, especially considering that Responsibility and Unacceptable Thoughts were no longer significant predictors of NJREs at Stage 2 of the regression analysis. This result calls into question the notion that NJREs are experienced exclusively by individuals with clinical psychiatric diagnoses because symmetrical concerns are very common among individuals in addition to those who display clinical-level OC symptoms (Cougle et al., 2013). Furthermore, the fact that Symmetry was significantly more strongly correlated with NJREs than was any other DOCS subscale may indicate that NJREs are not experienced equally across OCD symptom dimensions and in particular seem to be less prevalent or strong for OCD symptom dimensions driven by harm avoidance. Although it has already been theoretically touched upon as to why this may be the case, future studies should explore why the potential factors unique to symmetry may make NJREs such a common experience, as this and other research suggests (Ecker & Gönner, 2008; Fergus, 2014).

**Associations Between NJREs and Other OCRDs**

In addition to looking at the differential associations between NJREs and distinct OCD symptom dimensions, we also explored whether NJREs would be associated with symptoms of other DSM-5 defined OCRDs. Because OCRDs are posited to share features and potentially core etiologic mechanisms, we hypothesized that NJREs may also be associated with symptoms of other OCRDs beyond OCD.

**Hoarding and body dysmorphic disorder symptoms.** Hoarding and body dysmorphic disorder symptoms were the strongest correlates with NJRE after OCD itself. When regressed onto NJRE, OCD, hoarding disorder, and body dysmorphic disorder symptoms accounted for 61.6% of all variance in scores, which indicates that NJREs are strongly associated with these symptoms.

One potential explanation for the correlation between NJREs and both hoarding disorder and body dysmorphic disorder symptoms is the similarity between their and OCD symptoms. Unlike skin-picking and hair-pulling behaviors, hoarding disorder and body dysmorphic disorder are both characterized by classic obsessions in the form of repetitive, unwanted, and intrusive thoughts that often compel affected individuals to respond compulsively in order to alleviate anxiety, disgust, or shame (Ecker et al., 2014). Each of these emotions may be closely tied with the occurrence of NJREs in a similar way to how a cluttered desk may engender not-just-right sensations for an individual with OCD symmetry concerns. That is, in each of these disorders, NJREs present in response to obsessions, leading to compulsive response behaviors.

With regard to hoarding behavior, NJREs may arise in relation to a number of symptoms. For example, acquiring behaviors may occur in response to an NJRE regarding not having a certain item that becomes greatly desired or perceived as needed. Similarly, an individual may experience a significant amount of stress when faced with a decision of whether to discard a particular item. Because hoarding often involves excessive attachment to items, it is possible that the thought of being without an item that is perceived to be
important could engender feelings of incompleteness which, in turn, may comprise an NJRE.

Individuals who display body dysmorphic disorder symptoms, on the other hand, may experience body-focused NJREs. Individuals with body dysmorphic disorder perceive their bodies to be flawed in some way, even if that flaw is not perceptible to others or is not present in actuality. This can lead to distorted and perfectionistic patterns of cognition that lead to feelings of incompleteness and guilt that can be experienced as NJREs (Mancini, Gangemi, Perdighe, & Marini, 2008). In response, individuals may choose to engage in a variety of behaviors to mitigate the impact of the NJREs. For example, individuals concerned with hiding a flaw may check their appearance in the mirror as often as they can to make sure their features of concern stay out of sight and, therefore, possibly reduce a sense of imperfection.

**Skin-picking and hair-pulling symptoms.**

Whereas hoarding and body dysmorphic disorder symptoms demonstrated strong correlations with NJRE symptoms, the correlations between NJREs and skin-picking and hair-pulling behavior were weak to moderate. None of the MIDAS or MIST subscales approached a strong correlational magnitude with NJRE-Q-R scores. This is intriguing, especially because, at surface-level, it would seem that both skin-picking and hair-pulling symptoms revolve around strange physiological sensations (Cohen & Leckman, 1992). Indeed, several types of pullers have been identified; pleasure pullers primarily report pleasurable responses to hair-pulling, whereas relief pullers primarily pull their hair in order to relieve or cope with disturbing cognitions or sensations (O’Dlaug, Chamberlain, Schreiber, & Grant, 2013). Therefore, it is not unreasonable to assume that some of these sensations, particularly those experienced by relief pullers, would fit the definition of an NJRE. Although, to our knowledge, no literature to similarly categorize skin-picking exists, much of the extant body of research suggests that the two disorders are highly similar (Snorrason, Belleau, & Woods, 2012) and allow us to sensibly wonder why these symptoms had comparably lower associations with NJREs as compared to symptoms of other OCRDs.

One potential explanation is that the current study did not use a clinical skin-picking or hair-pulling sample and, therefore, there was not enough variance in the current sample to draw substantive conclusions about the relationships between NJREs and skin picking and hair pulling. The majority of our sample (70–86%) reported no experience of hair pulling or skin picking. Another explanation incorporates the inherent differences between skin-picking disorder, hair-pulling disorder, and other OCRDs such as OCD and body dysmorphic disorder. Skin-picking and hair-pulling disorders, both of which can be classified as body-focused repetitive disorders (O’Dlaug et al., 2013), may be less similar to OCRDs than current diagnostic classifications imply. For instance, OCD and body dysmorphic disorder compulsions are enacted in an attempt to reduce anxiety and distress levels that have been elevated by symptoms. In the context of this article, this means that alleviatory behaviors would be in direct response to a negative evaluation of NJREs.

However, not all individuals with skin-picking or hair-pulling disorder perform these behaviors in response to negative emotions. In fact, the behaviors are sometimes gratifying instead of alleviatory (O’Dlaug et al., 2013; Stasik, 2014). Moreover, OCD and body dysmorphic disorder are ego-dystonic in that the individual views the symptoms themselves as distressing and in conflict with one’s own sense of self. In contrast, hair-pulling and skin-picking symptoms, in some cases, are quite ego-syntonic, even leading to pleasure (O’Dlaug et al., 2013). Therefore, it is difficult to say for certain that these behaviors may be driven by NJREs because they may be enacted to acquire a positive sensation rather than alleviate negative ones such as NJREs.

These results pose some interesting questions that should be examined future studies. One such question involves the categorization of hair-pulling and skin-picking disorders in the DSM-5, both of which are classified as OCRDs. Given the potential differences between ego-dystonic disorders such as hoarding disorder, body dysmorphic disorder, and OCD and ego-syntonic (at least in part) disorders such as skin-picking and hair-pulling disorder, future research may help discern whether the body-focused nature of hair pulling and skin picking might better be classified as body-focused repetitive disorders (American Psychiatric Association, 2013).

**Study Limitations**

This study’s contributions should be interpreted within the following limitations. First, the present research did not make use of a clinical sample. Twenty-two percent of our sample self-reported a current psychiatric diagnosis and a moderate proportion (10–31%) scored at or above the cutoff on the OCRD measures. However, we were not
able to examine the association between NJREs and clinical diagnoses themselves. Therefore, the results of the present study reflect associations between NJREs and OCRD symptoms rather than disorders. In addition, the low base rate of several of the assessed disorders might have led to low endorsement rates and limited our ability to find significant associations between study variables. Although these disorders occur on a continuum, the results may not entirely reflect the experience of those with clinical diagnoses. In particular, scores on the measures of hair pulling, skin picking, and body dysmorphic disorder were quite low. Thus, results may not be generalizable to individuals with clinical presentations of these disorders. This study was also cross-sectional in nature, precluding causal or directional inferences. It will be important that future studies employ longitudinal design to determine whether NJREs indeed function as a risk factor for the development of OCD and related disorders.

In addition, the nature of self-report, and particularly electronically based self-report, has been known to be unreliable at times and might have led some participants to embellish or diminish their actual experiences (Austin, Deary, Gibson, McGregor, & Dent, 1998; Balakrishnan, 1999). Although data collected via Mechanical Turk samples have been found to be as reliable as traditional data collection methods (Buhrmester, Kwang, & Gosling, 2011) and we employed validity checks, Mechanical Turk participants have been found to pay less attention to experimental stimuli and to seek external information to ensure their responses are factually correct (Goodman, Cryder, & Cheema, 2013). Future studies might benefit from use of in-vivo NJRE inductions such as seating participants in front of a cluttered desk and instructing them not to organize anything. Previous studies have effectively used lab-based paradigms to induce NJREs and have found associations between in-vivo NJRE induction and both OCD (Coles, Heimberg, Frost, & Steketee, 2005; Cougle et al., 2013; Summers, Fitch, & Cougle, 2014) and body dysmorphic disorder symptoms (Summers et al., 2017). Studies that mirror such methodology and include measurement of other OCRDs can further enhance understanding of the associations between NJREs and OCRD symptoms.

Another limitation is that the NJRE-Q-R assesses NJREs in a manner that may tap into symptom dimensions specific to OCD (e.g., checking, washing; Fergus, 2014), thus artificially strengthening the association between NJREs and OCD, and weakening the associations between NJREs and other types of symptoms. Therefore, using a measure that does not encompass as much content overlap would be beneficial for future research that aims to assess relationships between NJREs and a range of psychopathology.

Finally, it is important to note that we found unexpectedly high correlations between the Focused and Automatic subscales of the MIDAS. This measure was developed to include separable subscales, and previous research has demonstrated low intersubscale correlations ($r = .01, p = .74$; Flessner et al., 2008). However, in the present sample, these subscales were very strongly correlated ($r = .78$). It is possible that this indicates that focused and automatic hair pulling may not be as discordant as previously thought. It may also be an anomaly of our sample. In particular, our participants did not endorse very many hair-pulling or skin-picking symptoms. The majority of our sample received the lowest score possible on each of the MIDAS and MIST subscales. In addition, as noted in our Method, the MIDAS Automatic subscale had very low reliability (Coefficient alpha = .22). It appears that the low reliability value was due to the fact that the two reverse-keyed MIDAS items (which are both part of the Automatic subscale) actually correlated positively with the other items prior to rescoring, leading to poor psychometric properties when the items were reverse-scored. This seems to be because many participants endorsed “not true at all for me” for all of the items, irrespective of the direction the item was keyed. We opted to drop this scale from further analysis. However, this limits the generalizability of these results and highlights that the measure may not be as applicable in a nonclinical sample. Therefore, results involving the MIST and MIDAS in the present study should be interpreted with caution because these measures did not function as expected.

Despite the study’s limitations, the yields of this research allow for further speculation not only into what spectra of disorders NJREs may present in, but in what specific types of OCRD symptoms. Furthermore, it may be that therapeutically addressing risk factors that underlie higher order problems regarding obsessive thoughts and compulsive behaviors could be useful in treating an array of disorders.

**Future Directions**

Although this research provided insight regarding the relationship between NJREs and symptoms...
of the OCRDs, several aspects of the phenomena are still unclear. First, although some studies have demonstrated the potential universality of NJREs for all individuals including those who do not suffer from psychiatric disorders (Coles et al., 2003; Jacoby et al., 2013), further evidence may be beneficial to support the strength of these assertions. If NJREs manifest in healthy individuals as well as individuals with psychiatric disorders, future research could delineate potential reasons why individuals focus on not-just-right sensations when they present. More specifically, why some people may become obsessed with their NJREs (such as someone with clinical hoarding, body dysmorphic, or OCD symptoms compared to the numerous others who potentially experience similar sensations but remain healthy) is not yet understood.

One possible explanation for the difference in attention paid to NJREs between those with and without a psychological disorder is a simple gap in expectations that individuals hold about their desired states. A highly perfectionistic individual, for example, may be more prone to developing clinical symptoms in response to a cluttered desk when compared to a nonperfectionistic counterpart. Because more situations are likely to be perceived as imperfect, the sheer experience of NJREs may expedite the development of clinical symptoms. Additionally, if most people do experience such sensations, it may be that only those who interpret NJREs as meaningful develop symptoms of various disorders (Ecker & Gönner, 2008; Moretz & McKay, 2009). Further research to either validate or invalidate such speculations would greatly increase support the strength of these assertions. If NJREs as OCD, results indicated that symptoms of each OCRD were significantly correlated (with varying magnitudes) to overall NJRE-Q-R scores. In particular, hoarding disorder and body dysmorphic disorder symptoms demonstrated strong positive correlations to NJREs both in correlational and hierarchical regression analyses. In summary, this article offers a new perspective on NJREs in relation to a variety of psychiatric symptoms that were previously unexplored in this context.

A final consideration of importance is the conceptualization of NJREs themselves. NJREs have broadly been defined as the need or drive for perfection, certainty, or "closure" (Ecker & Gönner, 2008; Rasmussen & Eisen, 1992). However, whether NJREs represent a vulnerability toward OCD and related symptomatology, or should be considered a symptom themselves is unclear. In some frameworks, NJREs are considered a trait-like underlying motivation that gives rise to symptoms such as those seen in OCD (Ecker & Gönner, 2008; Summerfeldt, 2004; Summerfeldt et al., 2014). As described earlier, "not just right" feelings may reflect one of the affective-motivational domains that underlie overt OCD symptoms. Defined this way, internal discomfort with one’s current sensory experience leads to the need to engage in compulsive behavior (Summerfeldt, 2004; Summerfeldt et al., 2014). This was the conceptualization we used in the current study because we hypothesized that this same underlying drive or motivation may underlie symptoms of other OCRDs.

In other frameworks, though, NJREs are considered to be a symptom of the disorder itself (Moretz & McKay, 2009). For instance, the Vancouver Obsessional Compulsive Inventory (Thordarson et al., 2004) includes a 12-item "Just Right" subscale (e.g., "I often have trouble getting things done because I try to do everything exactly right"). Here, NJREs may be more similar to an obsession, in which its appraisal (rather than its existence) leads to compulsive behavior (Belloch et al., 2016). Future work examining the potential transdiagnostic nature of NJREs may shed further light on the precise nature of this experience.

**Conclusion**

This study aimed to better understand how NJREs manifest in a variety of symptom dimensions outside of those that have been previously studied. This involved a correlational assessment between NJRE-Q-R scores and the current spectrum of OCRD symptoms as denoted by the DSM-5 including hair-pulling and skin-picking disorders, body dysmorphic disorder, hoarding disorder, and OCD, as well as their subscales. Although less strongly associated with NJREs as OCD, results indicated that symptoms of each OCRD were significantly correlated (with varying magnitudes) to overall NJRE-Q-R scores. In particular, hoarding disorder and body dysmorphic disorder symptoms demonstrated strong positive correlations to NJREs both in correlational and hierarchical regression analyses. In summary, this article offers a new perspective on NJREs in relation to a variety of psychiatric symptoms that were previously unexplored in this context.

**References**


Ecker, W., Kupfer, J., & Gönner, S. (2014). Incompleteness as a link
Ecker, W., & Gönner, S. (2008). Incompleteness and harm avoidance in
Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon’s Mechanical
https://doi.org/10.1177/0020764012463520
Perspectives on Psychological Science, 6, 3–5.
https://doi.org/10.1177/1745691610393980
cognitive-behavioral therapy for trichotillomania: A randomized controlled trial.
Journal of Nervous and Mental Disease, 202, 893–898.
https://doi.org/10.1097/NMD.0000000000000215
http://dx.doi.org/10.1016/j.janxdis.2013.07.008
Boston, MA: Little, Brown and Company.
Leckman, J. F., Grice, D. E., Barr, L. C., de Vries, A. L. C., Martin, C.,
http://dx.doi.org/10.1002/anxi.3070010504
http://dx.doi.org/10.1016/j.jbtep.2007.02.002
http://dx.doi.org/10.1521/ijct.2008.1.3.181
http://dx.doi.org/10.1016/j.janxdis.2009.01.015
http://dx.doi.org/10.1016/j.jpsychires.2013.06.033
http://dx.doi.org/10.1017/s1092852913001501.2013.828079
https://doi.org/10.1016/j.brat.2010.05.021
http://dx.doi.org/10.1017/bec.2015.11
https://doi.org/10.1017/s1092852900020319
http://dx.doi.org/10.1017/s0143208400011472
http://dx.doi.org/10.1177/2167702612469015
Clinical Psychology Review, 32, 618–629. https://doi.org/10.1016/j.cpr.2012.05.008


Author Note. Jeremy P. Schmidt, Department of Psychology, Knox College; Sara M. Stasik-O’Brien, Department of Psychology, Knox College.

Funding for this research was awarded to the first author by the Knox College McNair Scholars Program. Special thanks to Jon Crider for his assistance in using Amazon Mechanical Turk. Thanks also to all Psi Chi Journal reviewers for their support.

Correspondence concerning this article should be addressed to Jeremy P. Schmidt at jpschmidt@knox.edu
A traumatic life event often impacts future functioning, yet this relationship occurs in a complex environment with many interacting variables. One particularly salient and complex form of this relationship is that between childhood abuse and subsequent psychological functioning. The World Health Organization (WHO; 2014) estimated that 23% of people report having experienced physical abuse as children, and 20% of women and 5 to 10% of men report having experienced sexual abuse as children. Although the psychological outcomes of such childhood abuse have been investigated, there are specific areas that prior literature has failed to address. Previous researchers have tended to operationalize childhood abuse not as a generalized traumatic event, but as emotional, physical, or sexual abuse (Chartier, Walker, & Naimark, 2009; Chen et al., 2014; Shapero et al., 2014). Relevant research has predominantly focused on the relationship between these subtypes of childhood abuse and subsequent physical health, mental health, and risk-taking behavior (Chartier, Walker, & Naimark, 2009; Chen et al., 2014; Shapero et al., 2014). Although these prior approaches have been informative, child abuse as a composite construct is often excluded from analyses. As a result, there has been a lack of research investigating the effects of childhood abuse as a composite score on interpersonal relationships later in life. The cumulative effects
of childhood abuse and trauma are increasingly relevant to researchers and clinicians, particularly around the emerging discussion on developmental trauma disorder (Bremness & Polzin, 2014; van der Kolk, 2005; van der Kolk, et al., 2009). Further, child abuse has been related to behavioral outcomes such as substance abuse (Chartier et al., 2009; Keyes, Hatzenbuehler, Grant, & Hasin, 2012; Plant, Miller, & Plant, 2007). However, prior research has unearthed mixed findings on the relationship between childhood abuse and interpersonal problems (Keyes, McLaughlin, et al., 2012; Plant et al., 2007). The present study addressed these gaps by investigating the relationship between overall childhood abuse and later life interpersonal problems. In particular, the emphasis of the current study was on possible factors that helped to explain the effects of childhood abuse on social and relational difficulties. To move beyond simply identifying correlations between early abuse and later interpersonal problems, self-defeating beliefs (negative automatic thoughts) and binge-drinking behavior were investigated as cognitive and behavioral mediators, respectively, of this relationship.

**Childhood Maltreatment**

Vast numbers of children experience maltreatment. The United States Department of Health and Human Services, Administration for Children and Families (2015) estimated the victimization rate in the United States to be 9.2 per 1,000 children, or a total of 683,000 victims per year. By sex, the rate of victimization for girls is 9.5 per 1,000 children and 8.7 per 1,000 children for boys (CDC Violence Prevention, 2012). Rates of victimization also vary across race/ethnicity; per 1,000 children, rates are 14.2 for African Americans, 10.3 for Multiracial children, 8.4 for Hispanics, 8.0 for non-Hispanic Whites, and 1.7 for Asians (CDC Violence Prevention, 2012).

For many of these children, the consequences of their victimization pervade into later life. Victims of childhood abuse and neglect are more likely to experience cardiovascular disease, liver disease (Felitti & Anda, 2009), diabetes, lung disease (Widom, Czaja, Bentley, & Johnson, 2012), and obesity (Shin & Miller, 2012). Adults with a history of childhood physical or sexual abuse are at increased risk for malnutrition and vision problems (Widom et al., 2012). In addition to physical ailments, childhood abuse victims also experience psychological, social, and behavioral issues in adulthood.

Psychologically, childhood abuse is a risk factor in the development of depression, anxiety, and other psychiatric disorders (Kaplow & Widom, 2007). For adults who have experienced childhood emotional abuse, experiences of low self-esteem and depressive symptoms are more common (Briere & Runtz, 1990). Walker, Holman, and Busby (2009) found that, although childhood sexual abuse (CSA) only explained a portion of current negative impact regarding personal family history, it was directly related to depression. Further, childhood emotional abuse predicts higher levels of depressive symptomology when individuals experience current life stressors (e.g., health concerns, a fight with family member, etc.; Shapero et al., 2014). In a study of Han Chinese women, Chen et al. (2014) found that not only was CSA related to increased vulnerability for development of recurrent major depression (rMD), but also that the symptoms associated with rMD were more severe for victims of CSA than for rMD patients without a history of CSA. In addition, CSA was associated with suicidal ideation, suicide attempts, feelings of worthlessness, and feelings of guilt (Chen et al., 2014). Mental health problems also partially mediate the relationship between childhood abuse and adult health (Chartier et al., 2009). This means that mental health problems are predicted by childhood abuse and compound later life adult health issues.

In addition to psychological consequences and potential direct health effects, childhood abuse is predictive of risk-taking behavior and substance use (Chartier et al., 2009; Keyes, Hatzenbuehler, et al., 2012; Plant et al., 2007). Health risk behaviors such as alcohol problems, smoking, obesity, and having more than one sexual partner in the last year are all related to prior childhood abuse (Chartier et al., 2009). These health risk behaviors are also partial mediators of the relationship between childhood abuse and adult health (Chartier et al., 2009). Further, childhood abuse is positively related to adult medical costs (Reeve & van Gool, 2013), which may contribute to additional life stress.

Previous work has investigated childhood abuse in terms of interpersonal problems, but primarily within the context of romantic relationships. Walker et al. (2009) found that CSA is predictive of lower romantic relationship quality. Conversely, Larsen, Sandberg, Harper, and Bean (2011) found that frequency of CSA was not predictive of relationship quality. Despite the conflicting findings, other researchers have purported that victims of CSA have a more difficult time confiding in a romantic partner and report less happiness in their current relationship.
Childhood Abuse and Interpersonal Problems

relationship (Plant et al., 2007). Findings have also suggested that victims of childhood physical abuse experience relationship problems in later life (Larsen et al., 2011). Further, emotional dysregulation partially mediates the relationship between childhood emotional abuse and future relationship quality (Bradbury & Shaffer, 2012). Although there is some debate regarding CSA and romantic relationship quality, most evidence has indicated that the traumatic effects of childhood abuse carry onward into later life romantic relationships (Larsen et al., 2011; Plant et al., 2007; Walker et al., 2009).

Self-Defeating Beliefs

Self-defeating beliefs are cognitions that are highly similar to negative automatic thoughts. For example, a self-defeating belief could be “I am not good enough to succeed at this task,” whereas a negative automatic thought could be “I am bound to fail at this task.” For the present study, self-defeating beliefs were conceptualized interchangeably with negative automatic thoughts. Self-defeating beliefs included negative automatic thoughts regarding love, approval, achievement, perfectionism, self-blame, and helplessness. Such negative thoughts have been shown to be important in depressive disorders. According to Beck’s negative cognitive triad theory, negative automatic thoughts that are central to depression focus on three areas: the world, the self, and the future (Beck & Alford, 2009). Through this framework, depressed patients interpret ongoing experiences as negative and disparaging, see unpleasant experiences as internal defects, and anticipate that such a negative state of affairs will continue into the future (Beck & Alford, 2009). Beck theorized that these three thought patterns contribute to development and maintenance of depression.

This well-established connection between negative cognitions and depression has prompted interest in how negative cognitions may relate to other types of impaired functioning. Beck theorized that maladaptive thinking is a key feature of psychopathology in general, not just depression (Beck & Alford, 2009). Through current research, some have investigated the usefulness of the cognitive triad theory for comorbid cases of eating disorders and unipolar depression (Green et al., 2009). Although depression only predicts a fraction of behaviors associated with eating disorders (1%), the major predictors of eating disorders are low self-esteem, maladaptive social comparison, and body dissatisfaction. These three factors can be considered as self-defeating beliefs and are referred to as the eating disorder cognitive triad (Green et al., 2009).

Accordingly, self-defeating beliefs are predictive of negative future mental health outcomes. One study indicated that persons with a history of emotional abuse have higher levels of outwardly expressed negative thoughts (Harmelen et al., 2010). It has also been supported that, for people with depression in cognitive behavioral group therapy, a decrease in negative automatic thoughts is strongly associated with higher quality of life and lessened depressive symptomology (McEvoy, Burgess, & Nathan, 2013). This suggests that a decrease in negative automatic thoughts promotes more adaptive human functioning.

In terms of psychological functioning, increased negative automatic cognitions are associated with heightened anxious and depressed moods (Starr & Davila, 2012). In addition, interpersonal problems has been shown to moderate the relationship between anxious moods and later depressed moods. Briere and Runtz (1990) investigated psychosocial dysfunction as a function of childhood abuse; main findings included that childhood abuse was associated with generalized psychosocial problems. This further supports the notion that negative cognitions are not solely related to depressive symptoms, but also to the broader context of maladaptive behavior. In a study of college students, negative automatic thoughts were associated with both maladaptive coping strategies and anxiety (Mahmoud, Staten, Lennie, & Hall, 2015). Calvete, Orue, and Hankin (2013) found that early maladaptive schemas contribute to negative automatic thoughts that are directly related to increased symptoms of social anxiety and interpersonal problems. In addition, this relationship is bidirectional in that negative thoughts are strengthened by the underlying maladaptive schema. Flouri and Panourgia (2014) found that life stress among adolescents was indirectly associated with interpersonal problems with peers through negative automatic thoughts. Given this supporting literature, it is theoretically sound to hypothesize that negative cognitions (i.e., self-defeating beliefs) would be related to maladaptive social problems in addition to depression, anxiety, and eating concerns.

Alcohol Use

Alcohol use is an important behavioral factor in the relationship between childhood abuse and negative life outcomes including interpersonal problems (Chartier et al., 2009; Keyes, Hatzenbuehler, et al.,
disagreeing with others when needed or that they disagree with others too often. The relevance of interpersonal problems to diverse aspects of human development and functioning has spurred considerable research on the topic including the identification of types of interpersonal problems (for an overview, see Horowitz, 1996). Although the specific manifestations of interpersonal problems in part depend on the social context of an individual, the underlying core difficulties are central components of relationships in general. Specific to the present study, interpersonal problems have been found to be associated with childhood abuse (Larsen et al., 2011), negative automatic thoughts (Flouri & Panourgia, 2014), and alcohol use in romantic relationships (Fisher et al., 2005; MacDonald, Zanna, & Holmes, 2000). Gonzalez and Skewes (2013) found that college students who reported solitary binge-drinking experienced greater social discomfort than social binge-drinkers. Given this literature, it is expected that binge-drinking will be associated with interpersonal problems.

Present Study
Previous researchers have focused predominantly on childhood abuse in terms of the subcategories contained under the umbrella of negative childhood experiences (i.e., sexual, emotional, and physical). This led to specific knowledge of types of abuse, but not necessarily the construct of childhood abuse as a whole. In addition, maladaptive adult functioning has typically been measured through disorders, physical health, health-risk behavior, and romantic relationship quality. Thus, the effect of childhood abuse on generalized psychosocial functioning has not been fully explored in the context of interpersonal problems. Finally, the exact role of alcohol use and negative cognitions in the relationship between childhood abuse and interpersonal problems has yet to form into an empirically backed perspective.

The aim of the present study was to advance understanding regarding the complex relationship between childhood abuse and later life functioning. To address the limitations of previous research, childhood abuse was measured as a total score of three subscales; specifically, emotional abuse, sexual abuse, and physical abuse. Also, interpersonal problems were examined as general psychosocial functioning instead of as romantic relationship quality. Lastly, alcohol consumption was measured as binge-drinking behavior in order to clarify whether quantity of consumption per
session (binge drinking) was impactful on later life social functioning. Binge-drinking behavior and self-defeating beliefs are theoretically conceptualized as both cognitive (self-defeating beliefs) and behavioral (binge drinking) symptoms of generalized childhood abuse. The research hypothesis of this study was twofold: (a) self-defeating beliefs were expected to mediate the relationship between childhood abuse and interpersonal problems, and (b) binge-drinking behavior was expected to also mediate this relationship.

These hypotheses were investigated using mediation analysis, a method that allows the determination of both direct and indirect effects using a series of regression models (Baron & Kenny, 1986). In our first hypothesized model, the exposure variable was child abuse, the outcome variable was interpersonal problems, and the mediator variable was self-defeating beliefs (see Figure 1). In our second hypothesized model, the exposure variable was child abuse, the outcome variable was interpersonal problems, and the mediator variable was binge-drinking (see Figure 2). The first step to using this method is verifying that bivariate associations exist between the exposure variable and the outcome variable, and between the exposure variable and the mediator variable. The second step is to determine if an association exists between the mediator variable and the outcome variable when controlling for the exposure variable. The third step is to investigate the coefficient for the relationship between the exposure variable and the outcome variable when controlling for the indirect path from the exposure to the outcome through the mediator. If this coefficient is zero, there is evidence of full mediation. However, if this coefficient is attenuated, but still significantly different from zero, there is evidence of partial mediation (Baron & Kenny, 1986). For more information on mediation analysis, see Preacher and Hayes (2008).

**Method**

**Participants**

Participants were 171 undergraduate students who were part of a study investigating the relationship among personality, relationships, and alcohol consumption. Participants completed an online survey using SurveyMonkey, a website for survey design, collection, and analysis. This survey served as a screening method for participants who were possibly eligible for participation in related studies. Demographic information, personality questions, and behavioral questions were included in the survey. Only scales detailed in the measures section below were utilized for the current project (AUDIT, CTQ, IIP-C-IRT, and SDB). Participants’ data were included in the current study if it met criteria for five distractor questions included in the online survey. This was done to minimize the effect of participants who were randomly or inattentively answering questions. Originally, 264 participants completed the survey. However, after filtering out participants for distractor questions and with nonresponse for variables of interest, the final sample consisted of 171 participants (125 women, 46 men). Participants ranged from 18 to 69 years old ($M = 22.87, SD = 7.38$), although the 69-year-old participant was a substantial outlier and was dropped from the analytic sample. After exclusion, participant age ranged from 18 to 52 ($M = 22.60, SD = 6.49$), and the final analytic sample included 170 participants. Of the analytic sample, 122 reported being White (71.8%), 21 reported being Black (12.4%), 13 reported being Hispanic (7.6%), five reported being Asian (2.9%), and nine reported being other or multiracial (5.3%). One hundred six participants reported being single.
(62.4%), 41 reported having a significant other (24.1%), 21 reported being married (12.4%), and two reported being divorced (1.2%).

Measures

Binge drinking. Levels of binge drinking were measured using a modified version of the Alcohol Use Disorders Identification Test (AUDIT; Babor, Higgins-Biddle, Saunders, & Monteiro, 2001). The AUDIT is a 10-item scale designed and validated across age and gender as a screening measure for excessive alcohol consumption (Allen, Litten, Fertig, & Babor, 1997; Saunders, Aasland, Babor, de la Fuente, & Grant, 1993). Each item on the scale is a Likert-type scale with responses varying with question. For the present study, only questions measuring amount of alcohol consumed per drinking session were included due to the aforementioned findings of Plant et al. (2007). A subscale measuring amount of alcohol consumed per drinking session was computed by totaling the responses from the two questions that pertained to said construct. The first question, “How many drinks containing alcohol do you have on a typical day when you are drinking?” had response choices ranging from “1 or 2” to “10 or more.” The second question, “How often do you have six or more drinks on one occasion?” had response choices ranging from “never” to “daily or almost daily.” A Cronbach’s alpha measure of reliability was calculated for the binge-drinking subscale ($\alpha = .84$).

Childhood maltreatment. Childhood maltreatment was measured using the Childhood Trauma Questionnaire (CTQ; Bernstein & Fink, 1998). The CTQ is a 28-item retrospective self-report measure of childhood maltreatment that has been validated for use with adolescents and adults (Bernstein et al., 2003; Wright et al., 2001). There are six subscales for the CTQ: sexual abuse, emotional abuse, physical abuse, emotional neglect, physical neglect, and minimization/denial. The CTQ uses a Likert-type scale from 1 (never true) to 5 (very often true) with a minimum score of 1 (no abuse) on each subscale and a maximum score of 25 (severe abuse). For the purposes of the present study, only the sexual abuse (e.g., “Someone tried to make me do sexual things or watch sexual things”), emotional abuse (e.g., “People in my family said hurtful or insulting things to me”), and physical abuse (e.g., “I got hit so hard by someone in my family that I had to see a doctor or go to the hospital”) scales were calculated. The emotional neglect, physical neglect, and minimization/denial subscales were not included in the present analyses because these subscales measured constructs outside of the aims of the present study. The total abuse score was calculated as the sum of the three subscales with a minimum score of 15 (no abuse) and a maximum score of 75 (severe abuse). Reliability analyses were conducted with the current sample using Cronbach’s alpha for total abuse ($\alpha = .84$).

Interpersonal problems. Interpersonal problems were measured using the Inventory of Interpersonal Problems-Circumplex-Item Response Theory (IIP-C-IRT; Barkham et al., 1996; Sodano & Tracey, 2011). The IIP-C-IRT is a 32-item version of the 64-item IIP. The multiple versions of the IIP are each the result of rigorous development and numerous validation studies (Hopwood, Pincus, DeMoor, & Kooonce, 2008; Monsen, Hagtvet, Havik, & Eilersten, 2006; Ruiz et al., 2004). The IIP-C-IRT is designed to measure and identify an individual’s most salient interpersonal problems, or experienced difficulties in their relationships with others. The IIP-C-IRT is measured on two orthogonal dimensions (i.e., control and affiliation), and is comprised of eight subscales: domineering-controlling, vindictive-self-centered, cold-distant, socially inhibited, nonassertive, overly accommodating, self-sacrificing, and intrusive-needy. Each item on the individual subscales is a Likert-type statement that indicates for participants to specify how distressing a problem has been for them ranging from 1 (not at all) to 5 (extremely). An example statement measuring problems that arise from difficulty being assertive is “It is hard for me to be firm when I need to be,” whereas an example statement measuring problems that arise from being socially inhibited is “It is hard for me to join in on groups.” Given the focus on interpersonal problems in general, a total IIP-C-IRT score was calculated as the average of all 32 items with a range of scores from 1 (less interpersonal problems) to 5 (more interpersonal problems). A Cronbach’s alpha was calculated for the total IIP-C-IRT measure ($\alpha = .90$).

Self-defeating beliefs. Self-defeating beliefs were measured using a modified version of the Self-Defeating Belief Scale (SDBS; Healing Heart Counseling Center, 2008). For the purpose of this analysis, the subscales of approval addiction, love addiction, achievement addiction, perfectionism, and entitlement were not included. The addiction subscales do not appear to represent negative automatic cognitions, and perfectionism and entitlement appear to be separate constructs from those of interest in this study. The self-blame and
helplessness/hopelessness subscales of the SDBS were included due to their perceived ability to measure negative automatic cognitions. Each subscale consisted of five items. An example statement from the self-blame subscale is “I usually blame myself for the problems in my relationships with other people.” An example statement from the helplessness/hopelessness subscale is “It would be extremely difficult or impossible to solve the problems in my life.” A Cronbach’s alpha measure of reliability was calculated for the SDBS (α = .85).

### Procedure

Participants were undergraduate psychology students at a large comprehensive university in the southeast. Students were offered participation credit for participation. If students declined participation in research, an alternate assignment was provided for the class participation credit. Institutional review board approval was obtained prior to the data collection (Study #15-200), and informed consent was obtained at the beginning of the online survey. Demographics were obtained first, and the subsequent scales were administered in the following order: AUDIT, CTQ, IIP-C-IRT, and SDBS. In addition to the previously described measures, five distractor questions were imbedded in the survey to ensure participant attentiveness. For example, participants were asked to select a particular answer for a distractor questions (e.g., Please select “extremely” for this question.). All distractor questions had to be answered correctly for inclusion in the study. Because of the sensitive nature of the surveys, participants were given contact information for the counseling center on campus for referral to appropriate service providers if they felt any psychological or emotional stress from participation.

### Results

#### Preliminary Analyses

In regard to sex, an independent-samples t test was conducted to assess whether overall childhood abuse scores differed for men and women. The Levene’s test for equality of variances indicated that the variances were not assumed to be equal (F = 11.90, p < .001), thus degrees of freedom were adjusted from 168 to 113. Overall childhood abuse scores were higher for women (M = 24.74, SD = 8.18) than for men (M = 21.87, SD = 5.82), t(113) = -2.54, p = .012, d = 0.40. To assess differences in childhood abuse by race/ethnicity, a one-way Analysis of Variance (ANOVA) was conducted. The results indicated no significant difference in childhood abuse among race/ethnicity. Differences were also assessed for the relationship categories. The results of the one-way ANOVA indicated no significant difference in childhood abuse between relationship statuses.

#### TABLE 1

| Bivariate Correlations for Scores of CTQ, SDB, IIP-C-IRT, and Binge Drinking Measures |
|---------------------------------|---|---|---|---|
|                               | 1. | 2. | 3. | 4. |
| Childhood Abuse               |    | .25** | .31*** | -.13 |
| Self-Defeating Beliefs        |    |    | .66***| .02 |
| Interpersonal Problems Mean   |    |    |    | -.03 |
| Binge Drinking                |    |    |    |    |

Note: N = 170. **p < .01. ***p < .001.

#### TABLE 2

| Direct Effect Measures, OLS Regression of Child Abuse on Self-Defeating Beliefs, and Self-Defeating Beliefs on Interpersonal Problems |
|---------------------------------------------------------------|-------|-------|
|                                                          | Direct Effects on       | Direct Effects on       |
|                                                          | Self-Defeating Beliefs  | Interpersonal Problems  |
| Childhood Abuse                                             | 0.26** | (0.26) |
| Sex                                                          |       |       |
| Men (reference)                                             | 1.03  | (0.06) |
| Women                                                        |       |       |
| Race                                                         |       |       |
| White (ref)                                                  |       |       |
| Black                                                        | -4.87* | (-0.20) |
| Hispanic                                                     | -5.29* | (-0.18) |
| Asian                                                        | -3.14* | (-0.07) |
| Other/Multiracial                                            | -3.33* | (-0.09) |
| Relationship Status                                         |       |       |
| Single (reference)                                          |       |       |
| Significant Other                                           | -0.25 | (-0.01) |
| Married                                                     | -2.57 | (-0.11) |
| Divorced                                                    | 1.91  | (0.03) |
| Self-Defeating Beliefs                                      |       |       |
|               |       |       |
| Intercept                                                 | 21.33 |       |
| R²                                                        | 0.14  | 0.44  |
| Adj R²                                                    | 0.09  | 0.44  |
| F                                                         | 2.92* | 133.27*** |

Note: N = 170. Source: Alcohol and Stress Survey. **p < .01. ***p < .001 (two-tailed).
Bivariate Correlations

Correlations were conducted for all quantitative variables included in the study including total childhood abuse, self-defeating beliefs, mean interpersonal problems, and binge drinking (see Table 1). Significant correlation relationships were present between childhood abuse, self-defeating beliefs, and interpersonal problems. Childhood abuse was positively related to both self-defeating beliefs and average interpersonal problems, consistent with the hypothesis. Also expected, self-defeating beliefs were found to be moderately to highly correlated with interpersonal problems. However, in contradiction to the hypothesis that childhood abuse and binge drinking would be related, there was no apparent correlation between the CTQ and binge-drinking behavior (at least as it was measured in the current study), self-defeating beliefs, or interpersonal problems.

Self-Defeating Beliefs

Direct effects. Several regressions were conducted to assess the hypothesis that self-defeating beliefs mediate the relationship between childhood abuse and interpersonal problems in later life. In the first model, $F(9, 169) = 2.92, p = .003$ (see Table 2), when controlling for sex, race/ethnicity, and relationship status, childhood abuse was a significant predictor of self-defeating beliefs ($\beta = 0.26, p = .001$). Compared to White participants, Black participants ($\beta = -4.87, p = .010$) and Hispanic participants ($\beta = -5.29, p = .020$) reported lower levels of self-defeating beliefs on average. For the second model, $F(1, 169) = 132.27, p < .001$ (see Table 2), self-defeating beliefs predicted interpersonal problems ($\beta = 0.05, p < .001$).

Indirect effects. In the third model (nested regression, see Table 3), when self-defeating beliefs were added to the model, the beta coefficient for childhood abuse decreased from $\beta = .02 (p < .001)$ to $\beta = .01 (p = .032)$. In addition, self-defeating beliefs

<table>
<thead>
<tr>
<th>TABLE 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>OLS Regression on Interpersonal Problems as a Function of Childhood Abuse, Sex, Race/Ethnicity, Relationship Status, and Self-Defeating Beliefs</strong></td>
</tr>
<tr>
<td>---------------------------------</td>
</tr>
<tr>
<td><strong>Model 1</strong></td>
</tr>
<tr>
<td>---------------------------------</td>
</tr>
<tr>
<td><strong>Childhood Abuse</strong></td>
</tr>
<tr>
<td>---------------------------------</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
</tr>
<tr>
<td>Women</td>
</tr>
<tr>
<td><strong>Race</strong></td>
</tr>
<tr>
<td>White (reference)</td>
</tr>
<tr>
<td>Black</td>
</tr>
<tr>
<td>Hispanic</td>
</tr>
<tr>
<td>Asian</td>
</tr>
<tr>
<td>Other/Multiracial</td>
</tr>
<tr>
<td><strong>Relationship Status</strong></td>
</tr>
<tr>
<td>Single (reference)</td>
</tr>
<tr>
<td>Married</td>
</tr>
<tr>
<td>Divorced</td>
</tr>
<tr>
<td><strong>Self-Defeating Beliefs</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Intercept</td>
</tr>
<tr>
<td>$R^2$</td>
</tr>
<tr>
<td>Adj $R^2$</td>
</tr>
<tr>
<td>$F$ Ratio</td>
</tr>
<tr>
<td>$F$-Change</td>
</tr>
</tbody>
</table>

**Note.** $N = 170$. Source: Alcohol and Stress Survey. *$p < .05$, **$p < .01$, ***$p < .001$ (two-tailed).
accounted for additional variation explained by the model ($R^2 = .47, \Delta R^2 = .44$), $F(10, 169) = 14.34, p < .001$. These results indicate that self-defeating beliefs partially mediated the relationship between childhood abuse and interpersonal problems. Bootstrapping of these relationships was completed using the "PROCESS" program (Hayes, 2013). For the current analysis, the bias-corrected confidence interval construction method was employed with 5,000 bootstrap samples, the recommended amount of samples for bootstrapping (Preacher & Hayes, 2008). From the bootstrap procedure, the 95% confidence interval for the indirect effect of SDB on the relationship between childhood abuse and interpersonal problems is .0033 and .0236.

**Binge Drinking**

**Direct effects.** To test the second hypothesis that binge drinking would mediate the relationship between childhood abuse and interpersonal problems, several regressions were conducted. In the fourth model, $F(9, 169) = 0.85, p = .574$ (see Table 4), when controlling for sex, race/ethnicity, and relationship status, childhood abuse was not a significant predictor of binge-drinking behavior ($\beta = -0.03, p = .153$). For the fifth model, $F(1, 169) = 0.12, p = .733$ (see Table 4), binge-drinking behavior was not a significant predictor of interpersonal problems ($\beta = -0.01, p = .733$).

**Indirect effects of binge drinking.** In the sixth model (nested regression, see Table 5), when binge drinking was added to the model, the beta decreased from $\beta = 0.023$ ($p < .001$) to $\beta = 0.022$ ($p < .001$). Further, binge drinking produced no significant change in the amount of variation explained by the model ($R^2 = .158, \Delta R^2 = .084$), $F(10, 169) = 2.54, p = .007$. These results do not support that binge drinking mediates the relationship between childhood abuse and interpersonal problems.

**Analysis of types of interpersonal problems.** To further explore, we conducted an additional analysis that revealed interesting conclusions on the relationship between childhood abuse and interpersonal problems. To identify the interpersonal problems that were associated with high levels of childhood abuse, we divided the participants into quartiles based on childhood abuse scores, and conducted one way ANOVAs between abuse level and each of the eight interpersonal problems subscales. The ANOVA revealed that there were significant differences between the abuse groups on the subscales domineering-controlling, $F(3, 169) = 6.43, p < .001$, vindictive-self-centered, $F(3, 169) = 6.59, p < .001$, cold-distant, $F(3, 169) = 4.91, p = .003$, socially inhibited, $F(3, 169) = 3.48, p = .017$, and self-sacrificing, $F(3, 169) = 3.27, p = .023$. Bonferroni post-hoc tests indicated that the participants who scored in the highest quartile of childhood abuse also scored significantly higher than the two lowest quartiles on domineering-controlling (1st quartile vs. 4th quartile, $p < .001$, 2nd quartile vs. 4th quartile, $p = .027$), vindictive-self-centered (1st quartile vs. 4th quartile, $p < .001$, 2nd quartile vs. 4th quartile, $p = .012$), and on the cold-distant subscale (1st quartile vs. 4th quartile, $p = .005$, 2nd quartile vs. 4th quartile, $p = .008$). Post-hoc testing revealed no significant differences comparing quartiles pairwise on socially inhibited or self-sacrificing. No other differences were found among the remaining subscales.

<table>
<thead>
<tr>
<th>TABLE 4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Direct Effect Measures, OLS Regression of Child Abuse on Binge Drinking and Binge Drinking on Interpersonal Problems</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Childhood Abuse</td>
</tr>
<tr>
<td>Sex</td>
</tr>
<tr>
<td>Men (reference)</td>
</tr>
<tr>
<td>Women</td>
</tr>
<tr>
<td>Race</td>
</tr>
<tr>
<td>White (reference)</td>
</tr>
<tr>
<td>Black</td>
</tr>
<tr>
<td>Hispanic</td>
</tr>
<tr>
<td>Asian</td>
</tr>
<tr>
<td>Other/Multiracial</td>
</tr>
<tr>
<td>Relationship Status</td>
</tr>
<tr>
<td>Single (Reference)</td>
</tr>
<tr>
<td>Significant Other</td>
</tr>
<tr>
<td>Married</td>
</tr>
<tr>
<td>Divorced</td>
</tr>
<tr>
<td>Binge Drinking</td>
</tr>
<tr>
<td>Intercept</td>
</tr>
<tr>
<td>$R^2$</td>
</tr>
<tr>
<td>Adj $R^2$</td>
</tr>
<tr>
<td>F Ratio</td>
</tr>
</tbody>
</table>

Note: $N = 170$. Source: Alcohol and Stress Survey. *p < .05, “*p < .01, “*p < .001 (two-tailed).
In the current sample, women reported higher levels of abuse than men. This is in agreement with the previous findings of the CDC (2012) that childhood abuse has a higher prevalence rate for women than for men. However, it is also important to consider that there may be differences in self-disclosure rates of abuse by sex. No significant differences were detected in generalized childhood abuse by race/ethnicity, but this may be the result of a lack of statistical power because of the relatively few minority participants surveyed. Additionally, no significant differences were detected between relationship statuses. Theoretically, being in a relationship was expected to mitigate the negative life outcomes of childhood abuse, but this was not the case in the current sample. As with race/ethnicity, there was a lack of statistical power for participants who reported being married or divorced, which can be expected from a sample of college students. The possible role of relationship status was likely not detected because of this lack of power.

In agreement with the first hypothesis, self-defeating beliefs partially mediated the relationship between childhood abuse and interpersonal problems. This indicates that negative automatic cognitions may exacerbate the psychosocial problems that arise from experiences of childhood abuse. The traumatic experience of childhood abuse likely impacts the development of schemas and perceptions of the world. The thought patterns developed from such a harmful experience may give rise to negative beliefs and expectations or self-defeating beliefs. From these self-defeating beliefs, a cycle of negative interpretation of self can be perpetuated, influencing interpersonal relationships. Self-defeating beliefs may affect an individual’s behavior, which is interpreted by the

### TABLE 5

<table>
<thead>
<tr>
<th>OLS Regression on Interpersonal Problems as a Function of Childhood Abuse, Sex, Race/Ethnicity, Relationship Status, and Alcohol Consumption</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Model 1</strong></td>
</tr>
<tr>
<td><strong>b</strong></td>
</tr>
<tr>
<td>Childhood Abuse</td>
</tr>
<tr>
<td>Sex</td>
</tr>
<tr>
<td>Men (reference)</td>
</tr>
<tr>
<td>Women</td>
</tr>
<tr>
<td>Race</td>
</tr>
<tr>
<td>White (reference)</td>
</tr>
<tr>
<td>Black</td>
</tr>
<tr>
<td>Hispanic</td>
</tr>
<tr>
<td>Asian</td>
</tr>
<tr>
<td>Other/Multiracial</td>
</tr>
<tr>
<td>Relationship Status</td>
</tr>
<tr>
<td>Single (reference)</td>
</tr>
<tr>
<td>Significant Other</td>
</tr>
<tr>
<td>Married</td>
</tr>
<tr>
<td>Divorced</td>
</tr>
<tr>
<td>Binge Drinking</td>
</tr>
<tr>
<td>Intercept</td>
</tr>
<tr>
<td>( R^2 )</td>
</tr>
<tr>
<td>( \text{Adj} R^2 )</td>
</tr>
<tr>
<td>( F \text{-Ratio} )</td>
</tr>
<tr>
<td>( F \text{-Change} )</td>
</tr>
</tbody>
</table>

Note. \( N = 170 \). Source: Alcohol and Stress Survey. \( * p \leq .05, ** p \leq .01, *** p \leq .001 \). (two-tailed).

Allen, Rogers, and Pearcey | Childhood Abuse and Interpersonal Problems

FALL 2017

PSI CHI
JOURNAL OF
PSYCHOLOGICAL
RESEARCH

COPYRIGHT 2017 BY PSI CHI, THE INTERNATIONAL HONOR SOCIETY IN PSYCHOLOGY (VOL. 22, NO. 3/ISSN 2325-7342) 215
Childhood Abuse and Interpersonal Problems | Allen, Rogers, and Pearcey

individual as negatively received by others (the world), indicative of some internal defect (the self), and expected to continue beyond the present (the future; Beck & Alford, 2009). These findings support that negative cognitions are an important piece of the puzzle for understanding the connection between childhood abuse and interpersonal problems in later life.

In regard to the second hypothesis, binge drinking was not a significant mediator of the relationship between childhood abuse and interpersonal problems. Although the research of Plant et al. (2007) indicated that childhood abuse is predictive of quantity of alcohol consumed per drinking session (binge-drinking behavior), the results of the current study were unable to validate or support this finding. However, this finding is important because it mirrors previous research by revealing that the exact role of alcohol in the relationship between childhood abuse and negative life outcomes is far from simple. It is possible that alcohol consumption as measured by AUDIT is not truly indicative of binge-drinking behavior. The AUDIT is used for the identification of alcohol use disorders through questions predominantly about frequency of drinking, not amount of drinks per drinking session. Thus, it may not capture the construct of binge drinking. In addition, participants who reported abstinence were included in the present analyses. Although living situation was not included in the present analysis, 30.2% of participants reported living with their parents or guardians. Participants who live with their parents may not exhibit binge-drinking behavior due to proximity and oversight from their parents. Finally, because many college students depend on their parents for financial support, binge-drinking behavior may be low in the current sample because of family pressure to maintain low levels of drinking and high academic performance.

Upon further examination of interpersonal problems, we discovered that those who experienced substantial child abuse histories were more likely to have higher scores on the domineering-controlling, vindictive-self-centered, and cold-distant subscales of the IIP-C-IRT. These findings are particularly salient in that the three subscales are integrally related within the interpersonal problems circumplex because they compromise the three factors that are high on control and low on affiliation. Further, although there appeared to be differences in the socially inhibited and self-sacrificing subscales among those with varying levels of child abuse, no single category of child abuse was significantly different than others upon post-hoc comparisons. Additional exploration in this area could lead to a better understanding of the effects of childhood abuse on interpersonal functioning.

The present study was pertinent because the negative social effects of childhood abuse appear to be compounded by self-defeating beliefs. Also, this finding contributed to the literature by investigating interpersonal problems experienced by abuse victims in later life. Avoiding the narrow focus on romantic relationships aids understanding of interpersonal problems as they occur in less intimate relationships. The current findings may help to inform practitioners that negative automatic thoughts are an important problem to address in patients who have a history of childhood abuse. Therefore, if a patient has a history of childhood abuse, a practitioner would be justified in investigating the patient’s current interpersonal functioning and patterns of cognitive attributions. With this information, a therapist may be better informed as to the patient’s current state of affairs and possible treatment options.

There were some limitations to the present study. First, several features of the sample may limit the generalizability of the findings. A larger sample would be preferred in order to narrow the beta coefficient confidence interval predictions. Sample size was in part hindered by filtering out participants for distractor questions and with nonresponse for variables of interest. However, dropping these participants was necessary to ensure the quality of the data. The present sample was also age-restricted because it only included college students. A low representation of racial minorities and participants in married or divorced relationships also limited generalizability. Second, the study utilized self-report for data collection. As such, standard concerns about the extent and validity of participants’ self-disclosure apply. In particular, participants provided their retrospective recall of childhood abuse. Reported levels of abuse, or absence thereof, were not verified via records or corroborating family accounts. Third, the present analyses were conducted using scale measures of childhood abuse, interpersonal problems, and binge drinking. The use of qualitative interview methodology could offer additional insights into the nature of these variables and the types of interpersonal problems experienced by those with a history of child abuse. Fourth, analysis of binge-drinking behavior was measured using select
binge-drinking questions from the AUDIT measure. The AUDIT predominantly measures frequency of alcohol consumption instead of quantity (binge drinking). Additionally, validity information on the self-defeating beliefs scale was lacking in the literature. Although the questions on the scale are theoretically sound representations of negative automatic thoughts, future researchers should examine the validity of the instrument.

Future researchers should consider supplementing the present results with data gathered through qualitative methods. Such information may further elucidate the nature of relationships between abuse, binge-drinking behavior, and interpersonal problems. Another area of focus would be to examine the exact nature of each individual negative cognitive area (i.e., the world, the self, and the future) as mediators of the relationship between childhood abuse and later life interpersonal problems. This would allow investigation of the relative importance of each negative cognitive area to the indirect development of later life interpersonal problems associated with childhood abuse. Although self-defeating beliefs mediate childhood abuse and interpersonal problems, each negative cognitive area may affect interpersonal problems as individual predictors. Such research would illuminate what specific types of negative automatic cognitions are developed after adverse life events (e.g., childhood abuse). Further, researchers should consider the use of latent constructs and structural equation modeling for the investigation of the complex indirect relationships between childhood abuse, self-defeating beliefs, alcohol use, and interpersonal problems. These methods would allow for greater understanding of the underlying constructs, their associated dimensionalities, and potential measurement variability.

Applied to issues of mental illness, greater understanding of the factors that influence the development of negative thought patterns would inform relevant treatments and programs. In order to fully investigate the role of binge drinking, future studies should utilize a scale designed solely for the measurement of binge-drinking behavior. Measuring frequency of alcohol consumption may be problematic because childhood abuse has been linked to binge-drinking behavior in particular (Plant et al., 2007). Future researchers should examine possible mediation effects of binge drinking in a larger sample size that has more individuals who report binge-drinking behavior. Further, it may be pertinent to investigate the effect of binge drinking on the relationship between childhood abuse and interpersonal problems in an older population (perhaps those who have had years to develop a pattern of binge-drinking behavior). College students may exhibit less consistency of drinking behavior than adults who have had time to develop patterns of alcohol use.

References


Childhood Abuse and Interpersonal Problems


U. S. Department of Health and Human Services, Administration for

FALL 2017

PSI CHI JOURNAL OF PSYCHOLOGICAL RESEARCH

218 COPYRIGHT 2017 BY PSI CHI, THE INTERNATIONAL HONOR SOCIETY IN PSYCHOLOGY (VOL. 22, NO. 3/ISSN 2325-7342)


Author Note. Jacob E. Allen, Department of Psychology, Kennesaw State University; Daniel T. Rogers, Department of Psychology, Kennesaw State University; Sharon M. Pearcy, Department of Psychology, Kennesaw State University.

Jacob Allen is now at Department of Public Health, Georgia State University.

This research was supported in part by Kennesaw State University’s Center for Young Adult Addiction and Recovery. Special thanks to all Psi Chi Journal reviewers for their support.

Correspondence concerning this article should be addressed to Jacob E. Allen, Department of Public Health, Georgia State University, Atlanta, GA 30303. Contact: jallen69@gsu.edu
In the American classic children’s story, *The Little Engine That Could* by Watty Piper, a little railroad train engine accepts the difficult task of pulling an unusually large load up a steep mountain railroad and eventually succeeds. In this story, the self-beliefs and positive emotional attitude of the train aided it in accomplishing its task. Similarly, research has indicated that the beliefs of students—particularly those of apprehension and self-efficacy—are connected to academic performance (Faigley, Daly, & Witte, 1981; Pajares & Johnson, 1994; Pajares, Johnson, & Usher, 2007; Sanders-Reio, Alexander, Reio, & Newman, 2014; Zorbaz, 2015). We designed the current study to understand the relationships between apprehension, self-efficacy, and performance in student writing at an undergraduate level, and the effects of mood on writing apprehension, writing self-efficacy, and writing performance.

In academic settings, writing assignments remain tools used by educators to evaluate the progress and learning of students. Researchers also continue to explore the benefits of other uses for writing in the classroom. Although much of writing stems from learning-to-write paradigms, there is also a movement toward using writing-to-learn methods to help students comprehend the material taught in academic classroom settings (Bangert-Drowns, Hurley, & Wilkinson, 2004; Emig, 2007; Reynolds, Thaiss, Katkin, & Thompson, 2012). Mynlieff, Manogaran, Maurice, and Eddinger (2014) found that writing assignments with metacognitive components in undergraduate biology courses had a significant impact on student learning.
research focusing on the role of writing in the learning process in recent years, it is important to consider the factors that influence the writing beliefs and performance of students.

An individual’s apprehension toward writing may come from various sources. Daly and Miller (1975) applied theories of anxiety and apprehension to writing. They conceptualized writing apprehension as the tendency of an individual to avoid writing situations because of writing anxiety. This type of apprehension may stem from a fear of evaluation, lack of confidence, and poor experiences in the past. To measure these concepts, they created the Writing Apprehension Test (WAT). Wachholz and Etheridge (1996) found that undergraduate students with high levels of writing apprehension reported negative experiences with writing that often involved failure, unrealistic expectations of what good writing looks like, and a lack of confidence in what they had learned from previous teachers. Conversely, those with low levels of writing apprehension reported more positive and successful experiences with writing, reasonable expectations of what good writing looks like, and believed that their writing skills would influence their success in social and professional roles in the future. Additionally, Zorbaz (2015) demonstrated that a negative relationship between levels of writing activity and writing apprehension, where high levels of writing activity such as consistently writing in a journal were present in individuals with low levels of writing apprehension. Additionally, there was a negative relationship between good reading habits and low levels of writing apprehension.

Although many of these sources of writing apprehension seem predictive of writing performance, research has provided conflicting evidence on whether writing apprehension is actually predictive of writing performance. More specifically, in one study, writers with high apprehension performed worse than their low apprehension counterparts in all but two assessments. In instances where the highly apprehensive individuals scored better, the assessments were narrative essays (Faigley et al., 1981). Sanders-Reio et al. (2014) also demonstrated that writers with high writing apprehension received lower grades on writing assignments scored on grammar. However, their results using Daly and Miller’s (1975) WAT did not show a relationship between writing apprehension and writing performance. Similar to Sanders-Reio and her colleagues, Pajares and Johnson (1994), who also used the WAT, did not find writing apprehension to be predictive of writing performance. These conflicting results indicate a need for further research in order to make the relationship between writing apprehension and writing performance clearer. However, although this relationship is unclear, both Faigley et al. (1981) and Sanders-Reio et al. (2014) indicated that an individual’s writing self-efficacy beliefs were predictive of writing performance.

Prior to understanding the role of self-efficacy in terms of writing, it is important to understand the basic theoretical perspective on self-efficacy beliefs in general. Bandura (1977) defined an individual’s perceived self-efficacy as a judgment of whether or not one’s ability to accomplish a specific task or skill would produce success. It is a judgment of one’s own capabilities. Two decades later, Bandura (1997) theorized that an individual’s self-efficacy comes from four main areas: mastery experiences, where individuals experience success and attainment with a certain task, skill, or situation; vicarious experiences, where an individual witnesses another individual model the successful accomplishment of a task, skill, or situation; verbal persuasion, where an individual is convinced verbally of one’s own capabilities; and, physiological and affective states, where mood, stress, and arousal influence one’s own perceived self-efficacy. He also found that self-efficacy beliefs affected performance at various tasks.

These self-efficacy beliefs also exist within the context of writing. Shell, Murphy, and Bruning (1989) developed two instruments to measure self-efficacy beliefs regarding specific reading and writing tasks and reading and writing skills. The instruments were based on Bandura’s earlier conceptualizations of self-efficacy (Bandura, 1977, 1997). They found that high self-efficacy and outcome expectancy beliefs for both reading and writing were correlated with high achievement in undergraduate students. Their research supports the findings of Bandura (1997) in that self-efficacy beliefs in writing also impact performance in the same way that other beliefs do, and support the relationship between reading beliefs and performance with writing beliefs and performance (Zorbaz, 2015). Also, Pajares et al. (2007) postulated that, utilizing the model of Bandura’s sources of self-efficacy, mastery experience accounted for the greatest proportion of the variance in the writing performance of elementary, middle, and high school students. This finding suggests that students with high sources of writing self-efficacy gained their self-efficacy beliefs through mastery.
The Effects of Mood on Writing

The findings of the studies reviewed thus far support a connection between writing apprehension, self-efficacy, and performance. Writing apprehension and self-efficacy are perceived through self-judgments. However, these self-perceptions and self-judgments may be influenced by different factors within certain contexts. Affect may be one of these factors that could distort individuals’ abilities to accurately make judgments on themselves. Bandura (1997) stated that one’s affective state is one of the sources of an individual’s self-efficacy perception. Furthermore, because apprehension is a hesitation stemming from feelings of anxiety, changes in one’s affective state may directly impact one’s writing apprehension. For these reasons, a mood state may influence the writing apprehension, self-efficacy, and performance of an individual and student.

Mood states influence the perceptions and judgments of one’s own well-being. Schwarz and Clore (1983) found that individuals perceive their well-being differently depending on the weather. Individuals rated their well-being more positively when asked on a sunny day; individuals rated their well-being more negatively when asked on a rainy day. The sunny or rainy weather induced a positive or negative mood state, which altered the individuals’ perceptions of their well-being. Furthermore, they found that participants perceived themselves as happier after writing about a positive life event, and perceived themselves as sadder after writing about a negative life event. In these ways, a positive or negative mood state, induced by either the weather or the writing task, altered individuals’ perceptions and judgments of their own affective experience. If mood influences self-perceptions, then they may also alter individuals’ perceived confidence in their own writing abilities.

Perceptions of value and risk are also influenced by mood state. Individuals exposed to happy faces were more likely to pour and drink alcohol; those who were exposed to angry faces were less likely to pour and drink alcohol. The perceived value of the alcohol increased when primed with positive facial expressions and decreased when primed with negative facial expressions (Winkielman, Berridge, & Wilbarger, 2005). Johnson and Tversky (1983) obtained similar results by presenting participants with a fabricated news story involving death, thus creating a negative affective state. Participants were more likely to rate their perceived worry or concern for 18 different types of deaths higher than those who were presented.

experiences, or previous successes with writing assignments or projects. Interestingly, their results also showed a relationship between self-efficacy and writing anxiety. Higher levels of writing anxiety correlated with lower levels of self-efficacy. These results matched those reported in other studies (Pajares & Johnson, 1994; Sanders-Reio et al., 2014). Although students with high self-efficacy experienced mastery experiences, perhaps a lack of mastery experiences might have contributed to the presence of high anxiety in those students with low self-efficacy.

Writing self-efficacy beliefs may lead students to seek help in certain situations. Stewart, Seifert, and Rolheiser (2014) established that students with higher levels of writing self-efficacy are more likely to seek out help from others. Conversely, Williams and Takaku (2011) researched the connection between the help-seeking behavior, as defined by writing center visits, and self-efficacy beliefs of international English as a second language (ESL) students. They found that, although ESL students had lower self-efficacy beliefs than domestic students, the ESL students were more likely to seek out help for writing than their domestic counterparts. They explained this finding as the ESL students’ recognition of their own need for writing help in a new language. Although Williams and Takaku discussed the differences between ESL and domestic students, Stewart et al. (2014) did not address this possible difference. These findings would suggest that self-efficacy beliefs may produce varying effects within different populations.

Although the connection between help-seeking behaviors and self-efficacy is unclear, high levels of writing self-efficacy are positively correlated to high levels of writing performance. Pajares and Johnson (1994) used the same measure for writing self-efficacy that Shell et al. (1989) used in their study, the Writing Self-Efficacy Test (WSET). Pajares and Johnson (1994) found that an individual’s confidence in specific writing skills was predictive of writing performance. However, one’s confidence in completing a writing task was not predictive of writing performance. In this instance, the breakdown of the writing self-efficacy scale into two components, skills and tasks, was particularly useful in highlighting this specific difference in writing self-efficacy beliefs. Sanders-Reio et al. (2014) supported the results of Pajares and Johnson (1994) and also found that self-efficacy beliefs predicted writing performance, with high levels of self-efficacy predicting high levels of writing performance.
with an average news story not involving death. Also, these participants in the negative affect group were more likely to expect themselves to die from one of the 18 different types of deaths. Like with Schwarz and Clore (1983), mood influences one’s judgment or perception of risk. When students encounter a writing assignment or task, they may also be more likely to experience worry or concern over an impending evaluation or expected failure. Additionally, the value students place on the writing assignment or on the grade eventually obtained may also differ depending on the students’ affective states. In these ways, students’ perceptions and judgments of value and risk involved in a writing situation may influence their writing apprehension or self-efficacy.

Although mood may impact writing apprehension and writing self-efficacy, it may also directly impede individuals’ process of generation as they write. Gasper (2004) found that mood states impact the ability to generate new ideas. Gasper utilized a mood manipulation, which required participants either to write about a positive or negative life event to manipulate their mood (Schwarz & Clore, 1983). After this writing activity, participants then completed crossword puzzles in which each word was connected in some way to flying. Next, participants were told to write down as many things they could think of that could fly. She found that participants exposed to a sad mood prime thought of fewer new ideas than those participants exposed to a happy mood prime. This suggests that sad mood states inhibit the ability of the mind to generate new ideas and concepts represented by words. Frederickson (2001) described this effect of emotions on the mind’s ability to generate new ideas and concepts as the broaden-and-build theory. This theory posits that negative emotions narrow momentary thoughts and potential behaviors, or thought-action repertoires, while positive emotions broaden, or increase the availability of, thought-action repertoires. Essentially, negative emotions narrow an individual’s attention onto a limited number of ideas, and positive emotions broaden an individual’s attention to include many different types of ideas (Frederickson, 2001; Frederickson & Branigan, 2005). From the perspective of both Frederickson and Gasper, affective states impact an individual’s ability to generate new words and ideas. This generativity is integral in writing, creating, and organizing sentences and ideas (Crossley, Muldner, & McNamara, 2016; Flower & Hayes 1981), and so the mood state of the writer may inhibit access to ideas, and therefore the ability to translate thoughts into a cohesive and comprehensible narrative or argument.

Although recent researchers have sought to understand the connections between writing self-efficacy, apprehension, and performance and the influence of mood on perceptions, judgments, and generativity, little known research has combined both of these approaches. Furthermore, the relationships among writing apprehension, writing self-efficacy, and writing performance requires further exploration. The current study continued to explore the relationship between writing self-efficacy, apprehension, and performance, and examined if the mood state of an individual impacts that person’s writing beliefs and performance. Participants in this study experienced a process of mood induction prior to completing the WAT, WSET, and a 30-minute timed essay to measure writing apprehension, writing self-efficacy, and writing performance, respectively.

The main hypothesis was that a positive mood prime would produce lower levels of writing apprehension and higher levels of writing self-efficacy and writing performance, while a negative mood prime would result in higher levels of writing apprehension and lower levels of writing self-efficacy and performance. Additionally, it was hypothesized that, consistent with previous research, writing self-efficacy would be positively related to writing performance, but writing apprehension would not have a relationship with writing performance. Writing apprehension and writing self-efficacy would also be negatively related to each other.

We hoped to further the understanding of researchers and educators on the influences and implications of mood states, writing apprehension, self-efficacy, and performance. Findings on the impact of mood on these aspects of writing may help students, teachers, and curriculum coordinators understand the role of mood in the writing process and in both learning-to-write and writing-to-learn settings in academia.

Method

Participants
One hundred seventeen undergraduate students ($\bar{X} = 19.85, \ SD = 0.29$) enrolled in introductory psychology and English courses at a medium-sized southwestern university in the United States were recruited. Among participants, 71 were women and 46 were men; 82% identified as European American, 9% as Hispanic, and 9% of the sample
were small proportions of African American, Asian, and Pacific Islander. Participants were all proficient English language speakers and writers who were recruited through class visits and flyers containing the research details. Participants in the psychology courses received two credits toward their four required research participation assignments. All participants were entered into a drawing to receive a gift card to the campus bookstore: three $25 gift cards, and one $75 gift card.

Materials

Writing apprehension. Writing apprehension was measured using the WAT created by Daly and Miller (1975). The 26 questions on this instrument provide statements regarding teacher evaluation, emotional feelings, and self-beliefs in regard to writing. Responses are given on a 5-point Likert scale from 1 (strongly agree) to 5 (strongly disagree). Of the 26 questions, 12 are reverse coded. The total score combines the sum from responses to the eight coded questions with the sum from responses from the 12 reverse-coded questions, resulting in a total sum between 20 and 100. The WAT has been shown in past research to have adequate reliability (1975). According to their study, the split half technique produced a reliability coefficient of 0.94, which indicates high reliability. They also found that those with high writing apprehension selected professions that were less writing intensive, and those with low writing apprehension selected professions that were more writing intensive. These results suggest good predictive validity. Our sample also demonstrated adequate reliability (α = .94).

Writing self-efficacy. The writing self-efficacy of each participant was measured using a writing self-efficacy test developed by Shell et al. (1989). The writing self-efficacy scale is broken down into self-efficacy for writing skills and self-efficacy for writing tasks. The respondents are required to rate their confidence level to effectively communicate in a given writing situation or to effectively use a specific writing skill on a scale from 0% to 100%. The writing tasks section of this measure has 16 questions, and the writing skills section has eight questions. The WSET has been shown to be reliable (Shell et al., 1989). In our sample, a reliability of .87 was observed using Cronbach’s alpha for the task subscale and .93 for the skills subscale. This instrument, for measuring writing self-efficacy, has demonstrated good predictive validity. For example, high self-efficacy, as measured by this instrument, shared a positive relationship with writing performance (Pajares & Johnson, 1994; Sanders-Reio et al., 2014). These results matched the results found by Bandura (1997) predicting high self-efficacy beliefs associated with high performance.

Timed essay. Each participant responded to the following writing prompt: “What do you believe to be the qualities of a successful student?” Shell et al. (1989) used a similar question, except the question asked for the successful qualities of a teacher rather than a student. This topic change made the essay prompt more relatable for the student. This essay was a 30-minute timed essay, and was graded by three graders, all tutors at the university writing center. All writing center tutors are required to take a training course and receive additional monthly training. Because of the consistent training among the graders, the writing center tutors were seen as qualified to grade the essays. All graders used a grading rubric to improve consistency (see Appendix).

Procedure

After the Institutional Review Board of Southern Utah University gave its approval (#14-110215), participants were randomly assigned to one of three conditions: positive mood prime, negative mood prime, and neutral mood prime. Within the three mood groups, participants were assigned to one of three orders for the presentation of the dependent variables to control for order effects: (a) writing apprehension, writing self-efficacy, and writing performance; (b) writing self-efficacy, writing performance, and writing apprehension; (c) writing performance, writing self-efficacy, and writing apprehension. Before beginning the study, participants signed an informed consent document and were given initial instructions.

Next, participants completed a writing task intended to alter their mood states. They were randomly assigned to one of three writing conditions and asked to write about: (a) a sad or negative event, (b) a happy or positive event, or (c) a set of instructions explaining how to tie a shoe. These three writing subjects represented the negative, positive, and neutral mood manipulations respectively. Research conducted by Schwarz and Clore (1983) shows these writing events to be effective activities for mood manipulation. Participants then answered a few self-report questions as a mood manipulation check. Specifically, participants were asked describe their feelings of emotionality by completing a discrete emotions checklist styled
like the Positive and Negative Affect Schedule, and then to report their mood state by answering three bipolar questions anchored with happy to unhappy, positive to negative, and pleasant to unpleasant.

After completing the manipulation check, participants completed the WAT, WSET, and the 30-minute timed essay. The order for the next three tests varied according to the order group to which the participant was assigned, as mentioned earlier. All completed written works, tests, and short essays were given to the researcher conducting the session. Before leaving the experiment, participants were debriefed and thanked for their participation.

The short essays were given to three graders, who individually graded all of the short essays written by participants. The graders were blind to the scores of the other graders. The scores from the three graders were averaged together for an overall essay score. To determine the interrater reliability of the essay graders, an intraclass correlation (ICC) was conducted. The average measure ICC was .693 with a 95% confidence interval from .511 to .801, \( F(116, 232) = 4.025, p < .01 \).

**Results**

**Planned Analyses**

Three one-way independent Analyses of Variance (ANOVAs) were used to determine the effects of mood on each of the three dependent variables: writing apprehension, writing self-efficacy, and writing performance. For all three one-way independent ANOVAs, the between-subjects factor was the mood prime group—with the positive, neutral, and negative mood groups as the levels. In the first one-way independent ANOVA, the dependent variable was the score of the participants on the WAT. The results were not significant, \( F(2, 115) = 0.87, p = .42, \eta^2 = .02 \). For the second one-way independent ANOVA, a composite score of the participants on the skills and tasks scales on the WSET was the dependent variable; this combined self-efficacy score is consistent with previous research (Pajares & Johnson, 1994). The results were not significant, \( F(2, 115) = 1.13, p = .33, \eta^2 = .02 \). Lastly, the overall score for writing performance (i.e., the sum total of each grade on the rubric given by the essay graders) was the dependent variable in the final one-way independent ANOVA. The results were also not significant, \( F(2, 115) = 0.07, p = .94, \eta^2 = .00 \). See Table 1 for the means and standard deviations for these three ANOVA tests.

The effectiveness of the mood prime was analyzed by examining the self-report scores for the questions measuring positive emotions from the mood manipulation check by a one-way independent ANOVA. The test on the positive emotions indicated significant results, \( F(2, 115) = 14.02, p < .001, \eta^2 = .20 \), and a post-hoc comparison using the Scheffe test indicated a significant difference \((p < .001) \) between the positive prime \((M = 10.16, SD = 7.35) \) and the negative prime \((M = 4.08, SD = 3.65) \), and a significant difference \((p = .005) \) between the positive prime and the neutral prime \((M = 6.33, SD = 3.38) \). However, the difference between the neutral and negative prime was not significant \((p = .151) \). The one-way ANOVA for the neutral emotions showed a significant difference between the mood prime groups, \( F(2, 114) = 13.42, p < .001, \eta^2 = .19 \), and a post-hoc comparison with the Scheffe test produced a significant difference \((p < .001) \) between the positive mood prime \((M = 1.76, SD = 0.53) \) and negative mood prime \((M = 5.21, SD = 0.52) \), and a significant difference \((p < .001) \) between the negative mood prime and neutral mood prime \((M = 2.02, SD = 0.52) \). There was not a significant difference between the positive and neutral mood primes \((p = .939) \). A one-way ANOVA examining the negative emotions approached a significant difference between the mood prime groups, \( F(2, 115) = 3.01, p = .053, \eta^2 = .05 \). The means and standard deviations for the questions measuring neutral emotions for the mood prime groups are as follows: positive prime \((M = 4.21, SD = 0.52) \), neutral prime \((M = 5.48, SD = 0.50) \), and negative prime \((M = 5.92, SD = 0.50) \).

**TABLE 1**

<table>
<thead>
<tr>
<th>Mood Groups Across the Scores</th>
<th>( M )</th>
<th>( SD )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Writing Apprehension</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>67.68</td>
<td>19.19</td>
</tr>
<tr>
<td>Neutral</td>
<td>73.40</td>
<td>21.13</td>
</tr>
<tr>
<td>Negative</td>
<td>69.85</td>
<td>17.43</td>
</tr>
<tr>
<td>Writing Self-efficacy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>71.46</td>
<td>14.46</td>
</tr>
<tr>
<td>Neutral</td>
<td>67.77</td>
<td>13.74</td>
</tr>
<tr>
<td>Negative</td>
<td>67.14</td>
<td>12.65</td>
</tr>
<tr>
<td>Writing Performance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>19.85</td>
<td>2.37</td>
</tr>
<tr>
<td>Neutral</td>
<td>20.04</td>
<td>1.97</td>
</tr>
<tr>
<td>Negative</td>
<td>19.94</td>
<td>2.53</td>
</tr>
</tbody>
</table>

Note: The means and standard deviations for these three dependent measures reflect values for the raw scores.
The Effects of Mood on Writing | Vanhille, Gregory, and Corser

SD = 0.51). See Table 2 for the means and standard deviations of the reported emotionality of participants across mood prime groups.

A Pearson’s Correlation was used to test the relationships among writing apprehension, writing self-efficacy, and writing performance. Writing apprehension and writing self-efficacy shared a significant moderate negative relationship, r(115) = -0.56, p < .01. However, writing performance did not share a significant relationship with either writing apprehension or writing self-efficacy, r(115) = -0.17, p = .07 and r(115) = -0.14, p = .13, respectively. See Table 3 for the results of this test.

Exploratory Analyses

The essay prompts for the timed essay told students to “feel free to brainstorm or outline prior to writing, but [to] remain aware of how much time [they had] left to complete [their] essay.” Many participants used some form of prewriting, either with a concept map or with only words, to brainstorm and organize information. Although research questions regarding prewriting techniques were not an intended purpose of this study, the frequency of prewriting techniques by participants provided an opportunity for exploratory analyses. To determine whether the use of prewriting, and the type used, improved exam scores, a one-way independent ANOVA was conducted. The between-subjects factor was the presence of prewriting with three levels: no prewriting (N = 58), prewriting with words only (N = 36), and prewriting using a concept map (N = 23). The dependent variable was the overall score on the essay. The test showed a significant difference among the prewriting groups for the overall essay scores, F(2, 115) = 4.66, p = .011, η² = .08. A post-hoc comparison using the Scheffe test showed a significant difference (p = .018) between the no prewriting group (M = 19.33, SD = 0.29) and the words only prewriting group (M = 20.69, SD = 0.57). There was not a significant difference (p = .177) between the no prewriting group and concept map group (M = 20.35, SD = 0.461), nor was there a significant difference (p = .850) between the words only and concept map groups.

Because prewriting techniques often seem to help organizational structures in writing, an exploratory one-way independent ANOVA was conducted to test the effects of prewriting on organization in the essay. The between-subjects factor was the presence of prewriting, and the dependent variable was the organization score on the essay rubric. The test indicated a significant difference among the prewriting groups, F(2, 115) = 3.40, p = .037, η² = .06. A post-hoc comparison using the LSD test showed a significant difference (p = .024) between the word group (M = 4.24, SD = 0.12) and the no prewriting group (M = 3.89, SD = 0.10). There was not a significant difference (p = .965) between the word group and the concept map group (M = 4.23, SD = 0.15), nor a significant difference between the concept map group and the no prewriting group (p = .056).

Discussion

This study examined the effects of mood on writing apprehension, writing self-efficacy, and writing performance. The results of the one-way independent ANOVAs measuring the differences between mood groups for the WAT, WSET, and timed essay were not significant. Furthermore, the results of the one-way independent ANOVAs regarding the positive and negative mood induction indicated that the mood

<table>
<thead>
<tr>
<th>TABLE 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Means and Standard Deviations of Reported Emotionality Across Mood Prime Groups</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Positive Emotions</td>
</tr>
<tr>
<td>Positive Prime</td>
</tr>
<tr>
<td>Neutral Prime</td>
</tr>
<tr>
<td>Negative Prime</td>
</tr>
<tr>
<td>Neutral Emotions</td>
</tr>
<tr>
<td>Positive Prime</td>
</tr>
<tr>
<td>Neutral Prime</td>
</tr>
<tr>
<td>Negative Prime</td>
</tr>
<tr>
<td>Negative Emotions</td>
</tr>
<tr>
<td>Positive Prime</td>
</tr>
<tr>
<td>Neutral Prime</td>
</tr>
<tr>
<td>Negative Prime</td>
</tr>
</tbody>
</table>

Note. The means and standard deviations for these three dependent measures reflect the self-reported emotional state of participants.

<table>
<thead>
<tr>
<th>TABLE 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson’s Correlation for Testing the Relationships Among WAT, WSET, and WP</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>WAT</td>
</tr>
<tr>
<td>WSET</td>
</tr>
<tr>
<td>Overall WP</td>
</tr>
</tbody>
</table>

Note. WAT = Writing Apprehension Test; WSET = Writing Self-Efficacy Test; WP = timed essay. p < .01.
The Effects of Mood on Writing

Vanhille, Gregory, and Corser | The Effects of Mood on Writing

prime successfully influenced participants’ moods. These findings paired together seem to suggest that mood did not affect writing apprehension, writing self-efficacy, and writing performance. However, participants were likely made aware of their emotional state when they received the mood manipulation check immediately following the mood prime. This awareness might account for the lack of significant differences among the mood prime groups across the WAT, WSET, and overall writing performance score because it might have blunted or eliminated the effects that mood might have had on writing apprehension, writing self-efficacy, and writing performance. The original hypothesis that a positive mood prime would produce lower levels of writing apprehension and higher levels of writing self-efficacy and writing performance, although a negative mood prime would result in higher levels of writing apprehension and lower levels of writing self-efficacy and performance, was not supported by the results. However, the results of our study should not be interpreted as showing evidence that mood does not affect writing apprehension, writing self-efficacy, or writing performance. More specifically, mood had no effect on writing when participants are made aware of their emotional state. It is possible that the original hypothesis may be supported when individuals are not consciously aware of their emotional state.

The lack of support for mood effects are also particularly surprising for self-efficacy when viewing an individual’s affective and physiological state as a source of self-efficacy (Bandura, 1997). When self-efficacy is low, confidence in successfully completing a task leads to a more pronounced physiological (autonomic) arousal, or a greater physiological state. When self-efficacy is high, confidence in completing a task results in relatively lower physiological arousal due to lesser activation of the autonomic nervous system. Additionally, high self-efficacy leads to greater engagement, effort investment, and increased persistence due to the confidence that one can complete a task (Bandura, 1977, 1981). Given that a negative mood should result in more pronounced physiological arousal than a positive mood, relatively lesser engagement, and investment as well as decreased persistence, it is surprising to see that mood had no significant effect on writing self-efficacy in the current study. However, an acute mood prime may not be sufficiently powerful at significantly affecting deeply held academic self-efficacy beliefs, which may explain why the current mood manipulation design did not significantly affect writing self-efficacy beliefs.

Another part of the study aimed to identify the relationship among writing apprehension, writing self-efficacy, and writing performance. The significant moderate negative relationship between writing apprehension and writing self-efficacy coincides with the findings of other researchers and supports part of the hypothesis regarding writing beliefs (Pajares & Johnson, 1994; Sanders-Reio et al., 2014). The absence of a significant relationship between writing performance and apprehension, and between writing performance and self-efficacy has also been substantiated by other research (Pajares & Johnson, 1994; Sanders-Reio et al., 2014). Practical implications of the negative relationship observed in the current study may indicate that educators including tutors should, in addition to understanding the cognitive processes associated with writing, understand the psychological impacts of efficacy beliefs on writing and tailor their instruction to include these psychological effects. Given that greater self-efficacy leads to increases in engagement, investment, and persistence (Bandura, 1981), educators should strive to enhance student self-efficacy through appropriate constructive feedback. This feedback should include explicit information conveying what the student has done well, areas in which the student demonstrates weakness, along with ways to improve these areas of perceived weakness by focusing on changeable behaviors (Bandura, 1977). As students develop and strengthen their writing skills, educators should explicitly incorporate longitudinal self-comparison as evidence of improvements in areas that formerly were not as strong. Given that personal mastery experiences are the most powerful sources of self-efficacy (Bandura, 1977), educators should incorporate these comparisons between past performance and current performance to help empower students with greater confidence in their writing abilities.

Although not an original research question, the significant difference in overall essay scores between the no prewriting group and the prewriting using only words group adds to a different discussion than the main topic of this study. Much of the prewriting in the word prewriting group involved making an outline, or at least ordering the supporting points of their argument. Drawing from the cognitive process model of Flower and Hayes (1981), the word prewriting technique might have aided participants in translating their ideas into a cohesive argument. Furthermore, the difference between the organizational score for the two prewriting groups versus the no
The Effects of Mood on Writing  
Vanhille, Gregory, and Corser

Prewriting group would also suggest that prewriting leads to better organized papers. Although, just as probable, individuals who tend to write more organized papers may also be more likely to utilize prewriting techniques. However, it should again be noted that the roles and effects of prewriting were not included in the original research questions or hypotheses, thus the assumptions included above are purely speculative due to the exploratory nature of the effects of prewriting observed in the current study. Therefore, it cannot be assumed that the cognitive processes involved with prewriting caused increased cohesiveness in prewriters’ arguments.

Limitations
Because participants responded to three different measures and continued participating for up to 45 minutes after the end of the exposure to the mood prime, it is possible that participants experienced testing fatigue. The use of writing as the mood prime might also have contributed to testing fatigue, making the later measures less accurate in describing their performance or self-perceptions. Also, because the testing continued for 45 minutes after the exposure to the mood prime, the effects of the mood prime might not have been as potent after the first testing measure. This diminished potency may also help explain the lack of significant effects of mood on self-efficacy beliefs.

The lack of a filler task and participants’ awareness of their mood state might have led to the nonsignificant effects of mood. In some research, participants engage in a filler task after the mood manipulation has occurred (Gasper, 2004; Schwarz & Clore, 1983). This filler task distracts participants from the mood manipulation just used and allows time for the new mood to fully develop. With this in mind, it is likely that participants in the present study were alerted to their mood state by the mood manipulation check before they completed the WAT, WSET, and timed essay. As mentioned previously, it is possible that the affective states of individuals may in fact influence writing apprehension, writing self-efficacy, and writing performance when those individuals are not made consciously aware of their mood through the manipulation check prior to completing the WAT, WSET, and timed essay. This explicit alerting, combined with the absence of a filler task, might have contributed to the results of the study.

Another limitation may reside in the agreement among the essay graders. The researcher and the essay graders only met once for an hour and a half to review sample papers and the grading rubric. Although the graders are all tutors at the university writing center and receive the same training through the writing center, this training focuses on peer-tutoring and not on grading. Therefore, the lack of training on grading essays with the rubric might have contributed to the reliability score for the graders. The training of the writing center tutors and their working hours spent analyzing papers certainly qualifies the tutors to be able to evaluate a paper for the purposes of tutoring, but there is not sufficient training in grading assignments. If there had been more agreement among the essay graders, the results of the mood manipulation and multiple correlation might have been different with more reliable scores.

Future Recommendations
Researchers in the future should consider conducting three smaller experiments to test the effects of mood on writing apprehension, writing self-efficacy, and writing performance individually because this method may decrease test fatigue. Also, these smaller studies should administer the mood manipulation check at the end of the experiment and insert filler tasks between the exposure to the mood prime and the desired measure such as the WAT or WSET. By addressing these design issues, researchers may indeed find that mood affects writing apprehension, writing self-efficacy, and writing performance.

Other researchers should consider multiple meetings with the essay graders to review sample essays and discuss the essay rubric to increase consistency among graders. In fact, during the course of the meeting, the graders made adjustments to the grading rubric. It is important to establish a well-designed rubric, which encourages consistency, prior to training the essay graders. More than one meeting is advised to establish a tighter agreement among the graders.

Researchers studying writing should explore the influence of prewriting on the performance of individuals. This exploration may take the form of studies examining the factors that lead to prewriting in the first place, or it may seek to understand how different types of prewriting may influence writing performance in certain ways such as in specific areas of writing as designated by the grading rubric. Furthermore, the field may also benefit from research that draws on the broaden-and-build theory to examine the impact of affective states on generativity in prewriting (Frederickson, 2001; Frederickson & Branigan, 2005).
Although this study did not yield results indicating significant effects of mood on writing apprehension, self-efficacy, and performance when individuals are aware of their mood state, it does address important design components for testing affective influences on these aspects of writing. There is also room for further explanations regarding how self-efficacy and apprehension in writing develop, and perhaps on how these factors may influence prewriting behavior. As researchers continue to study various facets of writing, the purposes of writing instruction and applications of diverse methods should be reevaluated to integrate new findings to improve skills and experiences of students in writing.

References

Author Note. Jared VanHille, Department of Psychology, Southern Utah University; Bradley Gregory, Department of Psychology, Southern Utah University; Grant Corser, Department of Psychology, Southern Utah University.

Jared VanHille is now at Department of English, Brigham Young University.

This research was supported in part by grants from the Undergraduate Research and Scholarship Program at Southern Utah University and the Southern Utah University Student Association Senate. The researchers acknowledge the counsel provided by Lynn White at Southern Utah University regarding the general design of the study. Special thanks to all Psi Chi Journal reviewers for their support.

Correspondence concerning this article should be addressed to Jared VanHille and jaredvanhille@byu.net.
## APPENDIX

### Grading Rubric for the Timed Short Essay

<table>
<thead>
<tr>
<th>Punctuation and grammar</th>
<th>5 (superior)</th>
<th>4</th>
<th>3 (good)</th>
<th>2</th>
<th>1 (poor)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Three or fewer grammar or punctuation errors.</td>
<td>At least two or three grammar and/or punctuation errors occur in most paragraphs.</td>
<td>Grammar and punctuation errors occur in most sentences.</td>
<td>The grammar and punctuation errors drastically impede the readability of the essay.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Style</td>
<td>The essay has a wide variety of sentence structures and uses effective word choice.</td>
<td>Lack of sentence variety. Word choice is average.</td>
<td>The lack of sentence variety detracts from the overall flow and argument. Incorrect word choice is used.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evidence and support</td>
<td>The supporting points to the thesis are backed up by specific examples.</td>
<td>The evidence and support is more a detailed elaboration rather than concrete examples, but the elaboration connects to the point.</td>
<td>The supporting points to the thesis do not have any evidence or support backing the claims.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Organization</td>
<td>The thesis, supporting points, and conclusion are structured logically. Ideas within the paragraphs and throughout the paper connect together.</td>
<td>The paragraphs and/or sentences are structured reasonably well. Transitions are jarring from time to time.</td>
<td>The paragraph and sentence organization are basically nonexistent, making the argument and reasoning difficult to follow.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thesis and purpose</td>
<td>The essay has a clear and concise thesis and a strong focus throughout.</td>
<td>The paper may have a thesis and/or overall focus, but the writer does not articulate it well or continually stick to it.</td>
<td>The essay does not have a thesis, and the paper is disjointed and not cohesive.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### The Effects of Mood on Writing

| Vanhille, Gregory, and Corser |
The Interaction Effect of Facial and Vocal Attraction on Overall Perceived Attractiveness

Sheila Bonnough and Erin Moore*
Stetson University

ABSTRACT. Physical attractiveness helps people choose mates. Although many individuals have faces and voices considered equally attractive, this is not always true, in that some people have one feature that is perceived as more attractive than the other. Zuckerman and Sinicropi (2011) found that, when the levels of facial and vocal attractiveness are mismatched, participants are more likely to rate the mismatch as less attractive than matched attractiveness due to the dissonance of believing that the unattractive feature does not belong with the attractive feature. Whereas Zuckerman and Sinicropi examined the disappointment people felt and personality traits associated with targets with different levels of physical and vocal attractiveness, the present study sought to determine how overall attractiveness is influenced by having targets with mismatched faces and voices. It also sought to identify differences in perceived facial and vocal attractiveness between sexes. College students (N = 112) were recruited using probability sampling. They viewed 14 face/voice combinations and rated them on overall attractiveness. There was a significant four-way interaction between participant sex, target sex, target facial attractiveness, and target voice attractiveness (p = .018, ηp² = .05). Mismatched face/voice combinations were rated more attractive than matched unattractive face/unattractive voice and less attractive than matched attractive face/attractive voice. Attractive face/unattractive voice pairs were rated more attractive than unattractive face/attractive voice pairs. Although women’s ratings were higher than men’s, the degree of difference varied by target. These findings demonstrated that, when evaluating individual attractiveness, faces are given more consideration than voices.

A ttraction is the force that draws people together for the purpose of procreation, and physical attraction strongly influences mate choice (Fraccaro et al., 2010; Gallup & Frederick, 2010; O’Connor et al., 2012; Puts, Jones, & DeBruine, 2012; Saxton, Caryl, & Roberts, 2006). Evidence has suggested that attractiveness may be an underlying signal of health including how long individuals will live and their fertility (Gallup & Frederick, 2010), which indicates that perceived differences in attractiveness could serve an evolutionary advantage in identifying healthier mates. Attraction extends beyond mate selection, however, because people are able to perceive the attractiveness of others of the same sex. This helps to determine competition and the hierarchy of competing mates. It has been found that people who are considered more attractive are also believed to have more favorable personality traits (Zhao, Zou, Shi, & Zhang, 2015; Zuckerman & Sinicropi, 2011). However, attractiveness can be indicated by features besides the physical appearance of one’s face and body; one’s voice, for example, may be an indicator of attractiveness. The association of favorable characteristics with attractiveness extends beyond physical attractiveness to perceived vocal...
attractiveness. This can be attributed to the halo effect where a person who is judged as having an attractive quality is then considered to have other favorable characteristics (Jones et al., 2001; Rhodes, 2006). It is important to consider how the perceived attractiveness of multiple physical aspects of people—both their faces and voices—are weighted by potential mates and competitors.

Previous research has identified which factors are associated with greater perceived physical attractiveness when evaluating people’s faces, which differ depending on a person’s sex. Men generally report greater attraction to women with more feminine, delicate features and large eyes (Fraccaro et al., 2010; Puts et al., 2012). Tough faces, broad chins, and heavy brows are considered by women to be more attractive features in men (Puts et al., 2012; Saxton et al., 2006). Research has found that perceived facial attractiveness is independent of culture, sex, ethnicity, and age, meaning that men and women of all ethnicities and ages tend to agree on what is considered attractive when evaluating an individual (Gallup & Frederick, 2010).

Past studies have also identified features associated with greater perceived vocal attractiveness; there are differences in what vocal qualities are considered attractive based on sex (Hodges-Simeon, Gaulin, & Puts, 2010). Women generally have higher voice pitch than men (Puts et al., 2012), and women with higher pitched voices are generally rated to be more attractive by men than women with lower pitched voices (Hughes et al., 2010; Röder, Fink, Feinberg, & Neave, 2015). In contrast, women tend to prefer men with lower pitched, deeper sounding voices because this reflects more preferred qualities in a mate such as strength and dominant personality traits (Riding, Longsdale, & Brown, 2006). Men with deeper pitched voices have been found to have greater testosterone levels, which is also associated with having higher sperm count and motility (Weusthoff, Baucom, & Hahlweg, 2013). Attractive voices and faces both provide information about potential mate quality.

Attractive individuals’ faces unconsciously transmit nonverbal signals such as hormone levels and reproductive fitness to those around them, which indicate that they are healthier than their unattractive counterparts (Gallup & Fredrick, 2010; O’Connor et al., 2012; Puts et al., 2012; Saxton et al., 2006; Sprecher, Sullivan, & Harfield, 1994; Zhao et al., 2015). More attractive individuals tend to have better disease resistance and live longer, although the mechanism behind this is currently still under study. Furthermore, faces can contain clues suggesting their fertility rates because men with more attractive faces have a higher semen count, sperm motility, and sperm morphology (Gallup & Fredrick, 2010). It is also possible to determine a person’s physical strength based on facial features alone (Gallup & Fredrick, 2010). All of this means that the people who are perceived as having an attractive face have underlying biological characteristics that are desirable in a mate. It can be inferred that attraction is meant to foster reproduction for individuals who have healthier genes. Men look for women with a higher reproductive value, which is often shown in women with feminine faces because this is an indicator of higher estrogen levels (Fraccaro et al., 2010; Sprecher et al., 1994). Women find men with a strong jawline to be more attractive because this is a sign of higher testosterone, which can also make a man more dominant and physically strong, increasing his ability to acquire resources (Gallup & Fredrick, 2010; Saxton et al., 2006; Sprecher et al., 1994). Ultimately reproduction rates are what is more important, in reference to sociobiology, people want to maximize the number of genes that they pass on to future generations (Oliver & Hyde, 1995). It has been demonstrated that people with more attractive faces generally have more children, and the children of more attractive individuals are more likely to be healthier and have a higher reproductive fitness, which in turn makes it more likely for those children to pass those traits on to their children (Gallup & Fredrick, 2010; Saxton et al., 2006). Vocal attractiveness has been found to be correlated with physical attractiveness, in that people with attractive voices were found to have more symmetrical features and were therefore considered more physically attractive (Gallup & Fredrick, 2010; Hughes, Farley, & Rhodes, 2010). The correlation between attractive faces and attractive voices in an individual suggests that those who have attractive voices would also have the same underlying health and fertility benefits as those with attractive faces. This also indicates that they have more desirable genetic traits, increasing their overall appeal as possible mates. This further suggests that those with attractive voices would be perceived as more desirable by the other sex.

Individuals with more attractive features—both faces and voices—are considered more likeable, kinder, warmer, more trustworthy, and more likely to achieve (Hughes et al., 2010; Zhao et al., 2015); this perception of attractive people having
desirable personality characteristics (i.e., the halo effect; Jones et al., 2001; Rhodes, 2006) may be a deciding factor in choosing to establish nonssexual relationships with others. About judging someone of the same sex, research has suggested that men and women both evaluate others’ attractiveness in terms of mate selection. Men use facial features related to masculinity to assess another man’s ability to compete for potential mates (Puts et al., 2012). Additionally, women often view other women with more attractive faces as more promiscuous and women with voices considered more feminine to be more likely to flirt with men, suggesting that more attractive individuals provide more competition. Researchers have found that same-sex friends have similar levels of physical attractiveness (Bleske-Rechek & Lighthall, 2010). Further research is needed to explore how attractiveness may relate to relationship building beyond romantic relationships, such as how one’s attractiveness influences others’ interest in building a friendship.

There is less research on how vocal attractiveness contributes to perceived attractiveness compared to facial attractiveness, although an individual’s facial and vocal attractiveness are often perceived as equally attractive in their contributions to mate selection (Doll et al., 2014). Voices considered more attractive often belong to people with greater bilateral symmetry, which is partly responsible for how attractive an individual is perceived (Hughes et al., 2010). Nonetheless, there are individuals whose vocal and physical attractiveness are imbalanced. Zuckerman and Sinicropi’s (2011) study found that individuals whose facial and vocal attractiveness differed were perceived less favorably than individuals with similarly attractive faces and voices. However, people with an attractive face could have an unattractive voice and still be considered attractive overall because the attractive feature compensates for the unattractive feature. Nevertheless, it could be that the mismatched face and voice attractiveness causes those around them to feel as if these traits do not belong to the same person (Zuckerman & Sinicropi, 2011). This person may or may not be considered attractive, depending on how vocal and facial attractiveness interact to affect overall perceived attractiveness.

The current study served as an extension of Zuckerman and Sinicropi’s (2011) study, which focused primarily on the disappointment participants felt toward an individual with mismatched faces and voices, where disappointment indicates that their expectations based on the face did not line up with the voice they heard. The current study investigated whether a mismatch in the attractiveness of a target’s face and voice would lead to lower perceived attractiveness, or if the more attractive trait would compensate for the unattractive one. Furthermore, it contributes to the literature on the effects of perceived vocal attractiveness because research on this aspect of attractiveness is limited in comparison to facial attractiveness.

It was hypothesized that an attractive face would compensate more for an unattractive voice than the reverse. This hypothesis extends from research that has highlighted how humans react to visual information faster than auditory information (Jain, Bansal, Kumar, & Singh, 2015); thus, when making judgments based on limited sensory input, faces would be more influential than voices. Previous research has found that women were more likely than men to marry someone whom they did not consider attractive (Sprecher et al., 1994). As such, it was also hypothesized that women would rate targets with a mismatch in facial and vocal attractiveness as more attractive than men, regardless of whether the unattractive quality is face or voice.

Method

Participants

A sample of 112 participants was recruited from the Department of Psychology’s Subject Pool. Participants were invited to participate via e-mail using simple random sampling. The list of students was made for all psychology class rosters using the subject pool (N = 466) with removal of duplicates when students were enrolled in more than one course using the pool (n = 52) and removal of students who were familiar with the study’s purpose and methodology and/or had participated in the pilot study (n = 14). The final list contained 400 students (275 = women, 125 = men). A total of 100 female students and 100 male students were randomly selected and sent e-mail invitations to participate. To be eligible to participate, individuals were required to self-identify as heterosexual and have no visual or auditory impairments. Participants received course credit as compensation for participation. Approval for this study was obtained from Stetson University’s institutional review board.

Participants ranged in age from 17 to 30 (M = 19.56, SD = 1.77). There were equal numbers of men (n = 56, 50%) and women (n = 56, 50%). Eighty-seven participants self-identified as White (77.7%), 8 as Black (7.1%), 13 as Hispanic (11.6%), and 4 as Asian (3.6%). Half the sample (n = 57, 50%)
Facial and Vocal Attraction from two individuals had to be thrown out because of 48 participants completed the pilot study. Data in their physics class for their participation. A total of 39, 29.5%, the remaining 79 (70.5%) reported that their relationship status did not affect their ability to rate attractiveness in others.

Research Design
This study used a 4 x 2 x 2 mixed-factorial design. The primary variable of interest was the target face/voice combinations; photographs and voice clips were matched as both attractive, both unattractive, or mismatched with face attractive/voice unattractive or face unattractive/voice attractive. Target sex and participant sex were also analyzed. Partial counterbalancing using a balanced Latin square was implemented to control for possible sequence effects; participants were randomly assigned to one of 14 condition sequences using block randomization. The dependent variables included the attractiveness ratings of each face/voice combination.

Materials
A pilot study was conducted to determine which photos and voices to use. Thirty-nine color photographs (20 men, 19 women) were acquired from Radboud Faces Database (Langner et al., 2010). The photographs all displayed White individuals wearing black shirts from the neck up with the person facing forward with a happy expression. Eighty voice clips (40 male, 40 female) were obtained via Amazon Mechanical Turk. Voice clips were 10 seconds in length and consisted of individuals reading this script: “The restaurant is on the corner of 1st and 8th street. To get there from here, you turn right onto 3rd street and take a left at the first stoplight. The restaurant is two blocks down the road on your right.” Voice clips were screened for clarity/sound quality and to ensure that the script was followed, resulting in 80 voice clips being selected from the 142 submitted.

Pilot study participants were recruited from the university’s Physics Department using convenience sampling via announcements in the classroom. To be eligible to participate, individuals were required to be age 18 or older and to have no visual or auditory impairments. They were offered extra credit in their physics class for their participation. A total of 48 participants completed the pilot study. Data from two individuals had to be thrown out because they circled the same number for all questions, resulting in a final pilot sample of 46 participants. They ranged in age from 18 to 33 (M = 20.8, SD = 2.2). Approximately half were men (n = 24, 52.2%) and 47.8% (n = 22) were women. Thirty-eight self-identified as White (82.6%), 2 as Black (4.4%), 3 as Hispanic (6.5%), and 3 as Asian (6.5%). Nearly all (n = 45, 97.8%) identified as heterosexual, with one bisexual (n = 1, 2.2%). Eighteen (39.1%) were single, 3 (6.5%) were casually dating, and 25 (54.3%) were in a serious relationship.

Participants completed the study privately in a classroom to minimize distractions. Participants first completed a demographics survey before rating the stimuli on a computer while wearing headphones; voice clips could be replayed if necessary. The order in which the faces and voices were experienced was partially counterbalanced using a 4 x 4 balanced Latin square to determine the order of reviewing male faces, female faces, male voices, and female voices. Within each grouping, photographs and voice clips were randomly ordered. After each face or voice was presented, participants rated the face or voice’s perceived attractiveness on a 9-point Likert-type scale from 1 (not at all attractive) to 9 (extremely attractive), with 5 representing average attractiveness. Completion of the study took approximately 20 minutes.

We obtained z scores for each picture and voice clip. All data points more than two standard deviations away from the mean were removed for only that face or voice in order to remove outliers to create a more accurate majority rating of subjects. Next, the mean scores of faces and voices were calculated separately by sex, and then z scores of the means were found to identify the extremes. Attractive faces and voices were chosen from those receiving a z score of 1.0 or higher, and unattractive faces and voices were chosen from those receiving a z score of -1.0 or below. The extremes of the pilot study were used to create the face/voice combinations; to create the mismatched attractiveness groups, the most attractive faces were paired with the least attractive voices and vice versa. See Table 1 for z-score pairings. After calculating the extremes in the pilot study, an imbalance of male and female targets was discovered. Four men were identified on both the attractive and unattractive extremes for facial and vocal features. However, there were only three women on both the unattractive and attractive side for both facial and vocal features; the gap between the extremes and the next possible choice for the extremes was too large to reasonably
consider using four women clusters. This is why the study design was altered to include only six female targets instead of eight. A decision then had to be made on whether to cut the male targets down to six also; it was decided to retain eight because the extremes were so close and to strengthen the male findings.

Measures
Participants completed two items from the Perceived Attractiveness Scale (Elliot et al., 2010), which assessed information about perceived attractiveness of each target. Items were rated on a 9-point Likert-type scale from 1 (not at all attractive) to 9 (extremely attractive) using two items from the Perceived Attractiveness Scale: “How pleasant is this person to look at?” to evaluate facial attractiveness, and “How attractive do you think this person is?” An additional item was developed to evaluate vocal attractiveness: “How pleasant is this person to listen to?” with another general attractiveness item added to evaluate consistency in responding. The Perceived Attractiveness Scale is a 3-item scale; one item on the scale that asked about interest in meeting the person was not included in this study. Responses were averaged together; higher scores indicated greater perceived attractiveness. The Perceived Attractiveness Scale had an average $\alpha$ of .83 (max $\alpha = .88$, min $\alpha = .76$) across targets. A fifth question was asked that differed depending on whether the participant was the same sex as the face/voice combination; participants were asked to rate how much they would like to be friends with the target if they were the same sex, and asked how much they would like to date the target if they were the opposite sex. Participants also completed demographic questions regarding their age, sex, race, and relationship status.

In addition, participants responded to a question assessing whether they felt their relationship status affected their ability to evaluate the attractiveness of other people. This was an exploratory analysis due to the fact that no existing research had investigated whether relationship status might relate to ratings of perceived attractiveness. Mixed-design multivariate analyses of variance performed separately by sex for male participants’ attractiveness ratings of female targets and female participants’ attractiveness ratings of male targets found no significant relationship between their relationship status or whether they perceived that their relationship would affect their ratings of attractiveness. As such, these variables were not included in the main study analyses, and all participants were retained for analyses.

Procedure
Participants completed the study individually in a classroom to minimize distractions. Participants first provided informed consent and then completed the demographics survey. They viewed the 14 targets one at a time on a computer for 30 seconds; the 10-second voice clip was replayed three times through headphones while the photograph was displayed. After each combination was presented, participants rated the target’s perceived attractiveness on a paper-and-pencil survey. Participation took approximately 10 minutes.

Data Analysis
Data was analyzed with IBM SPSS software version 24. Given that the faces and voices were chosen in the pilot to represent the extremes of attractive and unattractive, there was the possibility that regression to the mean would be a threat to validity. Dependent $t$ tests were performed comparing pilot and main study mean ratings for faces and voices. There were no significant differences for women’s faces, $t(5) = 0.07, p = .99$, Cohen’s $d = 0.01$; women’s voices, $t(5) = 0.61, p = .57$, Cohen’s $d = 0.24$; and men’s faces, $t(7) = 0.36, p = .73$, Cohen’s $d = 0.03$. There was a significant difference for men’s voices, $t(7) = -2.53, p = .04$, with slightly higher ratings in

<table>
<thead>
<tr>
<th>TABLE 1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Pilot Study Face and Voice z-Score Pairings for Targets</strong></td>
</tr>
<tr>
<td>Male Conditions</td>
</tr>
<tr>
<td>Face-Attractive, Voice-Attractive</td>
</tr>
<tr>
<td>Face-Attractive, Voice-Unattractive</td>
</tr>
<tr>
<td>Face-Unattractive, Voice-Attractive</td>
</tr>
<tr>
<td>Face-Unattractive, Voice-Unattractive</td>
</tr>
<tr>
<td>Face-Unattractive, Voice-Attractive</td>
</tr>
<tr>
<td>Face-Unattractive, Voice-Unattractive</td>
</tr>
<tr>
<td>Female Conditions</td>
</tr>
<tr>
<td>Face-Attractive, Voice-Attractive</td>
</tr>
<tr>
<td>Face-Attractive, Voice-Unattractive</td>
</tr>
<tr>
<td>Face-Unattractive, Voice-Attractive</td>
</tr>
<tr>
<td>Face-Unattractive, Voice-Unattractive</td>
</tr>
<tr>
<td>Face-Unattractive, Voice-Attractive</td>
</tr>
<tr>
<td>Face-Unattractive, Voice-Unattractive</td>
</tr>
</tbody>
</table>
the main study ($M = 4.75, SD = 1.39$) compared to the pilot study ($M = 4.46, SD = 1.61$). However, this represented a small effect, Cohen’s $d = 0.14$. This suggests that regression to the mean was likely not a threat in this study.

The hypotheses were tested using a mixed-design multivariate analysis of variance; target sex, attractiveness of the face, and attractiveness of the voice were entered as within-subjects factors, and participant sex was entered as a between-subjects factor. Four repeated measures analyses of variance using the same independent variables were performed for men rating male combinations for friendship, men rating female combinations as potential dating partners, women rating female combinations for friendship, and women rating male combinations as potential dating partners.

**Results**

Box’s Test of Equality of Covariance Matrices was insignificant, indicating no issues across groups, $F(36, 40714.7) = 1.01, p = .46$. Mauchly’s test indicated that the assumption of sphericity was not violated ($\varepsilon = 1.0$). There was a significant main effect of target facial attractiveness, $F(1, 110) = 497.92, p < .001, \eta^2_p = .82$; those with attractive faces

### TABLE 2

Overall Attractiveness Ratings for Target Conditions

<table>
<thead>
<tr>
<th></th>
<th>Men $M$ (SD)</th>
<th>Women $M$ (SD)</th>
<th>$M$ Difference</th>
<th>Total $M$ (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Male Targets</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Face Attractive/Voice Attractive</td>
<td>5.32 (1.37)</td>
<td>5.73 (0.97)</td>
<td>0.41</td>
<td>5.33 (1.20)</td>
</tr>
<tr>
<td>Face Attractive/Voice Unattractive</td>
<td>4.47 (1.47)</td>
<td>5.27 (1.00)</td>
<td>0.80</td>
<td>4.87 (1.32)</td>
</tr>
<tr>
<td>Face Unattractive/Voice Attractive</td>
<td>3.90 (1.33)</td>
<td>4.38 (1.03)</td>
<td>0.48</td>
<td>4.14 (1.21)</td>
</tr>
<tr>
<td>Face Unattractive/Voice Unattractive</td>
<td>3.01 (1.32)</td>
<td>3.26 (1.26)</td>
<td>0.25</td>
<td>3.14 (1.29)</td>
</tr>
<tr>
<td><strong>Female Targets</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Face Attractive/Voice Attractive</td>
<td>5.99 (1.31)</td>
<td>6.31 (1.10)</td>
<td>0.32</td>
<td>6.15 (1.22)</td>
</tr>
<tr>
<td>Face Attractive/Voice Unattractive</td>
<td>5.29 (1.10)</td>
<td>5.37 (1.01)</td>
<td>0.08</td>
<td>5.33 (1.05)</td>
</tr>
<tr>
<td>Face Unattractive/Voice Attractive</td>
<td>3.83 (1.10)</td>
<td>4.45 (1.18)</td>
<td>0.62</td>
<td>4.14 (1.18)</td>
</tr>
<tr>
<td>Face Unattractive/Voice Unattractive</td>
<td>3.76 (1.28)</td>
<td>4.27 (1.15)</td>
<td>0.51</td>
<td>4.02 (1.24)</td>
</tr>
</tbody>
</table>

### TABLE 3

Dating and Friendship Desirability Target Ratings

*(Dating Ratings for Other-Sex Targets, Friendship Ratings for Same-Sex Targets)*

<table>
<thead>
<tr>
<th></th>
<th>Male Participants</th>
<th>Female Participants</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Dating Ratings $M$ (SD)</td>
<td>Friendship Ratings $M$ (SD)</td>
<td>Dating Ratings $M$ (SD)</td>
<td>Friendship Ratings $M$ (SD)</td>
</tr>
<tr>
<td><strong>Face Main Effects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attractive</td>
<td>4.35 (1.90)$^*$</td>
<td>5.31 (1.51)$^*$</td>
<td>4.31 (1.61)$^*$</td>
<td>5.83 (1.61)$^*$</td>
</tr>
<tr>
<td>Unattractive</td>
<td>2.50 (1.40)</td>
<td>4.51 (1.65)</td>
<td>2.50 (1.35)</td>
<td>5.30 (1.75)</td>
</tr>
<tr>
<td><strong>Voice Main Effects</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attractive</td>
<td>3.48 (1.68)</td>
<td>5.17 (1.48)$^*$</td>
<td>3.65 (1.59)$^*$</td>
<td>5.68 (1.66)$^*$</td>
</tr>
<tr>
<td>Unattractive</td>
<td>3.37 (1.63)</td>
<td>4.65 (1.69)</td>
<td>3.17 (1.37)</td>
<td>5.46 (1.70)</td>
</tr>
<tr>
<td><strong>Face and Voice Interaction</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Face-Attractive/Voice-Attractive</td>
<td>4.52 (2.04)</td>
<td>5.61 (1.44)</td>
<td>4.50 (1.69)</td>
<td>6.16 (1.62)$^*$</td>
</tr>
<tr>
<td>Face-Attractive/Voice-Unattractive</td>
<td>4.19 (1.77)</td>
<td>5.02 (1.59)</td>
<td>4.13 (1.53)</td>
<td>5.51 (1.60)</td>
</tr>
<tr>
<td>Face-Unattractive/Voice-Attractive</td>
<td>2.45 (1.32)</td>
<td>4.73 (1.52)</td>
<td>2.79 (1.49)</td>
<td>5.20 (1.70)</td>
</tr>
<tr>
<td>Face-Unattractive/Voice-Unattractive</td>
<td>2.55 (1.49)</td>
<td>4.29 (1.78)</td>
<td>2.21 (1.21)</td>
<td>5.41 (1.80)</td>
</tr>
</tbody>
</table>

*Note: $^*$ indicates significant difference between groups*
Given the number of interaction effects, only significant interactions are reported. There was a significant interaction between target sex and target voice attractiveness, \( F(1, 110) = 14.70, p < .001, \eta^2_p = .12 \); attractive female voices (\( M = 5.14, SD = 1.20 \)) were rated more attractive than attractive male voices (\( M = 4.83, SD = 1.20 \)), and unattractive female voices (\( M = 4.67, SD = 1.15 \)) were rated significantly more attractive than unattractive male voices (\( M = 4.00, SD = 1.30 \)).

There was a significant interaction between participant sex, target sex, and target facial attractiveness, \( F(1, 110) = 11.35, p = .001, \eta^2_p = .09 \); female targets with attractive faces were rated similarly by men (\( M = 5.64, SD = 1.20 \)) and women (\( M = 5.84, SD = 1.06 \)). However, women rated male targets with attractive faces (\( M = 5.50, SD = 0.98 \)), male targets with unattractive faces (\( M = 3.82, SD = 1.15 \)), and female targets with unattractive faces (\( M = 4.36, SD = 1.16 \)) more highly than men rated male targets with attractive faces (\( M = 4.90, SD = 1.42 \)), male targets with unattractive faces (\( M = 3.46, SD = 1.32 \)), or female targets with unattractive faces (\( M = 3.79, SD = 1.19 \)). There was a significant interaction between target sex, target facial attractiveness, and target voice attractiveness, \( F(1, 110) = 44.37, p < .001, \eta^2_p = .29 \); male and female mismatched targets with unattractive faces/attractive voices were rated the same level of attractiveness by the overall sample, but for all other face/voice combinations, female targets were rated higher than male targets; means and standard deviations are displayed in Table 2 in Total column.

There was a significant interaction between participant sex, target sex, target facial attractiveness, and target voice attractiveness, \( F(1, 110) = 5.79, p = .018, \eta^2_p = .05 \); Target 2 displays mean target ratings by sex. Although women’s ratings of every target were higher than the men’s, the degree of difference varied. Ratings were most similar for female mismatched targets with attractive face/unattractive voice and least similar for male targets with attractive face/unattractive voice. Figure 1 displays the four-way interaction.

Table 3 displays means and standard deviations for male and female participants’ ratings of target groups’ dating desirability (for other-sex targets) and friendship desirability (for same-sex targets). Mauchly’s test indicated that the assumption of sphericity was not violated (\( \epsilon = 1.0 \)) for any of these
analyses. There was a significant main effect of target facial attractiveness, $F(1, 55) = 96.35, p < .001$, $\eta^2 = .64$, for dating desirability of female targets. Female targets with attractive faces were rated as more desirable dating partners than those with unattractive faces. There was no significant main effect of target voice attractiveness, $F(1, 55) = 0.85, p = .36, \eta^2 = .02$, or interaction between face and voice attractiveness for female target dating desirability, $F(1, 55) = 3.31, p = .07, \eta^2 = .06$.

There was a significant main effect of target facial attractiveness, $F(1, 55) = 150.84, p < .001$, $\eta^2 = .73$, and target voice attractiveness, $F(1, 55) = 16.33, p < .001, \eta^2 = .23$, for dating desirability of male targets. Male targets with attractive faces and voices were rated as more desirable dating partners than those with unattractive faces and voices. There was no significant interaction between face and voice attractiveness for male target dating desirability, $F(1, 55) = 1.45, p = .25, \eta^2 = .03$.

There was a significant main effect of target facial attractiveness, $F(1, 55) = 24.23, p < .001, \eta^2 = .31$, and target voice attractiveness, $F(1, 55) = 22.49, p < .001, \eta^2 = .29$, for friendship desirability of male targets. Male targets with attractive faces and voices were rated as more desirable friends than those with unattractive faces and voices. There was no significant interaction between face and voice attractiveness for male target friendship desirability, $F(1, 55) = 0.52, p = .47, \eta^2 = .01$.

There was a significant main effect of target facial attractiveness, $F(1, 55) = 15.84, p < .001, \eta^2 = .22$, and target voice attractiveness, $F(1, 55) = 5.08, p = .03, \eta^2 = .09$, for friendship desirability of female targets. Female targets with attractive faces and voices were rated as more desirable friends than those with unattractive faces and voices. There was also a significant interaction between face and voice attractiveness for female target friendship desirability, $F(1, 55) = 14.83, p < .001, \eta^2 = .21$; the female matched attractiveness target received the highest ratings overall, while the mismatched targets with unattractive face/attractive voice received the lowest friendship rating.

**Discussion**

This study expanded on the work of Zuckerman and Sinicropi (2011), which found targets whose facial and vocal attractiveness differ were perceived less favorably than individuals with similarly attractive faces and voices. The current study explored how targets with mismatched facial and vocal attractiveness compared to those who have similarly attractive or unattractive features. It was found that mismatched targets were still perceived as more attractive than those with two unattractive features, but were perceived as less attractive than those with two matched attractive features. The hypothesis that an attractive face would be more likely to compensate for an unattractive voice, as opposed to an attractive voice compensating for an unattractive face, was supported. Additionally, when evaluating targets’ dating potential, both men and women assigned higher ratings to targets with attractive faces than those with unattractive faces, for both those with attractive and unattractive voices. Participants spent 30 seconds making their decision about the attractiveness of each target; these ratings were made based on a very brief evaluation period, a quick “first impression.”

Men generally focus on the physical reproductive potential of a mate first (March & Grieve, 2014), meaning men may rate women more harshly based on just their first impressions of them. Future studies could determine whether exposure for a longer period of time results in differences in perceived attractiveness.

It is particularly interesting to note that the matched attractiveness targets received the highest ratings of attractiveness despite those faces and voices not having the highest attractiveness $z$ scores in the pilot study. The most attractive faces were paired with the least attractive voices (and vice versa) based on $z$ scores. This suggests that, although the more attractive trait may significantly compensate for the unattractive one, having an unattractive feature does significantly decrease one’s perceived attractiveness compared to someone who has two attractive features. It may be that the unattractive feature indicates poorer gene quality (Gallup & Frederick, 2010) or is associated with undesirable personality traits (Zhao et al., 2015; Zuckerman & Sinicropi, 2011). Another possible explanation may be that individuals could have perceived that some faces and voices did not belong together, which would be consistent with Zuckerman and Sinicropi (2011). Participants might have perceived that features came from different people and might therefore have been disappointed, or that incongruence might have resulted in lower attractiveness ratings. A future replication of this study could include open-ended questions asking participants to explain why they found one target more or less attractive compared to another to explore their conscious decision-making process and whether incongruence is an explanation for attractiveness.
ratings. Alternatively, a future study could alternate stimuli exposure—exploring whether attractiveness ratings are different if participants first hear targets’ voices, and then see their faces and vice versa—to determine whether this is associated with different ratings or greater incongruence.

The hypothesis that women would rate mismatched targets higher than men was supported. Female participants always rated the targets as more attractive than men, regardless of target sex. This is a phenomenon that has not been explored much in the literature. It may be associated with gender differences in that women are socialized to be nicer than men or that men are more critical in evaluating physical attractiveness. However, it could also be because women did actually find all targets—even those considered less attractive in the pilot study—to be more attractive. Women may find a greater variety of features to be appealing. In terms of mate selection, men place more importance on good looks than women (Eastwick & Finkel, 2008; March & Grieve, 2014). The focus on visual attractiveness is supported by the finding that, although both face and voice attractiveness were significant in female participants’ ratings of male targets’ dating desirability, facial attractiveness was the only significant effect for men’s ratings of female targets’ dating desirability. Other research has also found that women have a different preference in male attractiveness during different stages of their menstrual cycle (Bossio, Suschinsky, Puts, & Chivers, 2013). During peak fertility in a woman’s cycle, the woman tends to favor more masculine characteristics. This should be explored in further research; moreover, asking participants for explanations on why they consider an individual to be attractive or unattractive in order to help identify qualitative differences in how men and women judge attractiveness.

Lastly, this study sought to explore how perceived attractiveness affected same-sex ratings in terms of how individuals may be evaluated as potential friends. Both men and women rated the friendship desirability of same-sex targets significantly different based on their perceived attractiveness; targets with attractive faces and attractive voices were rated as more desirable friends. Men saw more friendship potential in the matched attractive targets and the least friendship potential in the matched unattractive targets, which is consistent with previous research findings that more attractive individuals are considered kinder, warmer, and more trustworthy (Zhao et al., 2015).

However, there was a significant interaction effect of face and voice attractiveness for female participants; it was the female face unattractive/voice attractive targets that received the lowest friendship ratings, rather than the female matched unattractive target. This may be a result of the degree of facial unattractiveness. In the pilot study, the female face assigned to the matched unattractive target was ranked as less unattractive than the faces assigned to the mismatched targets with unattractive faces, which would remain consistent with Zhao et al. (2015). Other research found that same-sex friends normally have the same level of physical attractiveness (Bleske-Rechek & Lighthall, 2010), so participants in the present study might have indicated a preference to be friends with someone who they considered to be of similar attractiveness as themselves. Participants’ own perceived attractiveness was not evaluated in this study; future research should take that into account, whether participants’ perceptions of their own facial and vocal attractiveness (and whether they see these qualities as matched in terms of their appeal) are related to their perceptions of others’ attractiveness.

Limitations
This study significantly contributed to the literature on perceived attractiveness, but it is not without limitations. Participants in the pilot study were obtained using convenience sampling from the physics department and thus were not representative of the university as a whole. Although participants in the main study were randomly sampled, they were obtained from the department of psychology’s subject pool. It is not as diverse a group as the general student population because all students in the subject pool are currently enrolled in at least one psychology course. This could affect generalizability of the findings.

An equal number of men and women were recruited, but the sample was largely White (78%). This could actually be a strength of the study because all the photographs presented in the study depicted White men and women, meaning that most participants were of the same race as the visual subjects. However, 22% of participants were not White; although previous research has found that men and women of all ethnicities tend to agree that the same type of facial features are considered attractive (Gallup & Frederick, 2010), it is a limitation in this study. Future research should recruit a more racially diverse sample and use racially diverse photographs.
Facial and Vocal Attraction

Participants were required to have identified as heterosexual in this study so that they would only consider the other sex as a potential mate and consider the same sex as a potential friend. This decision was made to enhance validity of the data, but it means that findings are only applicable to heterosexual women and men. Future research should determine whether the same pattern of findings in responses to matched and mismatched attractive targets is found with samples of bisexual, gay, and lesbian individuals.

Lastly, a methodological limitation is that there were only six female targets included in this study as opposed to the eight male targets. In conducting the pilot study, fewer female faces and voices were considered less attractive as denoted by a z score below minus one. It was decided to maintain two targets in each condition for the male analyses rather than reduce to match the reduced female targets to strengthen those findings. This does mean that two of the female target conditions are based on a single target and not the average of two targets, possibly making the findings less reliable. However, because the hypotheses were primarily interested in the mismatched conditions, having only one female target in each matched condition is less of a weakness than if the mismatched conditions only had one target. Future research should endeavor to have an equal number of male and female subjects, as well as more targets in each combination category. This is particularly important given that, in some cases, one target in the same category was rated significantly differently from the other individual in that category; the two male matched unattractive targets were rated the two most unattractive male targets, but one received a significantly higher rating than the other.

Conclusion

This study found significant differences in the attractiveness ratings of targets with matched features and mismatched features. Targets with an unattractive feature (face or voice) and an attractive feature were considered less attractive than targets with two attractive features, but more attractive than targets with two unattractive features. Targets with an attractive face and unattractive voice were rated as more attractive than subjects with an unattractive face and an attractive voice, indicating the visual stimulus may be more influential in making initial attractiveness assessments. Furthermore, women’s ratings of both male and female targets were higher than men’s, which was consistent for matched and mismatched subjects. These findings are meaningful in understanding how men and women evaluate the attractiveness of both potential mates as well as potential friends.

References


https://doi.org/10.1146/annurev.psych.57.102904.190208


http://dx.doi.org/10.1177/147470491301100119

http://dx.doi.org/10.1111/j.1439-0310.2006.01278.x

http://dx.doi.org/10.1037/0022-3514.66.6.1074


http://dx.doi.org/10.2224/sbp.2015.43.5.855

http://dx.doi.org/10.1007/s10919-011-0106-5

Author Note. Sheila Bonnough, Department of Psychology, Stetson University; Erin Moore, Department of Psychology, Stetson University.

Sheila Bonnough is now at Department of Physics, University of Central Florida.

The first author would like to thank Erin M. Foster for her assistance in setting up the room for the experiment and Megan E. Pirrotta for her assistance in proofreading the paper. Special thanks also to Psi Chi Journal reviewers for their support.

Correspondence concerning this article should be addressed to Sheila Bonnough Department of Physics, University of Central Florida, Orlando, FL, 32826. E-mail: sbonnough@knights.ucf.edu
Rape is prevalent among college women, with reported rates ranging from 15% to 30% in some samples (Botta & Pingree, 1997; Cleere & Lynn, 2013; Kahn, Jackson, Kully, Badger, & Halvorsen, 2003; Osman, 2016). Rape has also been linked to a multitude of negative outcomes including substance abuse, self-objectification, sexual dissatisfaction, and symptoms of post-traumatic stress disorder, depression, anxiety, and disordered eating (Carr & Szymanski, 2011; Harned, 2000; Layman, Gidycz, & Lynn, 1996; Orlando & Koss, 1983; Oshodi, Macharia, Lachman, & Seedat, 2017). Another construct that has been examined for its relation to rape is body shame. Researchers have asserted that, because rape is a body-intrusive experience, negative feelings produced by the rape can be manifested as body-related concerns including body shame. Following rape, a woman might associate her negative feelings with her body and, thus, focus on negative aspects of her body, placing her at risk for elevated body shame (Carcirieri & Osman, 2011; Carr & Szymanski, 2011; Harned, 2000; Oppenheimer, Howells, Palmer, & Challoner, 1985; Schechter, Schwartz, & Greenfeld, 1987).

Body shame can be defined as the shame felt by a woman when her body does not fit with internalized idealistic cultural standards (McKinley & Hyde, 1996), and has been associated with sexual victimization in childhood and adulthood including rape (Andrews, 1995, 1997; Carcirieri & Osman, 2011; Carr & Szymanski, 2011; Davidson & Gervais, 2015; Vidal & Petrak, 2007). For example, Vidal and Petrak (2007) examined body shame in their clinical and community sample of women who experienced sexual victimization. They reported that body shame scores in their sample were greater than those reported for a validation sample of university students, although sexual victimization was not measured in this university sample (Andrews, Qian, & Valentine’s, 2002). Furthermore, a positive relationship between body shame and sexual victimization has been found in samples of female college students (Carcirieri & Osman, 2011; Carr & Szymanski, 2011; Harned, 2000; Oppenheimer, Howells, Palmer, & Challoner, 1985; Schechter, Schwartz, & Greenfeld, 1987).

**ABSTRACT.** Rape has been linked to a number of negative outcomes for college women, including body shame. The current study examined body shame levels among female rape victims based on status of personal acknowledgment that they experienced rape (yes; no) and recency since rape (recent, within the past year; earlier, between age 14 and the past year). Undergraduates (n = 251 women who reported behaviors consistent with legal definitions of rape) completed the Sexual Experiences Survey and Body Shame subscale of the Objectified Body Consciousness Scale. A significant interaction between acknowledgment status and recency since rape emerged, p = .02, partial $\eta^2 = .022$. Results demonstrated that body shame was greater for women in the acknowledged-recent group than women in the acknowledged-earlier group, but women in the unacknowledged-recent and -earlier groups did not differ on body shame scores. Findings suggest that acknowledgment may be associated with lower body shame for women who have experienced rape less recently, and could have clinical and educational implications.
trigger event may occur at some point that helps women to work through their rape, thus resulting in improved outcomes for women raped longer ago compared to those raped more recently.

Rape acknowledgment specifically refers to a woman’s acknowledgment of her own rape experience such that an acknowledged victim would label her experience as rape (Koss, 1985). However, many college women who have been raped are unacknowledged, which means that they have experienced rape, but do not label their experience as rape. Rape acknowledgment is a rape-related event that may occur soon after a woman experiences rape. However, evidence has suggested that it often occurs at a later point in time (Botta & Pingree, 1997; Gleere & Lynn, 2013; Peterson & Muehlenhard, 2011) and that it may be a trigger that motivates a woman to deal with her experience and eventually leads her to improved outcomes or recovery (Kelley & Gidyecz, 2015; Littleton et al., 2006). Although body shame based on rape acknowledgment status has not been examined, rape acknowledgment has been associated with other outcomes, both positive and negative.

Some researchers have found positive impacts associated with acknowledgment including better adjustment and coping, less distress, more social support and greater empathy for other women who have been raped (Botta & Pingree, 1997; Clements & Ogle, 2009; Littleton et al., 2006; Osman, 2016). Conversely, other have found negative impacts associated with acknowledgement including greater feelings of stigma, confusion and sadness, and more post-traumatic stress disorder and somatic symptoms (Conoscenti & McNally, 2006; Kahn et al., 2003; Layman et al., 1996; Littleton et al., 2008). Furthermore, some researchers have found no differences based on acknowledgment status such as in psychological well-being or sexual satisfaction (Layman et al., 1996; McMullin & White, 2006; Orlando & Koss, 1983), and other studies have produced inconsistent results. For example, Layman et al. (1996) did not find a difference in self-blame between acknowledged and unacknowledged women, Bondurant (2001) found that acknowledged women engaged in more self-blame, whereas others have found that acknowledged women engaged in less self-blame (Botta & Pingree, 1997; Pitts & Schwartz, 1993). Botta and Pingree (1997) also found that acknowledged women reported consuming less alcohol since rape than those who were unacknowledged, but in one of the only studies to consider both time and acknowledgment, McMullin and White (2006) found no differences...
in alcohol consumption between acknowledged and unacknowledged women (raped since age 14 and prior to the start date of the study) when assessed longitudinally at two different times separated by a period of about 10 months. However, they did find evidence suggesting that rape acknowledgment may be associated with decreased alcohol use over time. Thus, both recency since rape and acknowledgment status may be important to consider for other rape-related outcomes including body shame, but most studies examining acknowledgment did not investigate how it may interact with recency since rape. Considering how these factors may interact could shed light on the mixed results in the rape acknowledgment literature.

Consistent with the theoretical idea that negative feelings resulting from rape can be manifested as body shame given that rape is a violation of one’s body, increased body shame has been linked to sexual victimization in college women (Davidson & Gervais, 2015). However, Carcicieri and Osman (2011) found this only for women who had been victimized within the past year. Thus, recency since rape might have an impact on body shame. However, the potential influence of recency may co-occur with rape acknowledgment, given that women often do not acknowledge rape in the more immediate aftermath of its occurrence, but rather at some later point in time (Botta & Pingree, 1997; Cleere & Lynn, 2013; Littleton, et al., 2006; Peterson & Muehlenhard, 2011). Also, although rape experience may be a risk factor for body shame, body shame has never been examined based on rape acknowledgment status. Therefore, the purpose of the current study was to build on the literature by examining the role of each of these factors (recency since rape and acknowledgment status), and how they might interact to influence body shame. More specifically, body shame among women who reported behaviors consistent with legal definitions of rape was examined based on acknowledgment status (yes; no) and recency since rape (within the past year; between age 14 and the past year). These time frames were based both on past research (Carcicieri & Osman, 2011; Oshodi et al., 2017) and the standard time frame items used on the Sexual Experiences Survey (Koss et al., 2007).

If a woman’s acknowledgment of her own rape can occur at any point in time following rape, and eventually leads to improved outcomes or recovery, as past researchers have proposed (Kelley & Gidycz, 2015; Littleton, et al., 2006), then women raped more recently who have acknowledged rape are more likely to be actively trying to understand and deal with their experience, which may heighten body shame compared to other victim groups. Women who have acknowledged-earlier rapes are more likely to have moved further through the recovery process and, therefore, may have lower body shame compared to other victim groups. Finally, unacknowledged women may not be able to fully move through the recovery process so that their body shame scores may remain relatively more stable, regardless of recency since rape. Thus, we predicted that there would be an interaction between rape acknowledgment status and recency since rape such that (a) women in the acknowledged-earlier group would have the highest body shame scores, and (b) women in the acknowledged-earlier group would have the lowest body shame scores.

Method

Participants
Participants were 255 women who reported behaviors consistent with legal definitions of rape from a larger sample of 929 undergraduate women. These women were enrolled in either an introductory or upper-level psychology course at a midsize public university on the U.S. east coast. They volunteered as an option for extra credit, and were recruited via course announcements. Women who were raped in this sample were identified based on the Sexual Experiences Survey-Short Form Victimization (SES-SFV; Koss et al., 2007), which is described below. Participants were between 18 and 24 years old ($M_{age} = 19, SD = 1.5$). Eighty-one percent of participants identified as European American, 9% African American, 3% Hispanic, 2% Asian, and 5% Biracial or Multiracial. Fifty-six percent of participants identified as first-year students, 16% sophomores, 9% juniors, and 19% seniors.

Measures
To measure body shame, participants completed the Body Shame subscale (BSS) of the Objectified Body-Consciousness Scale (OBCS; McKinley & Hyde, 1996), rating their level of agreement with eight items on a 7-point scale ranging from 1 (strongly disagree) to 7 (strongly agree). Sample items include “I feel like I must be a bad person when I don’t look as good as I could,” “When I’m not the size I think I should be, I feel ashamed,” and “I would be ashamed for people to know what I really weigh.” Two items on the BSS were
reverse scored, and final scores were calculated as an average, with higher scores indicating greater body shame (McKinley & Hyde, 1996). The BSS of the OBCS is a reliable and valid measure. The Cronbach alpha in the current study was .83, which is similar to Cronbach alphas found in other samples of undergraduate women (ranging from .75 to .86; Carciriieri & Osman, 2011; Davidson & Gervais, 2015; McKinley & Hyde, 1996; Noser & Zeigler-Hill, 2014). With regard to validity, the BSS has been associated in the expected directions with body esteem, surveillance, self-objectification, sexual violence, and disordered eating (Davidson & Gervais, 2015; McKinley & Hyde, 1996; Moradi, Dirks, & Matteson, 2005).

The SES-SFV (Koss et al., 2007) was also completed to identify women who have been raped. On this measure, rape includes oral, anal, or vaginal penetration without consent due to intoxication, threatening harm, or using force. This survey includes nine questions that measure rape experience, but do not use the term rape. For example, “A man put his penis into my vagina, or someone inserted fingers or objects without my consent by threatening to physically harm me or someone close to me.” If the participant indicated rape experience on any of these items, they were classified as a rape victim and included in the analyses. Each of these nine items was followed by two questions asking when an experience (i.e., rape) took place, “How many times in the past 12 months” and “How many times since age 14,” with instructions specifying that the first question “refers to the past year going back from today,” and the second question “refers to your life starting on your 14th birthday and stopping 1 year ago from today.” These are the standard time frames for the SES-SFV. If rape occurred within the past year, recency since rape was classified as “recent.” If rape occurred since age 14 and prior to the past year, recency since rape was classified as “earlier.” If multiple rapes were reported, categorization was based on the most recent rape. Acknowledgment of rape was then measured by asking, “Have you ever been raped?” If the individual responded “yes,” the person was classified as acknowledged. If the individual responded “no,” the person was classified as unacknowledged. The SES-SFV is a valid measure for classifying rape and distinguishing between unacknowledged and acknowledged rape. For example, based on this measure, acknowledged rape has been associated with greater empathy for a rape victim than unacknowledged rape, and indicating victimization experience on the SES-SFV has been associated with expected outcomes such as intimate partner violence, body surveillance, body shame, somatization, depression, anxiety, and post-traumatic stress disorder intrusive experiences in previous studies (Davidson & Gervais, 2015; Davis, Gilmore, Stappenbeck, Balsan, & Norris, 2014; Osman, 2016).

Procedure
We obtained approval to conduct this study from an Institutional Review Board. Before administering the surveys, participants signed an informed consent form. We supplied participants with surveys that included demographic questions, followed by the BSS of the OBCS, and then the SES-SFV. Participants responded to these surveys in a large classroom setting (but not during a class) where they were instructed to sit every other seat apart to ensure privacy. Participant responses were anonymous. When participants finished, they placed their surveys in an anonymous drop bag and were provided references for counseling options as they left the room.

Results
Prevalence
As described above, based on responses to the SES-SFV, there were 255 women who experienced rape. We dropped four people from the analyses because they were missing data when asked about rape occurrence within the past year. Analyses, therefore, included 251 women who experienced rape. Participants who labeled their experience as rape were included in the acknowledged group (n = 46 or 18% of the sample). Of those, 65% reported acknowledged-recent experience (n = 30), and 35% reported acknowledged-earlier experience (n = 16). Participants who did not label their experience as rape were included in the unacknowledged group (n = 205). Of those, 60% reported unacknowledged-recent experience (n = 123), and 40% reported unacknowledged-earlier experience (n = 82). Given that those who are acknowledged are likely to be older and further along in their undergraduate careers than those who are unacknowledged (Botta & Pingree, 1997; Kahn et al., 2003), and that recency since rape may also influence these factors, we tested for group differences on our demographic variables (i.e., age, college class, and race). Women raped earlier were older F(1, 249) = 13.69 p < .0001, partial n² = .053, and further along in school, F(1, 249) =
9.85, \( p = .002 \), \( \text{partial } \eta^2 = .038 \), than women raped more recently. See Table 1 for demographic means and percentages based on group.

**Analysis of Covariance (ANCOVA)**

To test the hypotheses, we conducted a 2 x 2 (Acknowledgment Status x Recency Since Rape) ANCOVA on the body shame scores. Although there were unequal sample sizes, Levene’s test demonstrated that the assumption of homogeneity of variance for the ANCOVA was not violated, \( F(3, 247) = 1.67, \ p = .174 \). Age, \( p = .534 \), \( \text{partial } \eta^2 = .002 \), and college class, \( p = .393 \), \( \text{partial } \eta^2 = .003 \), were entered as covariates to control for their effects and were not significant. There was no significant main effect for acknowledgment status, \( F(1, 249) = 0.56, \ p = .455 \), \( \text{partial } \eta^2 = .002 \), but a significant main effect for recency since rape was revealed, \( F(1, 249) = 15.36, \ p < .0001 \), \( \text{partial } \eta^2 = .059 \). There was also a significant interaction between recency since rape and acknowledgment, \( F(3, 247) = 5.59, \ p = .019 \), \( \text{partial } \eta^2 = .022 \).

To examine the interaction, planned pairwise comparisons were conducted with a \( p \leq .01 \) significance level requirement using Dunn’s procedure to control for total experimentwise error. Although the means were in the expected directions, women in the acknowledged-earlier group did not differ significantly from those in the unacknowledged-earlier group, \( p = .088 \), \( \text{partial } \eta^2 = .031 \), and women in the acknowledged-recent group did not differ significantly from those in the unacknowledged-recent group, \( p = .175 \), \( \text{partial } \eta^2 = .012 \). As expected, women in the unacknowledged-recent and -earlier groups did not significantly differ, \( p = .094 \), \( \text{partial } \eta^2 = .014 \), but women in the acknowledged-recent and -earlier groups did, \( p = .001 \), \( \text{partial } \eta^2 = .223 \) (see Figure 1).

### Discussion

The current study examined the influence of rape acknowledgment status and recency since rape occurrence on body shame levels in a sample of college women who have been raped. It was predicted that acknowledged women raped within the past year would score higher on body shame than those in other groups, and that acknowledged women raped earlier would score lower on body shame than those in other groups. Although means were in the expected directions, these hypotheses were not statistically significant. Results revealed that women in the acknowledged-earlier group reported lower body shame than those in the acknowledged-recent group, but no other significant differences between groups were found.

Findings from the present study were consistent with results from McMullin and White (2006), who found no differences between acknowledged and unacknowledged rape victims on various outcomes measured at two different points in time about 10 months apart. However, their evidence regarding alcohol consumption following rape suggested that the benefits of acknowledgment may be seen over

### Table 1

Means and Percentages for Sample Demographics by Group

<table>
<thead>
<tr>
<th></th>
<th>Acknowledged</th>
<th>Unacknowledged</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Recent</td>
<td>Earlier</td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>30</td>
<td>16</td>
<td>46</td>
</tr>
<tr>
<td>M Age (SD)</td>
<td>18.90 (1.35)</td>
<td>19.88 (1.71)</td>
<td>19.40 (1.47)</td>
</tr>
<tr>
<td>Race</td>
<td>European American: 73% (81%)</td>
<td>84% (81%)</td>
<td>81% (81%)</td>
</tr>
<tr>
<td></td>
<td>African American: 10% (6%)</td>
<td>8% (10%)</td>
<td>9% (9%)</td>
</tr>
<tr>
<td></td>
<td>Hispanic: 7% (0%)</td>
<td>3% (2%)</td>
<td>3% (3%)</td>
</tr>
<tr>
<td></td>
<td>Asian: 3% (0%)</td>
<td>2% (2%)</td>
<td>2% (2%)</td>
</tr>
<tr>
<td></td>
<td>Biracial/Multiracial: 7% (13%)</td>
<td>3% (5%)</td>
<td>5% (5%)</td>
</tr>
<tr>
<td>Class</td>
<td>First year: 67% (44%)</td>
<td>63% (43%)</td>
<td>56% (56%)</td>
</tr>
<tr>
<td></td>
<td>Sophomore: 10% (13%)</td>
<td>18% (16%)</td>
<td>16% (16%)</td>
</tr>
<tr>
<td></td>
<td>Junior: 13% (19%)</td>
<td>7% (8%)</td>
<td>9% (9%)</td>
</tr>
<tr>
<td></td>
<td>Senior: 10% (25%)</td>
<td>11% (33%)</td>
<td>19% (19%)</td>
</tr>
</tbody>
</table>

Note. Earlier groups were older, \( p < .0001 \), and further along in school, \( p = .002 \), than recent groups.

### Figure 1

Ms, SDs, and cell sizes for body shame scores in Acknowledgment Status x Recency Since Rape Interaction. \( p < .001 \).
time. Although the present study was not longitudinal, results suggest that outcomes for acknowledged women measured at a later time period since rape may be better than for acknowledged women measured within a year of being raped. However, because acknowledged and unacknowledged participants did not differ on body shame, it is not clear if those who acknowledged their rape fare better or worse than those who did not acknowledge their rape when measured at a later or more recent time period.

It is also noteworthy that, although the difference between the two unacknowledged groups did not reach significance, there was a trend toward higher body shame scores for the recent group. The direction of the means also suggests a possibility that acknowledged-recent rape may be associated with heightened body shame compared to unacknowledged-recent rape, whereas acknowledged-earlier rape may be associated with lower body shame than unacknowledged-earlier rape. However, further research is needed to examine these potential relationships. Future study of both acknowledgment and recency since rape may also continue to shed light on the contradictory findings in the acknowledgment literature.

Results from the current study suggest that perhaps past researchers who reported positive outcomes (Botta & Pingree, 1997; Clements & Ogle, 2009; Littleton et al., 2006) had more acknowledged participants who were raped earlier, and those reporting negative outcomes (Conoscenti & McNally, 2006; Kahn et al., 2003; Layman et al., 1996) had more acknowledged participants who were raped more recently. This possibility highlights the importance of examining acknowledgment and recency since rape together.

Future researchers may also consider identifying more specific and different time frames since rape occurred because the current study was limited by the standard time frames utilized by the SESSFV. For example, women in the recent group who were raped within a month may differ from those raped 11 months before (Frazier et al., 2001), and women in the earlier group who were raped 2 years before may differ from those raped 5 years before. It may also be important to note that the current study examined recency since rape occurred, not recency since acknowledgment occurred, which could also be considered in future studies.

In the current study, participants completed the body shame measure prior to the sexual victimization measure. Thus, potential order effects should be considered. Also, although women in the acknowledged-earlier group reported lower body shame than those in the acknowledged-recent group, cause and effect cannot be assumed because participants could not be randomly assigned into groups. Therefore, other unmeasured characteristic factors may vary between the groups and potentially influence body shame. For example, research has found that women raped by a stranger are more likely to personally acknowledge their experience as rape than are women raped by an acquaintance (Koss, 1985; Koss, Dinero, Seibel, & Cox, 1988). Another limitation of the current study was the lack of control for experiential factors that might have impacted body shame during the time since rape occurred. For example, the acknowledged participants might have been more likely to utilize therapy, engage in coping strategies, or seek other types of support that may explain their lower body shame in the earlier compared to the recent group, and the relatively similar body shame levels reported by the two unacknowledged groups (Kelley & Gidycz, 2015; Littleton et al., 2006). These possibilities, as well as other rape factors (e.g., number of rape experiences, degree of physical injury, age at the time of rape, rape knowledge, police involvement) can be examined in future research for their influence on body shame and other outcomes. Future studies including various populations and obtaining higher rates of acknowledged participants may also help expand understanding of these associations, given that the current study examined female college students who were largely European American, and the sample size of acknowledged women was relatively small (18%), which warrants caution regarding generalizability and conclusions. Finally, although the effects observed in the current study were not large, which is reasonable given that multiple factors likely impact body shame, results may reflect subtle impacts on body shame that could have educational and clinical significance at the individual level.

This study contributes to the literature (Carcirieri & Osman, 2011; Davidson & Gervais, 2015) as among the first to examine body shame based on rape acknowledgment status and recency since rape. Although the present results may not be conclusive, they support the idea that time alone may not predict rape-related outcomes (Burt & Katz, 1987), and that acknowledgment status may also play a role in predicting body shame. Specifically, if a woman acknowledges that she has been raped, and she experienced rape over a year ago,
she is likely to report lower body shame than an acknowledged woman raped more recently. On the other hand, an unacknowledged woman may not fare better based on recency since rape. Thus, although it is not entirely clear that acknowledged women are better off compared to unacknowledged women for either time frame, results do suggest that acknowledged women may be better off regarding body shame if they were raped longer ago as compared to within a year. Furthermore, it is worth noting that women in the acknowledged-earlier group reported a mean body shame score (M = 3.03, SD = 1.32) that appears comparable to the score reported by Carcirieri and Osman’s (2011) nonvictimized group of college women (M = 3.4, SD = 1.2), as well as to scores reported in other female college samples (e.g., M = 3.25, SD = 1.04; M = 3.36, SD = 1.12) (McKinley & Hyde, 1996; Moradi et al., 2005), suggesting that acknowledged women raped over a year ago might not fare worse than nonvictims on body shame.

The current findings can be useful information for therapists working with rape victims who are experiencing body shame. Knowing an individual’s acknowledgment status and how much time has passed since rape, in conjunction with other factors that may be important or related to body shame (e.g., disordered eating, depression, substance abuse, sexual dysfunction; Carr & Szymbanski, 2011; McKinley & Hyde, 1996; Moradi et al., 2005; Sanchez & Kiefer, 2007), may help therapists optimize treatment plans for their patients. However, therapists should be aware that a patient’s acknowledgment status will not always be clear. Whereas women who seek therapy because they were raped are likely to be acknowledged, women undergoing therapy for some other reason could be an acknowledged or unacknowledged victim, if they have ever been raped. Results may also suggest considerations for rape education efforts on college campuses. The relatively low rates of acknowledgment found in the current study highlight the prevalence of unacknowledged rape in this population and may indicate a need for improved rape education. Also, it may be important to understand that many individuals targeted by these efforts may be unacknowledged rape victims who could become acknowledged as a result of exposure to the education (Peterson & Muehlenhard, 2011). Given this, the finding that acknowledged-earlier women may have lower body shame than acknowledged-recent women may be relevant and important for educators to share, while at the same time being sensitive to the possibility that both acknowledged and unacknowledged victims could be experiencing other negative rape-related outcomes to varying degrees.

**References**


Author Note. Carolyne Paige Merwin, Psychology Department, Salisbury University; Suzanne L. Osman, Psychology Department, Salisbury University.

Special thanks to Psi Chi Journal reviewers for their support.

Correspondence concerning this article should be addressed to Carolyne Paige Merwin, Psychology Department, Salisbury University, 812 Goshen Road #C29, West Chester, PA 19380. E-mail: cmerwin@mail.immaculata.edu
3 REASONS TO JOIN PSYCAS

PSYCAS BENEFITS STUDENTS AND YOUR PROGRAMS.
Most psychology students apply to more than one program. For students, the Centralized Application Service (CAS™) for psychology programs streamlines the application process, taking away the need to repeat application steps for each program of interest. A CAS can also streamline admissions processes for your psychology program, reducing the manual administrative work for admissions teams and faculty and facilitating the sharing of best practices in recruitment and retention among programs.

PSYCAS PAINTS THE BIG PICTURE.
PSYCAS provides data and analytics for all participating programs, collecting accurate statistics about the entire profession, such as a complete count of the number of applicants to psychology programs, their background and who from the applicant pool is accepted and enrolled into programs.

APPLICANTS WANT A MORE EFFICIENT APPLICATION PROCESS.
Today’s applicants are familiar with CAS technology, and having experienced when applying to undergraduate programs, expect it for their graduate applications too. Across other professions and disciplines, 90% of applicants express a preference for using a CAS when applying to graduate school, and prospective psychology graduate students say that having access to an efficient, streamlined, and less-stressful applicant process is worth the PSYCAS service fee.

SIMPLIFIED. STREAMLINED. SMART.
LEARN MORE ABOUT PSYCAS AT PSYCAS.APA.ORG
PREPARE TO CHANGE LIVES.

At the College of Clinical Psychology at Argosy University, we know that discovering your life’s purpose can make all the difference. Our supportive faculty is committed to helping you realize that purpose and succeed as a professional in one of the many areas of psychology research or clinical practice. Our Clinical Psychology programs offer a rigorous curriculum grounded in theory and research, while also fostering deep learning by incorporating experiential elements, classroom exercises, personal reflection and supervised field experiences.

And we’re proud to say that the Doctor of Psychology in Clinical Psychology (PsyD) program at each of our ten schools has received accreditation* from the American Psychological Association (APA), certifying that they meet their rigorous standards.

Each of our ten locations offers a degree in:

DOCTOR OF PSYCHOLOGY (PSYD) IN CLINICAL PSYCHOLOGY

We are now accepting applications for Spring and Fall 2018.

Learn more at clinical.argosy.edu/psichi

*The Doctor of Psychology in Clinical Psychology Program at Argosy University, Atlanta, Chicago, Hawaii, Orange County, Phoenix, San Francisco Bay Area, Schaumburg, Tampa, Twin Cities and Northern Virginia is accredited by the Commission on Accreditation of the American Psychological Association (APA). Questions related to the program’s accredited status should be directed to the Commission on Accreditation: Office of Program Consultation and Accreditation, American Psychological Association, 750 First Street, NE, Washington DC 20002. Phone: (202) 336-5879 / Email: apacredit@apa.org / Web: www.apa.org/ed/accreditation

Argosy University is accredited by the WASC Senior College and University Commission (888 Atlantic Ave., Suite 100, Alameda, CA 94501, wascac.org). Programs, credential levels, technology, and scheduling options are subject to change. Not all online programs are available to residents of all U.S. states. Administrative office, Argosy University, 601 South Lewis Street, Orange, CA 92866 ©2017 Argosy University. All rights reserved. Our email address is makearenew@argosy.edu

See auprograms.info for program duration, tuition, fees and other costs, median debt, salary data, alumni success, and other important information.
Publish Your Research in *Psi Chi Journal*

Undergraduate, graduate, and faculty submissions are welcome year round. Only the first author is required to be a Psi Chi member. All submissions are free. Reasons to submit include

- a unique, doctoral-level, peer-review process
- indexing in PsycINFO, EBSCO, and Crossref databases
- free access of all articles at psichi.org
- our efficient online submissions portal

View Submission Guidelines and submit your research at [www.psichi.org/?page=JN_Submissions](http://www.psichi.org/?page=JN_Submissions)

**Become a Journal Reviewer**

Doctoral-level faculty in psychology and related fields who are passionate about educating others on conducting and reporting quality empirical research are invited become reviewers for *Psi Chi Journal*. Our editorial team is uniquely dedicated to mentorship and promoting professional development of our authors—Please join us!

To become a reviewer, visit [www.psichi.org/page/JN_BecomeAReviewer](http://www.psichi.org/page/JN_BecomeAReviewer)

**Resources for Student Research**

Looking for solid examples of student manuscripts and educational editorials about conducting psychological research? Download as many free articles to share in your classrooms as you would like.

Search past issues, or articles by subject area or author at [www.psichi.org/?journal_past](http://www.psichi.org/?journal_past)

**Add Our Journal to Your Library**

Ask your librarian to store *Psi Chi Journal* issues in a database at your local institution. Librarians may also e-mail to request notifications when new issues are released.

Contact [PsiChiJournal@psichi.org](mailto:PsiChiJournal@psichi.org) for more information.

---

[Register an account:](http://pcj.msubmit.net/cgi-bin/main.plex)