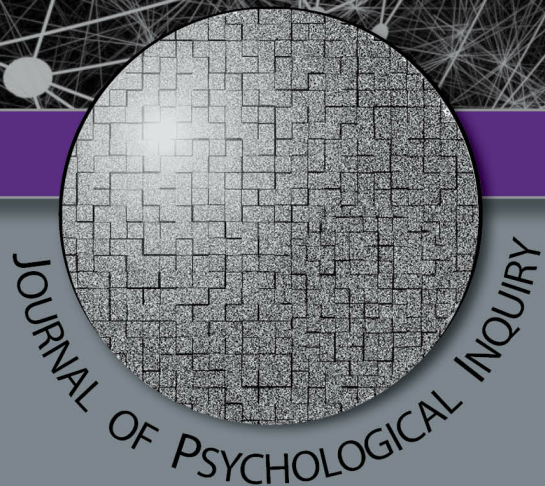




2025

Volume 29, Number 1, Spring

Issn 1085-6641



Journal of Psychological Inquiry

Editorial Board

Founding Editor

Mark E. Ware
Creighton University

Managing Editor

Ken Sobel
University of Central Arkansas

Graduate Assistant

Katelyn Pack
University of Central Arkansas

Associate Editors

Preston Bost
Wabash College

Drew A. Curtis
Angelo State University

Thomas J. Faulkenberry
Tarleton State University

Wind Goodfriend
Buena Vista University

Raymond Green
Texas A&M Commerce

J.T. Lacy
Fort Hays State University

Robyn Long
Baker University

Yeseul Nam
University of Central Arkansas

Donna Nelson
Winthrop University

Anna Park
University of Central Arkansas

Jennifer Peszka
Hendrix College

Mollie Price-Blackshear
University of Central Arkansas

Darren Ritzer
Winthrop University

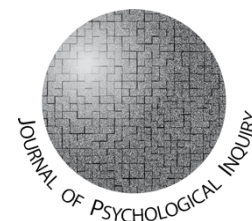
Merry Sleight
Winthrop University

Tammy Sonnentag
Xavier University

Amanda R. Stevens
Tarleton State University

Grace Sullivan
Colorado Mesa University

James Wages
University of Central Arkansas



Journal of Psychological Inquiry

Volume 29

Number 1

Spring, 2025

Journal Contents

From the Editor's Desk.....4

Empirical articles

Winner of this issue's

Elizabeth A. Dahl, Ph.D., Best Paper Award

Do people change the way they speak when they talk to you? A study of Chinese and Spanish-speakers' perceptions of and attitudes towards foreigner talk5

Ashna Mehra

University of Michigan

The father factor: Exploring the relationship between paternal involvement and young adult mental health22

Emily Gull, Amanda ElBassiouny, and Lauren Banks

California Lutheran University

Social support, age, and emotion regulation predict symptom severity of adults with major depressive disorder.....28

Alyssa M. Nelson¹, Joseph M. Diehl¹, Kevin S. LaBar¹, and Moria J. Smoski^{1, 2}

¹Duke University, ²Duke University Medical Center

"I support them, but...": Perceptions of transgender athletes depend on athlete gender and skill level35

Madden G. Ritzer, Merry J. Sleight, and Donna W. Nelson

Winthrop University

Empathy and perceptions of female speakers varying in confidence46

Meridee A. Ritzer and Donna W. Nelson

Winthrop University

Ethnic differences in seeking help after exposure to domestic violence55

Andrea Santibanez and Carmen Farrell

University of South Carolina Beaufort

Validity evidence for a computer-adaptive assessment of reading achievement at the secondary level 61

Catherine Windham and Giancarlo A. Anselmo

Winthrop University

Literature reviews

Psychedelic and dissociative-assisted therapy, current research and future potential..... 70

Alexander M. Singelis

California State University Chico

Call for Papers..... 82

FROM THE EDITOR'S DESK

Research serves a variety of purposes, including operating as a point of connection between disciplines and expanding the limits of existing knowledge. Further, research reflects the innately inquisitive nature of humankind. Such a nature makes research invaluable and profound, particularly within the field of psychology.

As both a graduate student and assistant, I have yet to read a research article that has not contributed to my understanding of myself, others, and the world itself. The research within this edition is no exception—the selections are impactful, insightful, and significantly benefit the field of psychology.

I cannot thank those who provided contributions for this issue enough, as each paper was notably crafted with care and thoughtfulness. I hope that our readers will find the selected papers as intellectually stimulating and inspiring as I did.

As this is my first time serving as a copyeditor, I hope to have met the standards of those who have filled this role before me. I have gratitude not only for the opportunity to serve as a copyeditor, but also for the chance to work on such incredible submissions.

Katelyn Pack

Katelyn Pack
Graduate Assistant
University of Central Arkansas

DO PEOPLE CHANGE THE WAY THEY SPEAK WHEN THEY TALK TO YOU? CHINESE AND SPANISH-SPEAKERS' PERCEPTIONS OF AND ATTITUDES TOWARDS FOREIGNER TALK

ASHNA MEHRA

UNIVERSITY OF MICHIGAN

Abstract – The study analyzes the perceptions and attitudes towards foreigner talk (FT), a speech register used when communicating with perceived foreigners. The study, unique in its focus on non-native English speakers, specifically those with Spanish or Chinese as their first language, recruited 64 participants through snowball sampling. The participants, native speakers of either Spanish or Chinese, all learned English after the age of 7. Through an online survey covering language background, attitudes toward native English speakers, experiences with and opinions of FT, and personal language ideologies, the study revealed significant differences between Chinese and Spanish speakers in their perceptions of FT. Chinese speakers displayed more favorable attitudes towards FT, feeling closer to a native speaker using FT compared to Spanish speakers. Conversely, Chinese participants experienced more social distance when interacting with someone in their own language compared to Spanish participants. These findings are true despite there being no significant difference in the amount of FT that Chinese-speaking and Spanish-speaking participants reported receiving. Correlations revealed that longer residence in the United States and lower self-reported accent correlated with a decreased amount of FT received. Participants with lower English proficiency found FT more helpful in both language groups. Additionally, individuals with stronger self-reported accents demonstrated lower prejudice towards others with accents. These findings contribute to understanding the impact of FT on non-native speakers and emphasize the importance of raising awareness among native speakers employing FT

Keywords: foreigner talk, non-native speakers, language perceptions, accents, belonging

Foreigner talk (FT) refers to a simplified version of language employed by native speakers when addressing non-native speakers. Many studies have explored native English speakers' attitudes towards foreign speakers in the United States, as well as their corresponding beliefs about them. Past authors inquired whether individuals make an effort to comprehend foreigners or merely dismiss them (Derwing et al., 2010). Furthermore, papers assessed whether individuals alter their speech patterns and personas when interacting with non-native speakers (Ferguson, 1975). These investigations were critical in determining the characteristics of and reasoning behind FT.

The current study approaches the subject from a different perspective by focusing on the foreign speaker's point of view, analyzing non-native speakers' perceptions and beliefs surrounding FT. The paper examines reception of FT, specifically asking what demographic groups are more likely to encounter it, and whether attitudes toward FT differ across these groups.

Additionally, the study aims to determine correlations between FT and feelings of belongingness in American culture, language ideologies, and social distance between foreigners and non-foreigners.

Features of, and Evidence for, FT

In 1971, Charles Ferguson coined the term foreigner talk as speech "used by speakers of a language to outsiders who are felt to have very limited command of the language or no knowledge of it at all." Ferguson suggested that many languages have particular features of pronunciation, grammar, and lexicon which are used when talking with foreigners. For example, a speaker of Spanish who wants to communicate with a foreigner will use the infinitive form of the verb rather than the correct conjugation (Ferguson, 1971). A few years later in 1975, Ferguson provided native students with ten sentences and asked them to write "equivalent phrases" directed towards a foreigner. All but two of the 36 students provided "foreigner talk sentences," suggesting

substantial agreement in basic features of FT and confirming that it is a widespread phenomenon (Ferguson, 1975). However, this work did not examine how people actually talk in real-life conversation.

To test the direct use of FT, Lugin and colleagues (2018) assessed whether people tend to adapt to their conversational partners in a mixed-cultural setting with a native and non-native speaker. The results revealed that participants (all native speakers of English) behaved significantly differently towards an English speaker with a foreign accent, reducing the number of words, speed of speech, time of interactions, and non-verbal cues. Lugin called this phenomenon AFC: adapted, foreigner-directed communication (Lugin et al., 2018).

Similarly, the sociolinguistic framework CAT (Communication Accommodation Theory) states that native speakers modify their speech to meet the communicative needs of non-native speakers (Beebe & Giles, 1984). Foreigner-directed speech (FDS) is a specific form of CAT characterized by general vowel hyperarticulation and a decrease in speech rates, intended to assist second-language learners (Uther et al., 2007). Another study used imaginal situations to find that FDS is also associated with exaggerated pitch range (Papoušek & Hwang, 1991).

Evelyn Hatch (1983) proposed three concrete characteristics of FT: regression, matching, and negotiation. With regression, native speakers regress back through stages of their own language development, until they find an appropriate level to match the listener. With matching, native speakers attempt to understand the state of the speaker and imitate their observations. Finally, with negotiation, native speakers simplify and clarify their speech according to feedback they obtain from the listeners they are communicating with.

Functions of, and Views on, FT

A double message exists regarding FT: FT may be perceived as helpful to non-native speakers, and/or it may be perceived as patronizing (Bobb et al., 2019). These contrasting opinions suggest that FT still has an inconclusive role in aiding second-language acquisition (Dela Rosa & Arguelles, 2016). The perception of FT may vary among individuals on the giving and receiving ends, and even across different cultural contexts.

For the argument of FT as a helpful phenomenon, many believe that native speakers adjust their speech to optimize comprehension and lighten the burden of non-native speakers. There are three ways that native speakers do this: omission, expansion, and replacement (Ferguson, 1975). In this regard, non-native speakers mean well to modify their speech to meet the

needs of foreigners and assist them (Bobb et al., 2019). For example, in a questionnaire sent to native English speakers, four fifths of respondents claimed that FT fostered communication, suggested respect and consideration for non-native speakers, and avoided conflict (Drljača Margić, 2017). Ensuring a smooth exchange of information, FT can be viewed as facilitative and positive.

On the other hand, using FT may imply the offensive idea that native speakers are superior to non-native speakers. Foreigners often receive more FT than they believe is necessary, suggesting their linguistic inadequacy (Bobb et al., 2019). This form of speech is seen as condescending, controlling, and dismissive, making foreigners feel incompetent (Zuengler, 1991). From the native speaker's point of view, Ferguson (1975) found that native English students had an overall negative attitude and disapproval towards the use of FT (Ferguson, 1975). A similar study asking native speakers to "transform sentences for foreigners" found that adults felt uncomfortable and sometimes even refused to take part in the research, supporting that native speakers as well hold negative views towards FT (Fedorova, 2018).

Another reason for negative sentiments towards FT is because it shares some features with Baby Talk, a form of language which is regarded by a speech community as being primarily appropriate for talking to young children, and not the normal adult use of language (Ferguson, 1964). FT is similar to baby talk in many ways; both are characterized by frequent pauses, a limited vocabulary, and brevity (Freed, 1981). Additionally, both speech registers include the repetition of words and imitation of foreign speakers (Freed, 1981).

FT Across Diverse Language Communities

Although prior work has looked at FT as a general phenomenon, there is reason to hypothesize that how FT is used and perceived may vary depending on the language and/or ethnic group of the foreign speaker. There are many stereotypes about people based on how they speak; conversely, there exist stereotypes about how people speak based on their identity. Developmental psychologist Katherine Kinzler suggested that someone's language and accent provides a tremendous amount of social information to a listener (Kinzler, 2021). Studying children as young as five years old, Kinzler found that children think about language as a marker for national identity, allowing it to be a vehicle for social grouping (Kinzler, 2013, 2021). As time has progressed, more studies have focused on examining FT towards specific accent groups to see if there is variation within attitudes depending on the foreign individual's race/ethnicity.

The focus of the present study is on English speakers whose first language was Chinese or Spanish, and who did not learn English until later in childhood (age seven or older). The inclusion of Spanish speakers in the study is particularly significant because of the prevalence of Latinos in the U.S. – 18.9% of the U.S. population identifies as Hispanic and/or Latino, many of whom are Spanish speakers (Office of Minority Health, 2023). Native Chinese speakers were included due to previous research indicating negative stereotyping towards Asians in the United States. The rationale for this comparison between languages is that many native English speakers in the U.S. have been found to hold implicit stereotypes that may differ between Chinese and Spanish speakers. Although FT is related to the accents of non-native speakers, it is also hypothesized to intersect with biases associated with beliefs, associations, and stereotypes about people of different national/ethnic/linguistic groups.

Asian FT as a Stereotype

Some studies show that if a person is a member of an Asian group, then they are assumed to have an accent (even when they don't), or are judged harder to understand (even when controlling for speech sample). When controlling for speech presented to participants, research suggests stereotypes about language intelligibility. In Rubin (1992), students heard speech associated with a picture of an Asian face and reported hearing a "heavy accent" that actually did not exist (Rubin, 1992). Another study was conducted with White and South Asian faces to determine if race impacted native English-listeners' perceptions of how another person speaks English. Results revealed that accentedness judgments increased for all varieties when speech was paired with South Asian faces (Kutlu et al., 2022). These cases suggest that upon hearing someone's accent, a listener makes predetermined judgments about how they *will* speak, judgments that are not based on how the person *actually* speaks.

More research, although not directly controlling for speech presented to participants, may also reflect language stereotypes. A study focused on Vietnamese-accented speakers found that native English participants required more attention when listening to a Vietnamese-accented speaker (Hosoda et al., 2007). Similarly, when researchers asked listeners of various backgrounds (Indian, Chinese, Korean, American) to rate the intelligibility of Chinese graduate students, all listeners rated English L1 students (students whose first language is English) as more intelligible than Mandarin L1 students. The only exception was Mandarin L1 listeners,

who found the two groups equally intelligible (Hardman, 2010).

Furthermore, research shows that if a person speaks with an accent from an Asian language, then people have certain stereotypes about what they are like. In an experiment directly testing how much people believe non-native speakers, participants judged trivia statements such as "ants don't sleep" as less true when stated by Korean-accented speakers versus native English speakers (Lev-Ari & Keysar, 2010). Relative to standard American-accented English speakers, Vietnamese-accented English speakers were perceived as poorer communicators who were less potent, less threatening, and more concerned about others (Hosoda et al., 2007). Native-English speakers also perceived Asians to be "highly competitive and technologically-skilled" (Carlson & McHenry, 2006). Another study found that Chinese-accented speakers were deemed less attractive than standard American-accented speakers in the context of a college classroom (Cargile, 1997). These cognitive reactions to Asian-accented English speakers suggest a set of stereotypes regarding characteristics deriving from the accent of someone whose first language was an Asian language.

Previous papers suggest that not all Asian subgroups perceive FT the same. A 2010 study reported that Malaysian students highly valued their accented English, while Japanese and Koreans disapproved of their own varieties of English and indicated their preference for native English pronunciation. Additionally, in the eyes of native speakers, different East Asian ethnicities such as Vietnamese and Mandarin are seen differently, even though they are both Asian – Vietnamese speakers were rated as more foreign by native speakers (Cargile et al., 2010). Due to this finding, rather than consolidating all Asian languages into a single category, the analysis centers on a single language. Given that Chinese is the most widely spoken East Asian language, this study includes only Chinese language participants as an exemplar for the East Asian group.

Spanish FT as a Stereotype

Stereotypes about Latinos in the U.S. are hypothesized to affect how people speak to someone whose first language is Spanish. Individuals speaking English with a Spanish accent are more likely to be judged negatively than someone speaking with an American-English accent: Spanish-accented speakers were rated as less competent, less knowledgeable, and of lower status compared to native English speakers (Cargile et al., 2010; Nelson et al., 2016). Spanish-accented individuals were also rated as less socially attractive than their American-English accented counterparts

(Montgomery & Zhang, 2018). Other perceptions exist surrounding Spanish-accented speakers, such as the conception that they are “poor, efficient, and hardworking” (DuBois, 2017).

In addition to facing character judgments, Latinos in the United States have encountered discrimination linked to their accents, notably in regards to employment opportunities. An empirical study revealed that, in the context of applying for a software engineering position, a hypothetical applicant with a Mexican-Spanish accent was disadvantaged compared to someone with a standard American-English accent (Hosoda et al., 2012). The individual with the Mexican-Spanish accent received lower ratings in terms of job suitability and was perceived as less likely to be promoted to a managerial role (Hosoda et al., 2012). In interviews with foreigners, Spanish-accented speakers said that discrimination took place “once in a while,” and more frequently at work or school, at government offices or in department stores (Korzenny and Schiff, 1987). We hypothesized that accent-based discrimination affects how people treat Latino individuals, and also how these people perceive themselves; microaggressions can be particularly cognitively and emotionally taxing for members of stigmatized groups (Sue et al., 2007).

Comparisons of Chinese and Spanish FT

Prior literature indicates that native English speakers perceive Spanish speakers as “less foreign” than Chinese speakers, leading to a comparison between the two groups. In a study asking Americans to rank “foreignness” of different languages, Asians were rated as “the most foreign,” while Latinos were only rated as “moderately foreign” (Cargile et al., 2010). Another experiment about perceived degrees of accentedness found that native Spanish speakers were judged not to have any accent when speaking English, whereas native Chinese subjects were judged as having a measurable accent – however, these studies did not control for age of English acquisition (AOL), and therefore effects may be due to AOL rather than the speaker’s native language (Flege & Fletcher, 1992). When students listened to verbal recordings paired with an American, Hispanic, or Asian phenotype (face), the students preferred to interact with the Hispanic phenotype when paired with a native English accent (McCrocklin et al., 2018). However, this may have been because 93% of participants in the study were Hispanic, and felt that they related better to a person from a similar background (McCrocklin et al., 2018).

To this date, only one major experiment, conducted by Podberesky (1990), has compared language discrimination directly between Chinese and Spanish

speakers – but this study also clumped Chinese into a category called “Oriental” which also included Japanese, Korean, Vietnamese, etc. Podberesky found that there was no difference in perception of traits amongst the groups. It is important to note, however, that the study group consisted of 19.5-year-old (mean age) college students living in a diverse university town with a high rate of non-native speakers of English, potentially skewing the results. Furthermore, the paper did not exclusively center on FT (Podberesky et al., 1990).

Rationale for Current Study

In light of recent research identifying racial/ethnic differences in how FT is perceived, the current study analyzes non-native speakers’ perceptions of FT, with a specific focus on individuals whose first language is either Chinese or Spanish. The primary investigation analyzes non-native English speakers’ experiences with FT, their attitudes toward FT, and whether these differ as a function of their linguistic/ethnic background.

This thesis adopts a novel perspective because it focuses on FT between these two groups specifically. Very few past studies have compared Chinese and Spanish speakers, and from these, none have directly analyzed FT (Podberesky et al., 1990). Amongst other papers that *did* study FT, none solely focused on a comparison between Chinese and Spanish speakers. Additionally, most past research predominantly explored how FT was used by native speakers, rather than the perception by non-native recipients. The current study will be the first of its kind, integrating a broader examination of language perceptions among Chinese and Spanish speakers with a specific emphasis on the phenomenon of FT.

Research Questions

Based on background knowledge and literature, three research questions were generated about non-native speakers’ perceptions of FT, outlined below:

Q1: Do Chinese and Spanish-speaking participants receive different amounts of FT? We hypothesize that Chinese speakers will receive more FT than Spanish speakers due to biases associated with beliefs/associations/stereotypes about people of Asian descent.

Q2: Do language beliefs and ideologies differ between Chinese versus Spanish speakers? One possibility is that Chinese speakers will have more negative perceptions of FT, feel less close to English speakers, and feel like they belong less in the U.S. due to receiving more FT than they believe is necessary. This would be consistent with prior findings that Chinese speakers are assumed to be less skilled in English than they actually are. Another possibility is that Chinese

speakers will have more positive perceptions of FT, feel more close to English speakers, and feel like they belong more in the U.S. due to having lower English abilities and finding FT helpful in aiding comprehension. This would be consistent with the greater linguistic relatedness and similarity between English and Spanish than between English and Chinese, thus potentially resulting in greater challenges with the English language.

Q3: Do we find language background factors correlated with FT? We hypothesize that age of English acquisition, length of stay in the U.S., English ability, and accent in English will correlate with frequency of receiving FT, attitudes toward FT, favor towards monolingualism, and accent prejudice.

To test the three research questions, an online survey was designed and distributed to U.S. adults both locally and nationwide. Participants were required to be non-native English speakers with either Chinese or Spanish as their first/native language.

Method

All study procedures were approved by the University of Michigan Institutional Review Board (HUM00140836). Data collection occurred in November 2023. A pilot study about non-native speakers' perceptions of FT was conducted in 2018 by Danielle Labotka. The current study focused on the differences in the frequency and attitudes of FT between non-native Spanish and Chinese speakers.

Power Analysis

Power was calculated with G*Power using an effect size of $d = 0.80$ to indicate high practical significance of the research outcome. Power with this sample size is 80%, as this is a generally-accepted power level to avoid Type II errors. The alpha level is 0.05. According to these criteria, 26 participants were needed in each group. To account for attrition and drop-criteria, the study aimed to have a sample size of 30 participants in each category.

Participants

Recruitment was via snowball sampling, a method in which initial participants recruit additional participants from their network, creating a "snowball" effect. Snowball sampling was advantageous in this situation due to the target population being difficult to reach through traditional sampling methods. The process began with the initial identification of a few participant groups (e.g., student groups on campus), then relied on referrals to others. To aim for diversity within the sample, the survey was sent to student organizations, faculty departments, and participants from outside the Ann Arbor area. Participants completed the survey hosted by

Qualtrics, an online platform designed for creating and distributing surveys. All participation was volunteer-based, therefore participants were not compensated for their time. Only residents of the US were allowed to participate. The survey took an average of 10 minutes to complete.

The final sample consisted of 41 Spanish speakers and 23 Chinese speakers. The sample of Spanish speakers is above the goal of 26 determined by the power analysis, but the sample of Chinese speakers is lower. No information was gathered on participant gender. Participants were on average 34 years old, with the youngest participant being 18 and oldest 65 years old ($M = 33.58$, $SD = 13.07$). The sample size learned English at an average age of 11 ($M = 11.11$, $SD = 4.92$) and has been living in the U.S. for around 11 years ($M = 11.36$, $SD = 9.99$). Most of the sample was highly educated, with 56.25% percent receiving a master's degree or higher. Spanish-speaking participants were from 12 different countries, and Chinese participants were from 2 different countries (see Appendix A).

Because the study required participants to be non-native English speakers, any participants who reported English as their first language ($n = 3$) or learned English before the age of 7 ($n = 4$) were removed from the survey. Additionally, participants who were not native Chinese or Spanish speakers were kicked out in prescreening. Once the survey began, to ensure that only participants who were engaged with the task were included in the final sample, two attention checks were presented. As an example, one attention check stated "This is an attention check. Copy the following sentence into the box below: Grass is green and the sky is blue." Participants who did not pass both attention checks were dropped from the sample ($n = 3$). Participants who did not respond to all questions in the survey were dropped ($n = 16$), unless they completely answered the first block of questions, in which case they were included in statistical analyses that pertained to the first block only ($n = 8$). Finally, one bot question was placed at the end of the survey to ensure participants were human, and participants who failed to answer the question were also dropped ($n = 18$).

Materials and Procedure

Each participant filled out a questionnaire that measured their experiences with and opinions of FT, and how FT related to different ideological beliefs. The questionnaire was based on prior research conducted by Danielle Labotka and Susan Gelman (unpublished). Since recruitment was completed via snowball sampling, we assumed that the majority of participants would fit the study's criteria. Nevertheless, before the questions began,

participants were presented with two pre-screening questions: one to ensure they were native Chinese or Spanish speakers, and another to confirm that they learned English after the age of 7. If participants did not pass these pre-screening questions, they were directed to an exit screen and thanked for their time. Participants who passed the screener questions were next asked to provide demographic information, including their age, race/ethnicity, education level, country of origin, and length of residence in the U.S.

Next, participants reported information about their language experience/background, including their native language and age of English acquisition. They rated dimensions of their English ability on a 7-point scale, including how comfortable they are speaking English (1 = Not comfortable at all, 7 = Very comfortable), how fluent they are in English (1 = Not fluent at all, 7 = Very fluent), and how grammatically correct their English is (1 = Not grammatically correct at all, 7 = Completely grammatically correct). Additionally, participants rated how accented their English is on a 7-point scale (1 = Not accented at all, 7 = Very accented). A composite subscale was created combining all four variables, but the alpha did not reach reliability ($\alpha = .66$). Consequently, “accentedness” was dropped, and the final subscale consisted of “comfortability,” “fluency,” and “grammatically correct” – this composite was called *English ability* ($\alpha = .76$).

After the demographics information, participants provided information about their personal attitudes and experiences talking to native English speakers. First, participants rated how much they felt they belonged in the U.S. on a 7-point scale (1 = Don’t belong at all, 7 = Completely belong). They then reported how close they felt to a person when they talked to them in both their native language and in English (1 = Not close at all, 7 = Very close). In the next question, participants were asked how much native speakers changed their speech when talking to them versus a native speaker on a 7-point scale (1 = No change, 7 = Change a lot). Finally, in an open-ended question, they were asked what kinds of changes they notice native speakers make in their speech when talking to them.

Then, the study provided an explanation of FT as “the changes in speech native speakers may make when talking to non-native speakers – characterized by slow, loud, over-enunciated, and simplified speech.” Participants were asked an open-ended question about what changes native speakers make to their language when using FT, then asked to indicate whether they had ever received FT (Yes, Maybe, No). In a check-the-box question, participants selected which aspect(s) of FT they

received (slow speech, loud speech, over-enunciated speech, simplified speech, other).

Participants then provided information regarding their experiences with FT. First, they indicated on a 7-point scale how often they receive FT (1 = Never, 7 = All the time). Second, they answered how close they felt to a person when they talked to them in FT (1 = Not close at all, 7 = Very close). Then, they rated on a 7-point scale how helpful they found FT (1 = Not helpful at all, 7 = Very helpful) and how offended they were by FT (1 = Not offended at all, 7 = Very offended). They also gave social ratings of a person who uses FT, including how much they like them (1 = Not like them at all, 7 = Like them a lot) and the positive/negative impact of using FT on their opinion of a person (1 = No positive/negative impact, 7 = Very positive/negative impact). In the last scale regarding FT, participants rated how much they felt they belonged in the U.S. when they received FT on a 7-point scale (1 = Don’t belong at all, 7 = Completely belong). The block ended with a free-response question asking what a person using FT thinks about them.

The final portion of the survey was dedicated to measures of participants’ American language ideologies. In a 2018 unpublished study, Danielle Labotka adapted scales of language ideologies to measure participants’ acceptance of monolingualism (Fitzsimmons-Doolan, 2011) and accent prejudice (Ura et al, 2015). Participants received 42 statements about language – 21 about monolingualism, 21 about accent prejudice – and rated how much they agreed with each statement on a 6-point scale (1 = Strongly disagree, 6 = Strongly agree). All of the included measures have been found to be valid and reliable in previous research (see Appendix B).

The overall interest of the survey was in attitudes, experiences, and evaluations of FT. Attitudes and experiences were measured with the first section of questions, while evaluations were measured with the second section of pre-scaled questions.

Statistical Analysis

The study used two different statistical tests to analyze the data: t-test and correlation. The first analysis examined the rates of receiving FT, attempting to answer the question “Is there a difference between the amount of FT received between Chinese speakers and Spanish speakers?” The question was posed in order to determine if factors such as racial biases came into play. T-tests were conducted to compare the two different non-native language groups on their frequency of receiving FT (see Appendix C). In a similar analysis, multiple paired-sample t-tests were conducted to investigate differences in how Chinese and Spanish participants evaluated FT (see Appendix C). A final analysis consisted of sets of

bivariate correlations (using Pearson's correlation coefficient) to analyze what factors relate to non-native speakers receiving FT (see Appendix D). All analyses were implemented in SPSS 29.

Results

How do Chinese versus Spanish Speakers Compare in Measures Unrelated to FT?

Before evaluating the proposed research questions, it is important to examine language background and demographic information among the participants to identify similarities and distinctions. The Chinese participants' English ability (a subscale created by combining three questions, $\alpha = .76$) was lower than the average Spanish speakers' English ability ($M_s (SDs) = 5.05 (1.0)$ vs. $5.90 (0.98)$, respectively), $t(62) = 3.26, p = .002$. More specifically, the Chinese speakers reported being less comfortable speaking English ($M_s (SDs) = 5.48 (1.50)$ vs. $6.20 (1.31)$, $p = .05$), less fluent in English ($M_s (SDs) = 5.26 (1.29)$ vs. $6.22 (1.01)$, $p = .002$), and less grammatically-correct when speaking English ($M_s (SDs) = 4.43 (0.99)$ vs. $5.29 (1.31)$, $p = .009$) when compared to Spanish speakers. There was no significant difference in self-reported accent between Chinese-speaking ($M = 4.04, SD = 1.55$) and Spanish-speaking participants ($M = 4.84, SD = 1.68$), $t(62) = 1.88, p = .07$. Additionally, there was no difference in length of stay in the United States between Chinese-speaking ($M = 9.17, SD = 7.41$) and Spanish-speaking participants ($M = 12.66, SD = 11.07$), $t(62) = -1.34, p = .07$, even though Spanish speakers had lived in the U.S. for an on-average greater time. Finally, there was no significant difference in age of English acquisition between Chinese-speaking ($M = 10.0, SD = 3.12$) and Spanish-speaking participants ($M = 11.78, SD = 5.58$), $t(62) = -1.41, p = .05$. In terms of education, 41% percent of Spanish speakers received higher than a bachelor's degree, while 83% percent of Chinese speakers did ($X^2 (1, N = 64) = 10.14, p = .001$). The two groups did not differ in their attitudes towards monolingualism ($M_{\text{Chinese}} = 2.94, SD = 0.55$; $M_{\text{Spanish}} = 2.65, SD = 0.61$; $t(54) = -1.79, p = .08$) or prejudice against accents ($M_{\text{Chinese}} = 2.06, SD = 0.64$; $M_{\text{Spanish}} = 1.82, SD = 0.53$; $t(54) = -1.52, p = .14$).

How do Chinese versus Spanish Speakers Compare in Frequency of Receiving FT?

Participants were asked if they ever received FT, given the choices of yes, no, or maybe. Only 3% said 'no'; 64% said 'yes' and 33% said 'maybe.' Research Question 1 asked whether we will find a difference in receiving FT between Chinese and Spanish-speaking participants. The hypothesis was not supported according to two t -tests. When calculating the difference in ever receiving FT (Yes

= 2, Maybe = 1, No = 0), there was no significant difference between Spanish speakers ($M = 1.61, SD = 0.59$) and Chinese speakers ($M = 1.61, SD = 0.50$), $t(61) = 0.007, p = .99$. Similarly, when calculating the difference in frequency of receiving FT (on a scale of 1-7), there was also no significant difference between Spanish speakers ($M = 3.00, SD = 1.39$) and Chinese speakers ($M = 2.83, SD = 0.98$), $t(62) = 0.53, p = .60$.

How do Chinese versus Spanish Speakers Evaluate FT?

Research Question 2 asked whether we find a difference in language beliefs and ideologies among Chinese versus Spanish speakers. This hypothesis was partially supported.

A subscale for *attitudes towards FT* combined seven questions, all relating to non-native speakers' attitudes towards receiving FT, with a higher number representing a more favorable opinion of FT ($\alpha = .91$). Chinese participants ($M = 4.14, SD = 1.21$) have more favorable attitudes towards receiving FT than Spanish speakers ($M = 3.18, SD = 1.61$), $t(62) = -2.48, p = .02$.

In terms of *social distance when receiving FT*, Chinese participants ($M = 3.43, SD = 1.47$) felt more close to a person speaking to them in FT than Spanish speakers ($M = 2.39, SD = 1.71$), $t(62) = -2.46, p = .01$. An additional analysis found that Chinese participants ($M = 5.43, SD = 1.53$) have a significantly lower average than Spanish participants ($M = 6.34, SD = 0.88$) in terms of *closeness to a person when speaking native language*, meaning that Chinese participants do not feel as close to others speaking Chinese as Spanish speakers feel close to others speaking Spanish, $t(62) = 3.01, p = .004$.

There was no significant effect for *feelings of belongingness* in the U.S., $t(62) = 1.11, p = .27$, despite Chinese speakers ($M = 4.07, SD = 1.78$) attaining higher scores than Spanish speakers ($M = 3.57, SD = 1.72$). Additionally, there was no significant difference in *feelings of belonging in the U.S. when receiving FT* between Chinese speakers ($M = 2.78, SD = 1.68$) and Spanish speakers ($M = 2.68, SD = 1.94$), $t(62) = -2.07, p = .04$.

What Factors Correlate with Non-Native Speakers Receiving FT?

The final Research Question asked whether we find language background factors correlated with FT. Specifically, Q3 hypothesized that age of English acquisition, length of stay in the U.S., English ability, and self-reported accent in English would correlate with frequency of receiving FT, positive attitudes toward FT, favorable attitudes towards monolingualism, and accent prejudice. The hypothesis was partially supported by the results.

As predicted, *frequency of receiving FT* was negatively correlated with English ability ($r = -.36$), meaning those with higher self-reported English ability received less FT. Also as predicted, frequency of receiving FT was negatively correlated with length of stay in the U.S. ($r = -.36$), suggesting that participants who have lived in the U.S. longer received less FT. Frequency of receiving FT was positively correlated with accent in English ($r = .32$), meaning the greater one's self-reported accent, the more FT they receive. There was no significant correlation between frequency of receiving FT and age of English acquisition ($r = .11$).

Positive attitudes toward FT were negatively correlated with English ability ($r = -.57$), suggesting those with higher self-reported English ability had more negative attitudes towards FT. There was no significant correlation between attitudes toward FT and accent in English ($r = .20$), age of English acquisition ($r = .23$), or length of stay in the U.S. (.02).

Favorable attitudes towards monolingualism were not significantly correlated with English ability ($r = -.12$), accent in English ($r = .07$), age of English acquisition ($r = -.1$), or length of stay in the U.S. ($r = .09$). However, favorable attitudes towards monolingualism were positively correlated with positive attitudes toward FT ($r = .34$), meaning participants who favored the use of one language felt more positively about using FT.

Finally, *accent prejudice* was negatively correlated with respondents' accent in English ($r = -.31$), suggesting that those with a greater self-reported accent had less prejudice towards different accents. Accent prejudice was not significantly correlated with English ability ($r = -.10$), age of English acquisition ($r = -.29$), or length of stay in the U.S. ($r = -.28$). Accent prejudice was, however, highly positively correlated with favorable attitudes towards monolingualism ($r = .53$), meaning participants with greater prejudice against accents also possessed monolingualistic attitudes, favoring the use of one language.

Open-Ended Responses

Open-ended responses were collected from participants to assess their thoughts on and views of FT from a non-native English speaker's perspective. First, participants were asked, "What do you think are some changes native speakers make to their language when using FT?" Most responses mentioned that native speakers use simplified speech, talk slower, and use less complicated language. One Chinese-speaking participant reported that native speakers are "trying to paraphrase certain terms or sentences and try to add explanations for cultural-specific or generation-specific terms used in the U.S." Similarly, a Spanish-speaking participant wrote that

native speakers use "intonation of certain words that 'sounds like a teacher.'" Some participants also thought native speakers' use of FT reflected negative attitudes, for example, "They think we are newcomers here and don't understand English... or they think someone with an accent is stupid. I've had to call it out exactly like that."

Participants were also asked, "What do you think a person using FT thinks about you?" Responses were coded as negative, positive, both, neither, or no response. The majority of participants (55%) reported negative responses, such as, "They think that because I'm foreign I'm stupid." In contrast, 19% of participants reported positive responses, such as, "They are considerate and willing to include me into their conversation." Finally, 9% of responses reported both negative and positive responses, depending on the person or situation. For example, one Chinese-speaking participant wrote, "Some of them are nice and try to have an inclusive, friendly conversation with me. However, there are also people who inherently think I don't belong to the same community they belong to and try to separate or 'distinct' them from me using FT." A Spanish-speaking participant likewise wrote, "I think it depends on the situation, sometimes we might need to simplify our speech. However, that person may think that I'm not smart enough to follow the pace of the conversation." The remaining 5% of participants fit into the 'neither' category, and 12% of participants did not respond to the question. Open-ended questions highlighted the double-message surrounding the use of FT.

Discussion

The phenomenon of FT has a daily impact on millions of Spanish and Chinese speakers in the United States. The current study is the first to examine FT from the perspective of the non-native speaker, involving a direct comparison of Chinese and Spanish speakers. The study found that the majority of participants (64%) reported receiving FT at least once in their lives, suggesting the concept is still apparent in the modern day. The study explored variations in FT across different language groups, attempting to understand perceptions and attitudes towards FT.

The primary finding of the study was that Chinese-speaking and Spanish-speaking participants had different perceptions of FT; Chinese-speaking participants exhibited significantly more favorable attitudes toward receiving FT and felt closer to a native speaker talking to them using FT compared to Spanish speakers. Conversely, Chinese participants felt more social distance to someone speaking their own language when compared to Spanish participants. This finding was true despite there being no significant difference in the

amount of FT received between Chinese-speaking and Spanish-speaking participants. Across both groups, the longer a participant had lived in the United States and the less of an accent they reported having, the less likely they were to receive FT. Participants with lower English ability found receiving FT more helpful among both language groups. Additionally, people with greater accents had less prejudice towards others with accents.

Frequency of receiving FT between Chinese and Spanish participants failed to support the original hypothesis and literature. The hypothesis presumed that, given the distinct stereotypes associated with Chinese and Spanish speakers in terms of language and personality, Chinese participants would encounter more instances of FT (Cargile et al., 2010; Flege & Fletcher, 1992; McCrocklin et al., 2018). However, the present study indicates that the participants' native language or origin did not significantly influence the amount of FT they receive; why this is the case will require more research. One possibility is that what matters more is the native-speaker's perception of the participant as "foreign," emphasizing the perceived difference and the need for an altered communication approach. Volker Hinnenkamp called this idea the generalized foreigner, explaining that non-native people are categorized and generalized as being "alien" (Hinnenkamp, 1987). Similarly, a 2013 study examining Asians and Latinos found that 91% of participants reported having experienced national origin identity neglect, a phenomenon when people treat members of distinct social groups as interchangeable parts of a broader category (Flores & Huo, 2013). Many native English speakers may simply see both Chinese and Spanish participants as "non-native," and chose to communicate with them both using FT.

Importantly, although the frequency of receiving FT may be comparable between both groups, the perceptions of FT differed among participants. The study revealed that Chinese participants found FT more helpful and felt closer to someone using FT with them. Given that Chinese participants in the study reported having lower levels of English ability, it is perhaps not surprising that Chinese-speakers would have more positive perceptions of someone using a form of speech with them that may be easier to process. One paper suggested that using FT "may serve to facilitate a foreign learner's comprehension by providing the morphophonemically fullest forms of the target language," possibly providing helpful language cues to someone with lower English ability (Valdman, 1981).

Another interesting finding from the current study was that Chinese speakers felt less close to other

people speaking their native language compared to Spanish speakers. This may suggest that Spanish-speaking participants exhibited a more pronounced in-group bias, feeling a greater sense of closeness to others who also speak Spanish. In a longitudinal study examining Spanish children's preferences, it was found that Spaniards significantly preferred their own group over others, allocating more positive adjectives towards their in-group (Enesco et al., 2011). On the other hand, when analyzing Chinese-speakers, numerous studies suggest that Asian-Americans experience a phenomenon called identity denial, in which they realize they are perceived as "less American than other Americans" and consequently claim greater participation in American practices, which could include rejecting people from their own native ethnicity (Cheryan & Monin, 2005; Wang et al., 2013).

Lastly, when examining correlations among characteristics related to receiving FT, it was found that the longer someone lived in the United States and the lower their self-reported accent, the less often they received FT. This outcome aligns with expectations, as both Chinese and Spanish speakers show cultural assimilation over the course of their stay in the United States. Previous research shows that the age at which someone learns a second language is a predictor for second-language accent (Moyer, 2018). Additionally, those who speak English with a non-native accent have been shown to receive stigma and discrimination in the United States, which could include receiving more FT (Sener, 2021). Participants in the current study who had lower English abilities actually found this use of FT as more helpful, suggesting that FT can also facilitate positive effects in aiding comprehension (Bobb et al., 2019).

As anticipated by Q3, participants with a higher self-reported accent showed less prejudice towards accents. One potential explanation for this finding is that people with accents tend to be more open-minded to how others speak, enabling them to understand language from another's point of view. Supporting this idea, a study evaluating attitudes towards others' foreign accents revealed that participants who grew up in an ethnically diverse environment, who had lived abroad and who were working in an ethnically diverse environment exhibited significantly more positive attitudes towards foreign accents (Dewaele & McCloskey, 2015).

Implications of Findings

The outcomes of this study can contribute to a deeper understanding of FT and highlight the sentiments of non-native speakers regarding interpersonal communication. Here we discuss two implications: the

impact of FT on non-native speakers, and potential consequences for native speakers employing FT.

Our data are consistent with Marlow's (2009) discussion of FT in terms of the receiver model of language criticism, suggesting that FT must be studied from the perspective of how it makes the non-native speaker feel. The model proposes that individual, contextual, and situational variables influence how people evaluate and respond to language criticism. Thus, it makes sense that some of the non-native speakers in the current study were adaptive and accommodating to receiving FT, while others may have experienced repressed anger, avoidance, or aggression (Marlow & Giles, 2010).

From the native-speaker's perspective, we hope that the findings of the current study can help speakers become cognizant of when they are using FT and whether their own preconceived notions are causing them to change their form of speech. One of the key findings of the CAT theory revealed that speakers accommodate not to where their recipients are in an objective sense, but rather where they are believed to be communicatively (Thakerar et al., 1982). Sometimes, speakers also change their speech to sound like the "prototype" of what they think a member of their recipient's in-group would sound like (Gallois & Callan, 1988). Through further education, as facilitated by the current study, native speakers can gain a deeper understanding that, despite positive intentions, their speech may inadvertently lead to miscommunication.

It is also important for all speakers to recognize that having an accent is not a hindrance, and foreign accents come with various social advantages. Over the years, research has demonstrated that non-native speakers are perceived as friendlier and more competent in the context of manual labor (Luhman, 1990; Giles et al., 1992). A recent study conducted in 2023 further revealed that speakers with accents were rated as more trustworthy and having more interpersonal appeal (Ip & Papafragou, 2023). Non-native speakers were partly protected against the social disadvantages associated with being "underinformative" compared to their native counterparts, indicating a phenomenon called social lenience towards non-native speakers (Ip & Papafragou, 2023).

Limitations

The study encountered a few limitations. For one, the sample size was notably small and the population of Chinese-speaking participants was below the goal determined by the power analysis. Consequently, generalizing results beyond the specific sample under investigation becomes challenging. Moreover, since the

survey was conducted in English, individuals with limited English proficiency were unable to participate due to language barriers. This limitation could be addressed by offering the study in multiple languages to accommodate participants from diverse linguistic backgrounds. While the survey aimed for nationwide coverage, a significant portion of participants resided in Michigan, potentially impacting external validity. When collecting demographic data, the gender of participants was never asked, resulting in the inability to analyze and report gender-specific results. Lastly, the study was survey-based, leaving room for response bias and limited depth of exploring complex sentiments.

Future Directions

This study represents an initial step toward comprehending the use of FT, as well as the attitudes and perceptions of native speakers regarding FT. Moving forward, further research is imperative, particularly in regards to the use of FT with Chinese and Spanish speakers, given their significant populations in the United States. Subsequent investigations could involve experimental studies where participants receive FT from an actor, allowing for a direct assessment of their real-time reactions; this experimental approach offers the potential for analyzing responses in the moment, contrasting with reliance on self-reported measures obtained through surveys. A future focus on FT research holds the promise of eventually facilitating meaningful changes in the way we communicate with others.

References

- Beebe, L. M., & Giles, H. (1984). Speech-accommodation theories: A discussion in terms of second-language acquisition. <https://doi.org/10.1515/ijsl.1984.46.5>
- Bobb, S. C., Mello, K., Turco, E., Lemes, L., Fernandez, E., & Rothermich, K. (2019). Second language learners' listener impressions of foreigner-directed speech. *Journal of Speech, Language, and Hearing Research*, 62(9), 3135-3148. https://doi.org/10.1044/2019_JSLHR-S-18-0392
- Cargile, A. C. (1997). Attitudes toward Chinese-accented speech: An investigation in two contexts. *Journal of Language and Social Psychology*, 16(4), 434-443. <https://doi.org/10.1177/0261927X970164004>
- Cargile, A. C., Maeda, E., Rodriguez, J., & Rich, M. (2010). "Oh, you speak English so well!": US American listeners' perceptions of "foreignness" among nonnative speakers. *Journal of Asian American Studies*, 13(1), 59-79. <https://doi.org/10.1353/jaas.0.0062>

- Carlson, H. K., & McHenry, M. A. (2006). Effect of accent and dialect on employability. *Journal of Employment Counseling*, 43(2), 70-83. <https://doi.org/10.1002/j.2161-1920.2006.tb00008.x>
- Cheryan, S., & Monin, B. (2005). Where are you really from?: Asian Americans and identity denial. *Journal of Personality and Social Psychology*, 89(5), 717. <https://doi.org/10.1037/0022-3514.89.5.717>
- Dela Rosa, J. P. O., & Arguelles, D. C. (2016). Do Modification and Interaction Work?--A Critical Review of Literature on the Role of Foreigner Talk in Second Language Acquisition. *Journal on English Language Teaching*, 6(3), 46-60. <https://eric.ed.gov/?id=EJ1131419>
- Derwing, T. M., Rossiter, M. J., & Munro, M. J. (2002). Teaching native speakers to listen to foreign-accented speech. *Journal of Multilingual and Multicultural Development*, 23(4), 245-259. <https://doi.org/10.1080/01434630208666468>
- Dewaele, J. M., & McCloskey, J. (2015). Attitudes towards foreign accents among adult multilingual language users. *Journal of Multilingual and Multicultural Development*, 36(3), 221-238. <https://doi.org/10.1080/01434632.2014.909445>
- Drljača Margić, B. (2017). Communication courtesy or condescension? Linguistic accommodation of native to non-native speakers of English. *Journal of English as a Lingua Franca*, 6(1), 29-55. <https://doi.org/10.1515/jelf-2017-0006>
- DuBois, S. (2017). Identification and stereotypes of foreign-accented English. *Notandum*, (46), 7-28. <http://dx.doi.org/10.4025/notandum.46.1>
- Enesco, I., Guerrero, S., Lago, M., & Rodríguez, P. (2011). How stable is children's affective orientation toward different ethnic groups? A longitudinal study of in-group and out-group attitudes among preschoolers. *Anales de Psicología*, 27(3). <http://hdl.handle.net/10201/26573>
- Fitzsimmons-Doolan, S. (2011). Language ideology dimensions of politically active Arizona voters: An exploratory study. *Language Awareness*, 20(4), 295-314. <https://doi.org/10.1080/09658416.2011.598526>
- Fedorova, K. (2018). 6. Foreigner Talk. *Registers of Communication*, 18.
- Ferguson, C. A. (1964). Baby Talk in Six Languages. *American Anthropologist*, 66(6), 103-114. <http://www.jstor.org/stable/668164>
- Ferguson, C. A. (1971). Absence of copula and the notion of simplicity. *Pidginization and Creolization of Languages*, 141-150.
- Ferguson, C. A. (1975). Toward a Characterization of English Foreigner Talk. *Anthropological Linguistics*, 17(1), 1-14. <http://www.jstor.org/stable/30027270>
- Flege, J. E., & Fletcher, K. L. (1992). Talker and listener effects on degree of perceived foreign accent. *The Journal of the Acoustical Society of America*, 91(1), 370-389. <https://doi.org/10.1121/1.402780>
- Flores, N. M., & Huo, Y. J. (2013). "We" are not all alike: Consequences of neglecting national origin identities among Asians and Latinos. *Social Psychological and Personality Science*, 4(2), 143-150. <https://doi.org/10.1177/1948550612449025>
- Freed, B. (1981). Foreigner Talk, Baby Talk, Native Talk. *International Journal of the Sociology of Language*, 1981(28), 19-40. <https://doi.org/10.1515/ijsl.1981.28.19>
- Gallois, C., & Callan, V. J. (1988). Communication accommodation and the prototypical speaker: Predicting evaluations of status and solidarity. *Language & Communication*. [https://doi.org/10.1016/0271-5309\(88\)90022-5](https://doi.org/10.1016/0271-5309(88)90022-5)
- Giles, H., Henwood, K., Coupland, N., Harriman, J., & Coupland, J. (1992). Language attitudes and cognitive mediation. *Human Communication Research*, 18(4), 500-527. <https://doi.org/10.1111/j.1468-2958.1992.tb00570.x>
- Hatch, E. M. (1983). Psycholinguistics: A second language perspective. Newbury House Publishers, Inc., Rowley, MA 01969.
- Hardman, J. B. (2010). *The intelligibility of Chinese-Accented English to International and American Students at a US University* (Doctoral dissertation, The Ohio State University). http://rave.ohiolink.edu/etdc/view?acc_num=os_u1276818520
- Hinnenkamp, V. (1987). Foreigner talk, code-switching and the concept of trouble. *Analyzing Intercultural Communication*, 137-180. <https://doi.org/10.1515/9783110874280.137>
- Hosoda, M., Nguyen, L. T., & Stone-Romero, E. F. (2012). The effect of Hispanic accents on employment decisions. *Journal of Managerial Psychology*, 27(4), 347-364. <https://doi.org/10.1108/02683941211220162>
- Hosoda, M., Stone-Romero, E. F., & Walter, J. N. (2007). Listeners' cognitive and affective reactions to English speakers with standard American English and Asian accents. *Perceptual and Motor Skills*, 104(1), 307-326. <https://doi.org/10.2466/pms.104.1.307-326>
- Ip, M., & Papafragou, A. (2023). The pragmatics of accents: The social costs and benefits of being a non-native speaker. *Journal of Experimental Psychology: Language, Memory and Cognition*. <https://doi.org/10.1037/xlm0001197>
- Kinzler, K. D. (2021). Language as a social cue. *Annual review of psychology*, 72, 241-264.

- <https://doi.org/10.1146/annurev-psych-010418-103034>
- Kinzler, K. D., & DeJesus, J. M. (2013). Children's sociolinguistic evaluations of nice foreigners and mean Americans. *Developmental Psychology*, 49(4), 655. <https://doi.org/10.1037/a0028740>
- Korzenny, F., & Schiff, E. (1987). Hispanic perceptions of communication discrimination. *Hispanic Journal of Behavioral Sciences*, 9(1), 33-48. <https://doi.org/10.1177/073998638703090103>
- Kutlu, E., Tiv, M., Wulff, S., & Titone, D. (2022). The impact of race on speech perception and accentedness judgements in racially diverse and non-diverse groups. *Applied Linguistics*, 43(5), 867-890. <https://doi.org/10.1093/applin/amab072>
- Lev-Ari, S., & Keysar, B. (2010). Why don't we believe non-native speakers? The influence of accent on credibility. *Journal of Experimental Social Psychology*, 46(6), 1093-1096. <https://doi.org/10.1016/j.jesp.2010.05.025>
- Lugrin, B., Eckstein, B., Bergmann, K., & Heindl, C. (2018). Adapted foreigner-directed communication towards virtual agents. In *Proceedings of the 18th International Conference on Intelligent Virtual Agents* (pp. 59-64). <https://doi.org/10.1145/3267851.3267859>
- Luhman, R. (1990). Appalachian English stereotypes: language attitudes in Kentucky. *Language in Society*, 19(3), 331-348. <https://doi.org/10.1017/S0047404500014548>
- Marlow, M. L. (2009). Race, power, and language criticism: The case of Hawai'i. *University of California, Santa Barbara*. <https://www.proquest.com/dissertations-theses/race-power-language-criticism-case-hawaii/docview/304857357/se-2>
- Marlow, M. L., & Giles, H. (2010). "We won't get ahead speaking like that!" Expressing and managing language criticism in Hawai'i. *Journal of Multilingual and Multicultural Development*, 31(3), 237-251. <https://doi.org/10.1080/01434630903582714>
- McCrocklin, S. M., Blanquera, K. P., & Loera, D. (2018). Student perceptions of university instructor accent in a linguistically diverse area. *Pronunciation in Second Language Learning and Teaching Proceedings*, 9(1).
- McGlone, M. S., & Giles, H. (2011). Language and interpersonal communication. *The SAGE Handbook of Interpersonal Communication*, 4, 201-238.
- Montgomery, G., & Zhang, Y. B. (2018). Intergroup anxiety and willingness to accommodate: Exploring the effects of accent stereotyping and social attraction. *Journal of Language and Social Psychology*, 37(3), 330-349. <https://doi.org/10.1177/0261927X17728361>
- Moyer, A. (2018). An Advantage for Age? Self-Concept and Self-Regulation as Teachable Foundations in Second Language Accent. *CATESOL Journal*, 30(1), 95-112.
- Nelson Jr, L. R., Signorella, M. L., & Botti, K. G. (2016). Accent, gender, and perceived competence. *Hispanic Journal of Behavioral Sciences*, 38(2), 166-185. <https://doi.org/10.1177/0739986316632319>
- Office of Minority Health. (2023). Hispanic/Latino Health. [https://minorityhealth.hhs.gov/hispaniclatino-health#:~:text=Overview%20\(Demographics\)&text=According%20to%202020%20Census%20data,a.group%20after%20non%20Hispanic%20white](https://minorityhealth.hhs.gov/hispaniclatino-health#:~:text=Overview%20(Demographics)&text=According%20to%202020%20Census%20data,a.group%20after%20non%20Hispanic%20white)
- Papoušek, M., & Hwang, S. F. C. (1991). Tone and intonation in Mandarin babytalk to presyllabic infants: Comparison with registers of adult conversation and foreign language instruction. *Applied Psycholinguistics*, 12(4), 481-504. <https://doi.org/10.1017/S0142716400005889>
- Podberesky, R., Deluty, R. H., & Feldstein, S. (1990). Evaluations of Spanish-and Oriental-accented English speakers. *Social Behavior and Personality: an International Journal*, 18(1), 53-63. <https://doi.org/10.2224/sbp.1990.18.1.53>
- Rubin, D. L. (1992). Nonlanguage factors affecting undergraduates' judgments of nonnative English-speaking teaching assistants. *Research in Higher Education*, 33, 511-531. <https://doi.org/10.1007/BF00973770>
- Sener, M. Y. (2021). English with a Non-Native Accent as a Basis for Stigma and Discrimination in the United States. *Dignity in Movement*, 99.
- Sue, D. W., Capodilupo, C. M., Torino, G. C., Bucceri, J. M., Holder, A., Nadal, K. L., & Esquilin, M. (2007). Racial microaggressions in everyday life: implications for clinical practice. *American Psychologist*, 62(4), 271. <https://doi.org/10.1037/0003-066X.62.4.271>
- Thakerar, J. N., Giles, H., & Cheshire, J. (1982). Psychological and linguistic parameters of speech accommodation theory. In C. Fraser & K.R. Scherer (Eds.), *Advances in the Social Psychology of Language*, 205-255. https://10.1007/978-1-349-25582-5_19
- Ura, M., Preston, K. S., & Mearns, J. (2015). A measure of prejudice against accented English (MPAAE) scale development and validation. *Journal of Language and Social Psychology*, 34(5), 539-563. <https://doi.org/10.1177/0261927X15571537>
- Uther, M., Knoll, M. A., & Burnham, D. (2007). Do you speak E-NG-LI-SH? A comparison of foreigner- and infant-directed speech. *Speech*

Communication, 49(1), 2-7.
<https://doi.org/10.1016/j.specom.2006.10.003>

Valdman, A. (1981). Sociolinguistic aspects of foreigner talk. <https://doi.org/10.1515/ijsl.1981.28.41>

Wang, J., Minervino, C., & Cheryan, S. (2013). Generational differences in vulnerability to identity denial: The role of group identification. *Group Processes & Intergroup Relations*, 16(5), 600-617.

<https://doi.org/10.1177/1368430212461963>

Zuengler, J. (1991). Accommodation in native-nonnative interactions: Going beyond the “what” to the “why” in second-language research. *Contexts of accommodation: Developments in Applied Sociolinguistics*, 223-244.

Author Note

Correspondence regarding this paper should be addressed to Ashna Mehra by email: ashmehra@umich.edu

Appendix A
Countries of origin

Spanish speakers	Colombia, Mexico, Argentina, Spain, Honduras, Chile, Ecuador, Uruguay, Dominican Republic, Puerto Rico, USA, Venezuela
Chinese speakers	Taiwan, China

Appendix B
Participants’ Language Ideologies, Scaled Questions

<u>Question</u>	<u>Ideology tested</u>	<u>Source</u>
I think that speakers with accents should learn to speak English better	Accent prejudice	Ura et al., 2015
Speakers with accents cannot speak English as well as native English speakers.	Accent prejudice	Ura et al., 2015
Speakers with accents cannot express their opinions as well as native English speakers.	Accent prejudice	Ura et al., 2015
Speakers with accents have more trouble interacting with others than native English speakers.	Accent prejudice	Ura et al., 2015
Speakers with accents cannot lead a group discussion as well as native English speakers.	Accent prejudice	Ura et al., 2015
Speakers with accents are less informative than native English speakers.	Accent prejudice	Ura et al., 2015
Speakers with accents are less adapted to American culture than native English speakers.	Accent prejudice	Ura et al., 2015
Speakers with accents cannot be Americans.	Accent prejudice	Ura et al., 2015

Speakers with accents do not belong to America as much as native English speakers.	Accent prejudice	Ura et al., 2015
In America, it is best to speak English without a foreign accent.	Accent prejudice	Labtoka, 2018
Speaking native sounding English will bring more respect than speaking with accents.	Accent prejudice	Ura et al., 2015
An eyewitness with an accent is less trustworthy than an eyewitness who is a native English speaker.	Accent prejudice	Ura et al., 2015
Those speaking English with accents hold lower status than native English speakers.	Accent prejudice	Ura et al., 2015
Those speaking English with accents are less prestigious than native English speakers.	Accent prejudice	Ura et al., 2015
Speakers with accents are less intelligent than native English speakers.	Accent prejudice	Ura et al., 2015
It is harder to be friends with a person who speaks with an accent.	Accent prejudice	Labtoka, 2018
When speakers get rid of their accent, they can be more successful in their career.	Accent prejudice	Ura et al., 2015
I can work better with an accent-free coworker than a coworker with an accent.	Accent prejudice	Ura et al., 2015
Speakers with accents should not attain as many management positions as native English speakers.	Accent prejudice	Ura et al., 2015
I do not prefer to work under a boss with an accent.	Accent prejudice	Ura et al., 2015
A professor with an accent is less effective than a professor with native English.	Accent prejudice	Ura et al., 2015
The use of more than one language creates social problems.	Acceptance of monolingualism	Fitzsimmons-Doolan, 2011
The use of more than one language makes social unity difficult.	Acceptance of monolingualism	Fitzsimmons-Doolan, 2011
The use of more than one language makes social mobility difficult.	Acceptance of monolingualism	Fitzsimmons-Doolan, 2011
Using one language to complete a task is better than using two languages.	Acceptance of monolingualism	Fitzsimmons-Doolan, 2011
In the United States, knowing English helps a person to be American.	Acceptance of monolingualism	Fitzsimmons-Doolan, 2011
Language represents national identity.	Acceptance of monolingualism	Fitzsimmons-Doolan, 2011
The standard or model form of a language is the most appropriate form for school.	Acceptance of monolingualism	Fitzsimmons-Doolan, 2011

In the United States, English is more normal than other languages.	Acceptance of monolingualism	Fitzsimmons-Doolan, 2011
The use of language is a human right.	Acceptance of monolingualism	Fitzsimmons-Doolan, 2011
Native languages are beautiful.	Acceptance of monolingualism	Fitzsimmons-Doolan, 2011
In the United States, the use of native languages other than English is helpful for sharing tradition.	Acceptance of monolingualism	Fitzsimmons-Doolan, 2011
Speakers have the right to choose the language they will use in any situation.	Acceptance of monolingualism	Labotka, 2018
Schools must teach native languages of students.	Acceptance of monolingualism	Fitzsimmons-Doolan, 2011
In the United States, the use of multiple languages is an economic asset.	Acceptance of monolingualism	Fitzsimmons-Doolan, 2011
In the United States, the use of more than one language should be promoted.	Acceptance of monolingualism	Fitzsimmons-Doolan, 2011
In the United States, public communication should occur in English.	Acceptance of monolingualism	Fitzsimmons-Doolan, 2011
The success of a nation depends on the use of a national language.	Acceptance of monolingualism	Fitzsimmons-Doolan, 2011
English is the official language of the United States.	Acceptance of monolingualism	Labotka, 2018
It is patriotic to speak English in the United States.	Acceptance of monolingualism	Labotka, 2018
Immigrants to the United States should learn English.	Acceptance of monolingualism	Labotka, 2018
Immigrants to the United States should teach their children their native language.	Acceptance of monolingualism	Labotka, 2018

Appendix C

T-test: How do Spanish versus Chinese speakers evaluate FT, and who receives it more?

	Spanish	English	t-Test
	<i>M (SD)</i>	<i>M (SD)</i>	<i>t (df)</i>
Receiving FT	1.61 (0.59)	1.61 (0.50)	$t(61) = 0.007, p = .99$
Frequency of receiving FT	3.00 (1.39)	2.83 (0.98)	$t(62) = 0.53, p = .60$
English ability	5.90 (1.00)	5.05 (0.98)	$t(62) = 3.26, p = .002^*$
Attitudes toward receiving FT	3.18 (1.61)	4.14 (1.21)	$t(62) = -2.48, p = .02^*$
Favor towards monolingualism	2.65 (0.61)	2.94 (0.55)	$t(54) = -1.79, p = .08$
Prejudice against accents	1.82 (0.53)	2.06 (0.64)	$t(54) = -1.52, p = .14$
Closeness to someone speaking native language	6.34 (0.88)	5.43 (1.53)	$t(62) = 3.01, p = .004^*$
Closeness to someone speaking English	4.34 (1.68)	4.48 (1.20)	$t(62) = -0.34, p = .73$
Closeness to someone after receiving FT	2.39 (1.71)	2.43 (1.47)	$t(62) = -2.46, p = .02^*$
Feelings of belonging in the U.S.	4.07 (1.78)	3.57 (1.72)	$t(62) = 1.11, p = .27$
Feelings of belonging after receiving FT	2.68 (1.94)	2.78 (1.68)	$t(62) = -2.07, p = .04$

* T-test is significant at the 0.05 level (2-tailed)

Appendix D
Correlations: What factors correlate to non-native speakers receiving FT?

	Positive attitudes toward FT	English ability	Accent in English	Age of English acquisition	Length of stay in the U.S.	Accent prejudice	Favor towards monolingualism
Frequency of receiving FT	.22	-.36*	.32*	-.11	-.36*	.04	.01
Positive attitudes toward FT		-.57*	.20	.23	.02	.15	.34*
English ability			-.24	-.06	.27	-.10	-.12
Accent in English				.41*	.03	-.31*	-.07
Age of English acquisition					.56*	-.29	-.10
Length of stay in the U.S.						-.28	-.09
Accent prejudice							.53*

* Correlation is significant at the 0.01 level (2-tailed)

THE FATHER FACTOR: EXPLORING THE RELATIONSHIP BETWEEN PATERNAL INVOLVEMENT AND YOUNG ADULT MENTAL HEALTH

EMILY GULL, AMANDA ELBASSIOUNY, AND LAUREN BANKS

CALIFORNIA LUTHERAN UNIVERSITY

Abstract – The current study sought to explore the relationship between involvement from a father figure and young adults' reports on their experience of anxiety during childhood. Method: This was examined by utilizing Wave IV of the National Longitudinal Study of Adolescent to Adult Health (ADD Health) dataset. Various aspects of paternal involvement, including feelings of closeness and living close by, and symptoms of anxiety were analyzed. Results: Young adults who were not diagnosed with anxiety or panic disorder were seen to be very close to their father figure. Young adults who were not at all close or somewhat close to their father figure reported feeling more isolated than those who were quite close or very close with their father figure. Further, the closer a young adult lived to their father figure, the less they worried about things, but only for those who also live with or close to their mother figure. Discussion: The findings of this study have implications for the fields of psychology, sociology, education, and social work on the important for the role of the father figure in young adults' mental health. Implications and outcomes when fathers are included in care were discussed.

Keywords: paternal involvement, father figure, anxiety, mental health, young adult, ADD Health

During adolescent years, it is common for teens to experience mental health issues because of stressors due to home, school, and extracurricular activities, which can ultimately affect one's daily life (Berger, 2018). About 32% of adolescents claim to have an anxiety disorder and 19.1% of adults reported that they had an anxiety disorder in the past (National Institute of Mental Health, n.d.). Based on the American Psychological Association (n.d.), anxiety can be defined as an emotion that is associated with feeling tense, having thoughts about things that cause one to be worried, and physiological responses (such as an elevated blood pressure). Many of those symptoms are present in other anxiety disorders too. It is important to understand that those symptoms affect adolescents and young adults because it can lead to harm, whether mentally or physically to themselves or others. In addition to understanding adolescents' and young adults' possible signs of anxiety, it is important to study the relationship parents can have, specifically looking at father figures, on their child and if that is related to how anxiety is experienced. Understanding the connection between fathers and anxiety in children and adolescents has not been studied extensively in the research, however, the role of fathers can play an integral role in

buffering against anxiety (Bögels & Phares, 2008). Therefore, the purpose of this study was to understand the association between young adults' reports of their childhood connections with their fathers and how it relates to their experiences with anxiety.

There are possible factors that have been linked to parental support and anxiety among adolescents since anxiety is known to be associated with possible factors of paternal support. Therefore, if one or more of these factors are present, the more anxious these adolescents may feel. These parental factors range from parental overprotectiveness, lack of parental support, negative parenting styles, and parental over-controlling tendencies (Gere et al., 2012). For example, based on the Parental Bonding Instrument (PBI), adult males stated that they felt a lack of care from their fathers, while adult females felt a lack of autonomy from their fathers (Kullberg et al., 2020). A longitudinal study about social anxiety and parental support revealed adolescents with more social anxiety lacked autonomy and had higher levels of psychological control from their parents, while the mothers of those adolescents also stated similar conclusions (Nelemans et al., 2020). However, fathers were not a focus in the study and as such, there is a gap in

the understanding of the relationship between reports of one's childhood connections with their fathers and how it relates to their experiences with anxiety.

A variety of factors related to parenting have been studied in relation to anxiety. For example, a correlational study examined the association between overprotective parenting and child anxiety, finding that parental overprotection was related to a child's behavior, indicating it is not uniquely because of anxiety, but rather, from additional behavior issues such as child anxiety stemming from parental involvement (Gere et al., 2012). Similarly, a study investigated parental controlling behaviors with anxious and non-anxious children and found that with an increase in anxiety from both the child and parent, the more control parents had over their child (Edison et al., 2011). However, most research studied on these parental factors focused on the amount and/or type of parental support displayed in an adolescent's life and is related to how much anxiety they experience (Simpson et al., 2018; Sluis et al., 2015; Vulliez-Coady et al., 2013; Wang et al., 2019). Therefore, the current study will be adding to the extant literature by specifically exploring how young adults reflect on their childhood connections with their fathers and experiences with anxiety.

Parental involvement and support were associated with adolescents' development of anxiety. Adolescents who were hospitalized for psychiatric disorders and often separated from their parents were found to have issues with parental support, but particularly with their mothers. They perceived their mothers to not be supportive and weren't close to them, in addition to feeling their interactions were negative (Vulliez-Coady et al., 2013). However, this highlights a specific gap in the literature, which is the lack of research exploring the relationships of paternal support with various psychosocial outcomes, particularly among those who are experiencing symptoms or diagnoses of psychological disorders. Similarly, patterns in previous literature focused on maternal figures, while not addressing paternal support, stating social anxiety symptoms were associated with higher levels of maternal psychological control and lower autonomy (Kullberg et al., 2020; Nelemans et al., 2020; Vulliez-Coady et al., 2013). An alternative topic, paternal support, may play a role in adolescents' experience with anxiety. Although adolescents' maternal support is important to development, it is one factor among many that can influence a family unit's dynamic. Each member of a family unit adds something to it as a whole, so without looking at the role of the paternal figure, it may not provide all the information necessary to understand families as a system.

There is a lack of literature on the role of fathers in specific, therefore, the goal of this study was to analyze the relationship between paternal support and young adults' reports of anxiety from their childhood. This was explored by examining anxiety in young adults (ranging in age from 25 to 34) through Wave IV of the National Longitudinal Study of Adolescent to Adult Health (ADD Health) dataset (which specifically gathered responses made by young adults who have been contributing to this longitudinal data collection since they were adolescents; Harris et al., 2019). This study investigated the amount of paternal involvement and if support fosters growth to then minimize anxiety. Specifically, is there an association between closeness to one's father figure and having an anxiety or panic disorder and is this moderated by whether their father figure served time in jail or prison? This moderator was explored since previous research using the same dataset found that when a parent was incarcerated during one's childhood, it was associated with negative health outcomes in one's young adulthood (Lee et al., 2013). Is there a difference in how isolated a young adult feels based on the type of father figure they had in their life? Does this feeling of isolation vary based on how close a young adult feels towards their father figure? Is there an association between having a relationship with a father figure and symptoms of anxiety in young adults and is this moderated by how close they live to their mothers? Previous research has noted that transitions throughout the course of one's life and various individual characteristics are associated with the quality of the relationship between a parent and their child (Kaufman & Uhlenberg, 1998). As such, proximity to living to one's parent was studied as a moderator.

Method

Participants

The sample was taken from the fourth wave of ADD Health, the national longitudinal study that observed the development of health from adolescence to adult health, representing the population within the United States (Harris et al., 2019). This wave was composed of 15,701 participants, ranging from ages 25 to 34. The interview questions were administered on laptops and utilized the Blaise Survey Software, which is a computer-assisted interviewing (CAI) program. Most of these interviews were done in the respondent's homes, but there were additional interviews conducted in restaurants, coffee shops, libraries, and parents' or friends' homes (Harris et al., 2019).

Participants were included in the sample for the current study if they had a paternal figure, therefore excluding other participants' responses within the Add Health dataset. Therefore, the final sample included

4,401 participants, ranging from the ages of 25 to 34. Of the participants included, 85.8% had a biological father, 2.2% had an adoptive father, 1.4% had a step-father who adopted you, and 10.6% had a step-father.

Materials

Questions related to paternal support included, father figures having done jail time (“(Has/Did) your biological father ever (spent/spend) time in jail or prison?”), what father figure they have (“What is this person’s relationship to you?”), contact through electronic devices/mail (“How often do you and your (father figure) talk on the telephone, exchange letters, or exchange email?” with choices ranging from one a year or less to almost every day), satisfaction with communication (“You are satisfied with the way your (father figure) and you communicate with each other?” with choices ranging from strongly agree to strongly disagree), closeness (“How close to you feel to your (father figure)?” with choices ranging from not at all close to very close), frequency of interaction (“How often do you and your (father figure) see each other?” with choices ranging from never to almost every day), and proximity (“How far do you and your (father figure) live from one another?” with choices ranging from live together to more than 200 miles).

The items related to anxiety were applicable to those who may have an anxiety disorder, along with those who feel they experience common symptoms of anxiety, but were never formally diagnosed. Participants were asked if they had an anxiety disorder diagnosed (“Has a doctor, nurse, or other healthcare provider ever told you that you have had: anxiety or panic disorder?”). In addition, they were asked about the behaviors related to anxiety, including worrying (“I worry about things.”), communication (“I talk to a lot of different people at parties.”), worrying about past experiences (“I don’t worry about things that have already happened.”), avoiding conflict (I go out of my way to avoid having to deal with problems in my life.”), stress (“I get stressed out easily.”), expectations (“I hardly ever expect things to go my way.”), isolation (“I keep in the background.”), mood (I have frequent mood swings.”), control (“There is little I can do to change the important things in my life.”), and distrust (“I rarely count on good things happening to me.”). These items were answered on a 5-point scale, ranging from 1 (strongly agree) to 5 (strongly disagree).

Results

Close to Father Figure and Diagnosed with Anxiety

A Chi Square Test of Independence revealed that how close participants felt to their father figure (not at all

to somewhat close vs. quite close to very close) was significantly associated with whether a doctor, nurse, or other healthcare provider ever told the participants that they have/had an anxiety or panic disorder, $\chi^2 (1, N = 4,004) = 9.32, p = .002$. These results reveal a significant association between young adults who were not diagnosed with anxiety or panic reporting being closer to their father figure when compared to those who were diagnosed with either anxiety or panic.

Relationship with Father Figure and Symptoms of Anxiety

A Pearson Correlation indicated there was not a significant relationship between participants’ reports of how often they see their father figure and having frequent mood swings, $r(3516) = -.028, p = .102$.

A Pearson correlation indicated there was not a significant correlation between participants’ reports on how far they lived from their father figure and how much they worry about things, $r(4378) = -.016, p = .291$. This relationship was further explored if it was moderated by how far these young adults live from their mother. It was found that living with or close to their mother figure (within 50 miles), moderated the relationship between how far they lived from their dad and how much they worry about things, $r(2762) = -.037, p = .050$. The closer a young adult lived to their father figure, the less they worried about things, but only for those who also live with or close to their mother figure. However, this relationship among those who live with or close to their mother figure was weak and only 0.13% (i.e., $r^2 = .0013$) of the variance in worry about things was explained by how close they lived to their father figure. This relationship was not moderated if a young adult lived further from their mother figure, $r(1411) = .008, p = .761$.

Having a Father Figure and Feeling Isolated

Participants were split into four categories, based on what type of father figure they had in their life, including biological, adoptive, step-father who adopted you, and step-father. Then, a one-way between-subjects analysis of variance (ANOVA) was conducted to observe the difference between having different types of father figures on feeling isolated, $F(3, 4380) = 0.35, p = .79$. The difference in the type of father figures including biological fathers ($M = 3.30, SD = 0.97$), adoptive fathers ($M = 3.25, SD = 1.03$), step-fathers who adopted you ($M = 3.39, SD = 1.03$), and step-fathers ($M = 3.28, SD = 0.97$) on feelings of isolation was not significant. Based on the results, having a father figure does not play a significant role in feeling isolated.

Closeness to Father Figure and Feeling Isolated

A one-way between-subjects analysis of variance (ANOVA) was conducted to analyze the difference between how close a young adult feels towards their father figure and how isolated they were. Participants were split into two categories based on how close they feel to their father figure (not at all to somewhat close or quite close to very close). There was a significant difference in feelings of isolation based on how close they were with their father figure, $F(1, 4214) = 8.60, p = .003$. Young adults who were not at all close to somewhat close to their father figure ($M = 3.24, SD = .99$) reported feeling more isolated than those who were quite close to very close with their father figure ($M = 3.34, SD = .96$).

Further analyses revealed that this relationship was not moderated by having a father figure who has served time in prison or jail ($F(1, 95) = 2.67, p = .105$) or has not served any time ($F(1, 688) = .50, p = .480$).

Discussion

The purpose of this study was to explore the association between paternal support from father figures and young adult mental health. It was found that certain aspects of the relationship with a father figure was related to symptoms of anxiety in young adults. Young adults who were not diagnosed with anxiety or panic disorder seemed to also be very close to their father figure. Young adults who were not at all close or somewhat close to their father figure felt more isolated than those who were quite close or very close with their father figure. Further, if a young adult lived closer to their father figure, they worried less about things, but only for those who also live with or close to their mother figure.

While the current study elucidated the importance of a paternal figure and its relationship to displaying less symptoms of anxiety, most of the past research has focused on the importance of the maternal figure in this context (Kullberg et al., 2020; Nelemans et al. 2020; Vulliez-Coady et al., 2013). Further, general parental involvement has been explored, such that adolescents perceiving their parents being involved was associated with less mental health difficulties and suicidal thoughts and behaviors (Wang et al., 2019). As the current study demonstrated, it is critical to understand the role of one's father figure, in specific, and how it relates to symptoms of anxiety in young adults. Being close with one's father has positive associations with a young adult's anxiety. This is consistent with a study on British adolescents that found that involvement of their mothers and fathers independently were integral for their happiness (Flouri & Buchanan, 2003). However, it was also found that involvement of an adolescent's father had

a stronger effect in this relationship (Flouri & Buchanan, 2003), indicating the importance of further understanding the role of one's father figure independently in a U.S. context in the trajectory of one's life and how it relates to their levels of anxiety. Further, it has been found that the role of a father's is integral, yet different from that of a mother's, in how children are socialized, along with being able to help buffer against anxiety symptoms (Bögels & Phares, 2008). These results together relate to the current study in that paternal involvement does significantly relate to a decrease in struggles with anxiety among adolescents.

This study concluded that paternal involvement was related to minimizing some feelings of anxiety. Having a father figure involved in young adults' lives did reveal positive associations with mental health outcomes; therefore, contributing to the overall mental health of young adults. Both mother and father figures play a role in young adults' well-being and contributing to their growth and into their adult lives (Wang et al., 2019). However, the role of one's father can be integral to one's development. For example, among adolescent boys, having a father who was involved was related to greater life satisfaction, particularly among those who have been victimized by bullying (Flouri & Buchanan, 2002).

Limitations and Future Research

When using archival data, there are limitations as to what relationships can be studied based on being constrained to the data that was previously collected. As such, the scope of the variables explored in the current study were limited to how they were limited to how they were operationalized in the original dataset. Therefore, a more holistic understanding of the topic could be achieved by studying different aspects of the variables. Future studies should explore the relationship between paternal support and young adults' anxiety, based on the parenting style they were reared with and the attachment style that have developed. Since the importance of parenting and attachment styles have been well-established in the literature, it would be integral to understand how they moderate the relationships father figures have with the anxiety experienced by their children and as they grow into young adults.

Future studies could explore how these paternal relationships vary based on culture. Since the current study focused on father figures living in the United States, which is a predominantly individualistic culture, these results could change if investigated in a collectivistic culture because of differences in the role that father figures are expected to uphold. Further, future studies should explore how these parental relationships can vary based on having parents of the same gender and

how these role expectations, along with relationships with their children, can vary based on their cultural expectations, too.

In conclusion, parents have an important role as caregivers since they allow affordances to their children, such that without their support, the child's development may be inhibited, possibly increasing stressors in the future. From a clinician's perspective, the results of the current study could be used as background knowledge that not only maternal support, but paternal support is crucial for a young adult's development of anxiety. It is important that father figures are involved in the lives of young adults, regardless of their family situation. If there are conflicts within the family, it is recommended to work with clinicians and social workers on strengthening the relational bond between the father, child, and the family as a unit (Thomas, 2010). It is also recommended for parents and future parents to work through their own struggles with anxiety in hopes of preventing projection onto their current relationships/upcoming relationship with their children. The current study can also add to sociological literature, as it helps to understand the role of a father in the lives of their offspring and in the larger family unit. In addition, schools and the education system at large can try to develop programs to foster paternal engagement and teach adaptive parenting skills to improve a child's experiences with anxiety.

References

- American Psychological Association. (n.d.). *Anxiety*. <https://www.apa.org/topics/anxiety#:~:text=Anxiety%20is%20an%20emotion%20characterized,they%20are%20often%20used%20interchangeably>.
- Berger, K. S. (2018). *The developing person through childhood and adolescence* (11th edition). Worth Publishers.
- Bögels, S., & Phares, V. (2008). Father's role in the etiology, prevention and treatment of child anxiety: A review and new model. *Clinical Psychology Review*, 28, 539-558. <https://doi.org/10.1016/j.cpr.2007.07.011>
- Edison, S. C., Evans, M. A., McHolm, A. E., Cunningham, C. E., Nowakowski, M. E., Boyle, M., & Schmidt, L. A. (2011). An investigation of control among parents of selectively mute, anxious, and non-anxious children. *Child Psychiatry and Human Development*, 42(3), 270-290. <https://doi.org/10.1007/s10578-010-0214-1>
- Flouri, E., & Buchanan, A. (2002). Life satisfaction in teenage boys: The moderating role of father involvement and bullying. *Aggressive Behavior*, 28, 126-133. <https://doi.org/10.1002/ab.90014>
- Flouri, E., & Buchanan, A. (2003). The role of father involvement and mother involvement in adolescents' psychological well-being. *British Journal of Social Work*, 33, 399-406. <https://doi.org/10.1093/bjsw/33.3.399>
- Gere, M. K., Villabø, M. A., Torgersen, S., & Kendall, P. C. (2012). Overprotective parenting and child anxiety: The role of co-occurring child behavior problems. *Journal of Anxiety Disorders*, 26(6), 642-649. <https://doi.org/10.1016/j.janxdis.2012.04.003>
- Harris, K.M., Halpern, C.T., Whitsel, E.A., Hussey, J.M., Killea-Jones, L., Tabor, J., & Dean, S.C. (2019). Cohort profile: The National Longitudinal Study of Adolescent to Adult Health (Add Health). *International Journal of Epidemiology*, 48, 1415-1425. <https://doi.org/10.1093/ije/dyz115>
- Kaufman, G., & Uhlenberg, P. (1998). Effects of life course transition on the quality of relationships between adult children and their parents. *Journal of Marriage and Family*, 60, 924-938. <https://doi.org/10.2307/353635>
- Kullberg, M.-L., Maciejewski, D., van Schie, C. C., Penninx, B. W. J. H., & Elzinga, B. M. (2020). Parental bonding: Psychometric properties and association with lifetime depression and anxiety disorders. *Psychological Assessment*, 32(8), 780-795. <https://doi.org/10.1037/pas0000864>
- Lee, R.D., Fang, X., & Luo, F. (2013). The impact of parental incarceration on the physical and mental health of young adults. *Pediatrics*, 131, e1188-e1195. <https://doi.org/10.1542/peds.2012-0627>
- Nelemans, S. A., Keijsers, L., Colpin, H., Leeuwen, K., Bijttebier, P., Verschueren, K., & Goossens, L. (2020). Transactional links between social anxiety symptoms and parenting across adolescence: Between-and within-person associations. *Child Development*, 91(3), 814-828. <https://doi.org/10.1111/cdev.13236>
- National Institute of Mental Health. (n.d.). *Any anxiety disorder*. <https://www.nimh.nih.gov/health/statistics/any-anxiety-disorder.shtml#:~:text=Prevalence%20of%20Any%20Anxiety%20Disorder%20Among%20Adolescents,-Based%20on%20diagnostic&text=An%20estimated%2031.9%25%20of%20adolescents,estimated%208.3%25%20had%20severe%20impairment>
- Simpson, D., Suarez, L., Cox, L., & Connolly, S. (2018). The role of coping strategies in understanding the relationship between parental support and psychological outcomes in anxious youth. *Child*

- and *Adolescent Social Work Journal*, 35(4), 407-421.
<https://doi.org/10.1007/s10560-018-0531-y>
- Sluis, C. M., Steensel, F. J. A., & Bögels, S. M. (2015). Parenting and children's internalizing symptoms: How important are parents? *Journal of Child and Family Studies*, 24(12), 3652–3661.
<https://doi.org/10.1007/s10826-015-0174-y>
- Substance Abuse and Mental Health Services Administration. (2016, June). *Impact of the DSM-IV to DSM-V changes on the national survey on drug use and health*. National center of biotechnology information.
<https://www.ncbi.nlm.nih.gov/books/NBK519704/table/ch3.t15/>
- Thomas, L.K. (2010). Relational psychotherapy: The significance of father. *Psychodynamic Practice: Individuals, Groups and Organisations*, 16, 61-75.
<https://doi.org/10.1080/14753630903474579>
- Vulliez-Coady, L., Vidal, C., Vivot, A., Monnin, J., Nezelof, S., & Bifulco, A. (2013). Anxiety and hospitalization in adolescence: Relations to attachment style and parental support. *Psychiatria Danubina*, 25, 135-142.
https://www.psychiatria-danubina.com/UserDocsImages/pdf/dnb_vol25%20Suppl%202_no/dnb_vol25_noSuppl%202_13_5.pdf
- Wang, C., La Salle, T. P., Do, K. A., Wu, C., & Sullivan, K. E. (2019). Does parental involvement matter for students' mental health in middle school. *School Psychology*, 34(2), 222-232.
<https://doi.org/10.1037/spq0000300>

SOCIAL SUPPORT, AGE, AND EMOTION REGULATION PREDICT SYMPTOM SEVERITY OF ADULTS WITH MAJOR DEPRESSIVE DISORDER

ALYSSA M. NELSON¹, JOSEPH M. DIEHL¹, KEVIN S. LABAR¹, AND MORIA J. SMOSKI^{1,2}

¹DUKE UNIVERSITY, ²DUKE UNIVERSITY MEDICAL CENTER

Abstract – The aim of this study was to examine the relationships among social support, age, emotion regulation, and depressive symptoms in adults with Major Depressive Disorder (MDD). A better understanding of remediable factors that are related to MDD outcomes could inform effective treatments and preventative strategies. A total of 42 participants aged 35-68 responded to survey items assessing social support, distraction use, reappraisal use, and depressive symptom severity. Depression severity was more strongly associated with lower social support for older participants. Lower emotion regulation strategy use was associated with greater depressive symptoms, a relationship that was accounted for by lower social support. Social support is a key contributor to the positive impact of emotion regulation ability on mitigating depression symptoms in mid-life to older adults with MDD. Emotion regulation-based interventions may benefit from a focus on enhancing social support and positive social interactions. Future research should utilize experimental manipulations and longitudinal data analysis to establish causality in the observed relationships.

Keywords: adulthood, emotion regulation, Major Depressive Disorder, social support

Major Depressive Disorder (MDD) is a pervasive mental health condition with significant adverse consequences across the adult lifespan. About 20% of American adults experience a depressive episode in their lives, rendering MDD one of the most common mental health conditions in the United States (Hasin et al., 2018). In addition, MDD has significant detrimental consequences for those affected as well as society at large (Charlson et al., 2013; Lawrence et al., 2022; Russo et al., 2015; Pérez-Sola et al., 2021). The negative ramifications of depression generalize beyond the realms of emotional well-being and daily functioning. A robust literature has demonstrated that depression has a substantial negative impact on somatic health, increasing risk of morbidity and mortality (Penninx et al., 2013; Chiu et al., 2018).

Effective interventions for depression are vital due to its prevalence and devastating consequences. Although prior research supports the effectiveness of various treatments for MDD (Cuijpers et al., 2023), available therapies have limitations with respect to initial treatment response (Lemmens et al., 2011) and relapse (Paykel, 2008). A better understanding of factors that are related to depression outcomes could inform effective treatments and preventative strategies for MDD. Positive

relationships and connections with others, as well as the ability to intentionally regulate one's emotions, are known to mitigate depression severity (Leskela et al., 2006; Campbell-Sills et al., 2006; Daros et al., 2023). However, the relationship between social support and emotion regulation, especially in older adults, is less clear.

A robust literature confirms that social support is linked to positive clinical outcomes for those with MDD. Subjective perceptions of strong social support, as well as frequent social interactions, are associated with a lower risk of developing a recurrent episode of MDD and reductions in depressive symptoms over time (Sherbourne et al., 1995; Leskela et al., 2006). Research has also found social support to be associated with favorable treatment outcomes for those undergoing therapy for MDD (Solomonov et al., 2019). Unfortunately, the size of social networks tends to diminish as individuals progress through adulthood (Schnittker, 2007). This reduction occurs as social support may be more salient, with one study demonstrating that the protective effects of social support for onset of MDD and the clinical course of the disease were most pronounced for older patients (Sherbourne et

al., 1995). It is unclear if other age-related changes in cognition and emotion may impact the relationships among social support, age, and depression.

Emotion regulation is another factor that has important implications for MDD. Impaired emotion regulation is a characteristic feature of those who have MDD (Joormann & Stanton, 2016). Specifically, individuals with MDD seem to have greater difficulty regulating negative affect after adverse life experiences, and this has been found to be implicated in the onset of many individuals' first depressive episode (Hammen, 2005). Additionally, those with MDD report less frequent use of effective emotion regulation strategies such as reappraisal (thinking differently about a negative stimulus to alter emotional response) (Joormann & Stanton, 2016). Clinical populations who are struggling with depression also derive less benefit from some emotion regulation strategies (Smoski et al., 2014) and are often unable to return to their emotional baseline as quickly as non-clinical populations (Joormann, 2010). This may be due in part to the fact that those with MDD have increased ruminative thinking, meaning their thoughts tend to recur, or "recycle" themselves (Flett et al., 2002; Spasojevic & Alloy, 2001). This pattern of thinking is particularly detrimental in the context of depressed moods, as rumination enhances negative thoughts, interferes with instrumental behavior that could improve mood, and increases the likelihood of social rejection during depressive episodes (Nolen-Hoeksema et al., 2008).

To date, few studies have explored the relationship between social support and emotion regulation in MDD. One theoretical link between the constructs of social support and emotion regulation is Social Baseline Theory, which posits that violations of an expectancy of social support tax cognitive and affective resources (Coan & Sbarra, 2015), such that fewer resources are available for emotion regulation in the context of low social support. Conversely, strong social support seems to facilitate attempts to regulate emotion. Morawetz et al. (2021) demonstrated that reappraisal strategies in response to disturbing images are enhanced with the support of a friend compared to the support of a stranger or when alone. Thus, the inverse relationship between social support and depression may be partially explained by emotion regulation.

The aim of the present study was to more comprehensively examine the relationships among social support, age, emotion regulation, and depressive symptoms in middle aged to older adults with MDD. We hypothesized that the negative association between social support and depressive symptoms would get stronger as

age increased in adulthood. In addition, we hypothesized that effective emotion regulation strategy use would at least partially account for the relationship between social support and depression.

Methods

This study is a secondary analysis of an observational study examining the neurobehavioral mechanisms of emotion regulation across the adult lifespan [NCT03207503]. The parent study protocol was approved by the Duke Health IRB. Some participants included in the present study's analysis did not complete the entire protocol for the parent study. Their data was included if they met screening, clinical, and neuropsychosocial requirements.

Participants

Participants included 42 adults diagnosed with MDD. The mean age in the sample was 49.67 ($SD = 10.71$, range = 35-68). The majority of participants identified as female (86%) and white (79%). Black or African American participants made up 12% of the sample, with 5% who identified as Hispanic/Latinx. Inclusion criteria included the following: not currently pregnant, ambulatory with no known uncorrected sensory deficits, an estimated verbal IQ of at least 88, and a Beck Depression Inventory-2 (BDI-2; Beck et al., 1996) score of at least 14 with current MDD or Persistent Depressive Disorder (PDD) with either persistent major depressive episode or intermittent major depressive episodes. Exclusionary criteria included current substance use disorder; current or history of psychosis, mania, or eating disorder; disorder with impact on brain characteristics; contraindications to MRI scanning; unstable medicine regimen of antidepressants, antipsychotics, or mood stabilizers; and indication of mild cognitive impairment or dementia. Participants were recruited through online advertisements, letters from their primary care physicians, and research recruitment databases, including the Duke Brain Imaging and Analysis Center control subject registry, ResearchMatch, and the Duke Cognitive-Behavioral Research Treatment Program participant registry.

Procedure

The current study is a secondary data analysis in which existing data from a parent study was analyzed to address a novel research question. The parent study of the current project was an observational study that incorporated interview, behavioral, and neuroimaging measures of depression as well as survey measures of social support and emotion regulation. Data for the present study was drawn from the aforementioned survey responses. Participants completed surveys online using

REDCap (Research Electronic Data Capture), an online application that collects and stores data for clinical research projects (Patridge & Bardyn, 2018). They also completed surveys in-person at Duke University. Due to COVID-19 restrictions, some participants also completed components of the study remotely over the phone or online over Zoom or WebEx.

Participants completed a pre-screen over the phone or through REDCap (based on participant preference) to determine potential study eligibility. Those who were eligible were directed to an online consent form hosted by REDCap. Once the participants had given informed consent, they were scheduled for up to two assessment visits to further determine study eligibility. Assessments included clinical interview and a brief neuropsychosocial screener. Eligible participants were then asked to respond to survey questions administered through REDCap that assessed demographics, social support, and emotion regulation. Two more study visits, which involved data that will not be analyzed in this study, were scheduled after this.

Measures

MDD diagnosis was confirmed via the *Structured Clinical Interview for DSM-5–Research Version* (SCID-5-RV; First et al., 2015), a structured interview to assess lifetime psychiatric disorders and to document duration, onset, and number of depressive episodes. Depression severity was indexed by the *Beck Depression Inventory-II* (BDI-II; Beck et al., 1996), a 21 item, self-report instrument that measures severity of depressive symptoms. Cronbach's α for this measure was .97, suggesting acceptable internal reliability. A total composite score of the *Duke Social Support Index* (DSSI; Koenig et al., 1993) indexed social support. The DSSI is a 23 item, self-report instrument that measures social networks, social interactions, subjective support, and instrumental support. Cronbach's α for this measure was .89, suggesting acceptable internal reliability. Emotion regulation was measured as a summed standard score of distraction use (the 11 distraction items of the *Response Styles Questionnaire* (RSQ; Nolen-Hoeksema & Morrow, 1991)) and reappraisal use (the 6 reappraisal items of the *Emotion Regulation Questionnaire* (ERQ; Gross & John, 2003)). Cronbach's α for this measure was .89, suggesting acceptable internal reliability. Finally, participants provided information on gender, age, race, ethnicity, education, city of birth, and marital status.

Analyses

A summary emotion regulation variable was created by summing the Z-scores of the RSQ-distraction and ERQ-reappraisal scores. Bivariate correlations were conducted to establish associations between social

support scores and depression scores and between social support scores and age. Moderation analysis was conducted using PROCESS in SPSS and reverified using the ggplot2 package in R to test whether the association between depression scores and social support became stronger as age increased. Next, bivariate correlations were conducted to establish associations between emotion regulation scores and social support scores and between emotion regulation scores and depression scores. Partial correlations were then conducted between depression scores and social support scores while accounting for summary emotion regulation scores. Finally, partial correlations were conducted between depression scores and emotion regulation scores while accounting for social support scores.

Results

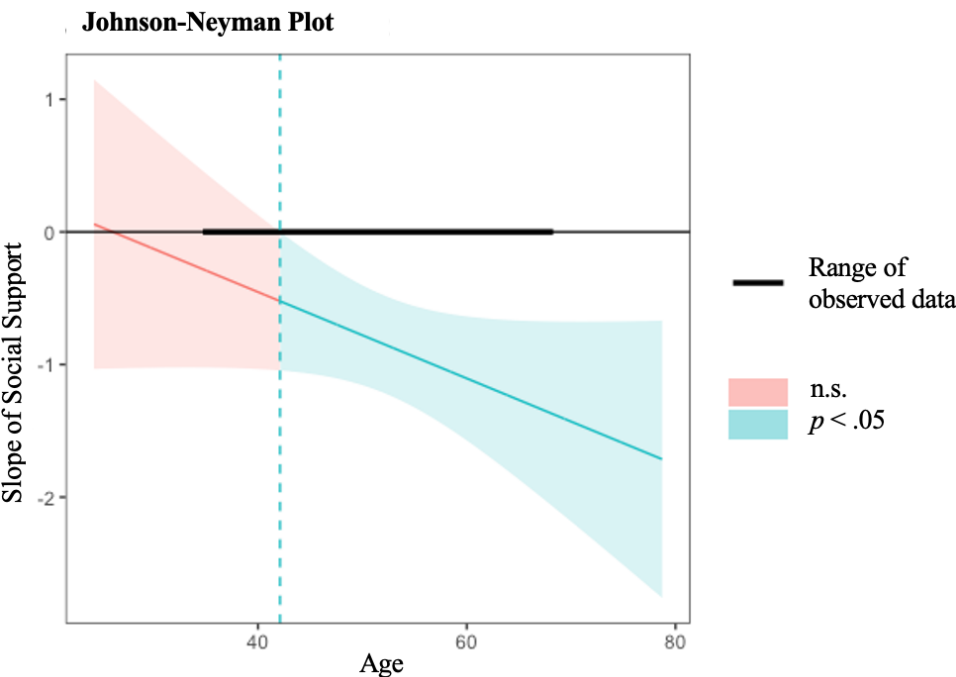
Associations between social support and depression ($r = -.57, p < .001$) and between social support and age ($r = -.35, p = .02$) were statistically significant. A multiple regression analysis with social support, age, and their interaction term as predictors of depression severity was computed. The overall model was a good fit, $F(3, 42) = 8.46, p < .001, R^2 = .40$. Higher social support was a significant predictor of lower depression ($b = 0.76, p < .01$) but age was not ($b = 0.16, p = 0.15$). The interaction term trended toward significance ($b = -0.33, p = .08$), and thus, the moderation analysis was probed using the Johnson-Neyman procedure. Beginning at age 42, the effect of age on the association between depression and social support was significant ($b = -.52, p < .05$). As age increased from this point, the negative association between social support and depression grew stronger until it reached the highest age in the sample ($b = -1.37, p < .001$). A Johnson-Neyman plot of these results is presented in Figure 1. A conceptual model of age as a moderator of the association between social support and depression is depicted in Figure 2.

Associations between emotion regulation and social support ($r = .38, p = .01$) and between emotion regulation and depression ($r = -.35, p = .02$) were statistically significant. Contrary to our hypothesis, when accounting for emotion regulation, the inverse association between social support and depression severity remained significant ($r = -0.51, p < .001$). However, when accounting for social support, the inverse association between emotion regulation and depression became nonsignificant ($r = -.173, p = .278$).

Discussion

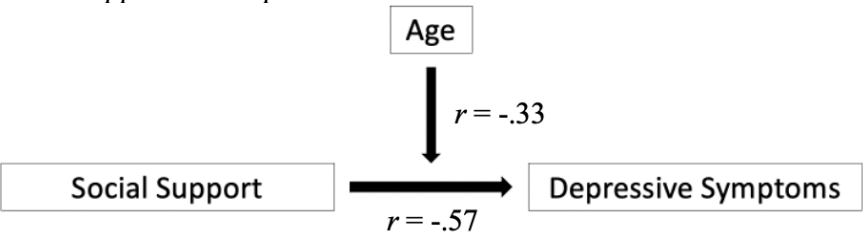
The objective of this study was to examine the relationships among social support, age, emotion regulation, and depressive symptoms in adults with

Figure 1
Johnson-Neyman Plot of Age as a Moderator of the Association Between Social Support and Depression



Note. Beginning at age 42 in our sample, the effect of age on the association between depression and social support was significant. As age increased from this point, the negative association between social support and depression grew stronger until it reached the highest age (68) in the sample.

Figure 2
Conceptual Model of Age as a Moderator of the Association Between Social Support and Depression



MDD. The results suggest that social support and age are negatively associated and that age moderates the negative association between social support and depression. Additionally, findings indicated that social support and use of effective emotion regulation strategies are positively associated. Furthermore, the present study discovered that the negative association between emotion regulation and depression is accounted for by the positive relationship between emotion regulation and social support.

Consistent with hypotheses, greater social support was more strongly related to diminished

depressive symptoms for the oldest adults in this sample. These findings are in line with earlier work (Sherbourne et al., 1995), which demonstrated that social support elicited the greatest protective effects for the onset and course of MDD for older adults. It is noteworthy to mention that the present findings also align with prior research indicating an inverse relationship between social support and age. As social networks shrink, the risk of experiencing depression tends to increase (Gianfredi et al., 2021). This underscores the heightened necessity of cultivating resilient social connections and support systems to enhance the mental well-being of older adults.

This study also provides evidence that social support and emotion regulation are positively related, consistent with prior research (Cooke et al., 2019; Morawetz et al., 2021). Specifically, greater social engagement and satisfaction with support received were associated with greater use of distraction and reappraisal

strategies. One potential explanation of these findings is that engaging in social interactions and feeling supported by one's social network may provide opportunities to leverage strategies that diminish negative affect. This is consistent with recent research that indicated leveraging social resources is a promising avenue for regulating emotion in MDD (Liu et al., 2024). Alternatively, leveraging strategies to regulate negative emotions may make it easier to form and maintain relationships. A final possibility is that a third variable not measured in the current study could be related to both social support and

emotion regulation, explaining their positive association. For instance, a factor of psychological health may enable individuals to establish and maintain social relationships as well as engage in strategies that diminish their negative emotions.

Concerning depression severity, the present findings indicate a negative association with use of effective emotion regulation strategies. However, contrary to expectations, better emotion regulation did not account for the negative association between social support and depression. Instead, the association between emotion regulation and depression became nonsignificant after accounting for social support. This pattern suggests that leveraging distraction and reappraisal diminishes depressive symptoms by enhancing satisfying social engagement. These results are consistent with Coyne's (1976) empirically supported interpersonal theory (Marcus & Nardone, 1992), which posits that individuals with persistent negative affect are eventually distanced from family and friends, leading to the loss of social support and the onset of depressive episodes.

The present findings have clinical implications, as they provide insight regarding interventions for depression. Consistent with prior research (Sherbourne et al., 1995; Leskela et al., 2006), having a strong social network and feeling supported by this network was associated with fewer depressive symptoms for those with MDD, particularly in older age. This may be related to the idea that social support "buffers" against the impacts of stressors and strengthens the ability to bounce back from stressful life events that might otherwise lead to depression (Bovard, 1959). In addition, the present findings indicate that without strong social support, emotion regulation may not be a protective factor for depressive symptom severity in older depressed adults. Emotion regulation-based interventions may benefit from a focus on enhancing social support and positive social interactions, as demonstrated in the ENGAGE intervention for late-life MDD (Solomonov et al., 2019).

While the present study has notable strengths, such as implementation of rigorous exclusion criteria and control measures during enrollment, it is important to note methodological limitations. The lack of experimental manipulations and longitudinal data analysis prevents establishing causality in the observed relationships. Additionally, the results have limited generalizability to those with marginalized identities, as the sample was predominantly white. Finally, the statistical power of the study was limited due to the small sample size. Future research should address these limitations by increasing the sample size and the diversity of their samples with

respect to age, race, ethnicity, and gender. Also, longitudinal and/or experimental methodology would provide greater clarity as to the directionality of the relationships observed among the variables in this study.

Despite the aforementioned limitations, this study contributes to a greater understanding of the intricate relationships among social support, age, emotion regulation, and depression. Most notably, the study emphasizes that social support is a key contributor to the positive impact of emotion regulation ability on mitigating depression symptoms in mid-life to older adults with MDD.

The findings provide future directions for a more nuanced clinical approach to MDD across the adult lifespan. These results underscore the importance of developing therapeutic interventions for MDD that consider age-related factors to address the specific needs of different age groups. For older adults in particular, therapeutic programs for MDD should prioritize bolstering social support through strengthening of existing relationships and fostering new sources of social connection. This approach can enhance emotion regulation abilities and improve mental health outcomes in this demographic.

References

- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck depression inventory-II*. Psychological Corporation.
https://doi.org/10.1007/978-1-4419-1005-9_441
- Bovard E. W. (1959). The effects of social stimuli on the response to stress. *Psychological Review*, 66(5), 267–277. <https://doi.org/10.1037/h0045556>
- Campbell-Sills, L., Barlow, D., Brown, T., & Hofmann, S. (2006). Acceptability and suppression of negative emotion in anxiety and mood disorders. *Emotion*, 6(4), 587–595.
<https://doi.org/10.1037/1528-3542.6.4.587>
- Charlson, F. J., Ferrari, A. J., Flaxman, A. D., & Whiteford, H. A. (2013). The epidemiological modelling of dysthymia: Application for the global burden of disease 2010 study. *Journal of Affective Disorders*, 151(1), 111–120.
<https://doi.org/10.1016/j.jad.2013.05.060>
- Chiu, M., Vigod, S., Rahman, F., Wilton, A. S., Lebenbaum, M., & Kurdyak, P. (2018). Mortality risk associated with psychological distress and major depression: A population-based cohort study. *Journal of affective disorders*, 234, 117–123. <https://doi.org/10.1016/j.jad.2018.02.075>
- Coan, J. A., & Sbarra, D. A. (2015). Social Baseline Theory: The Social Regulation of Risk and Effort. *Current opinion in psychology*, 1, 87–91.
<https://doi.org/10.1016/j.copsyc.2014.12.021>

- Cooke, J. E., Kochendorfer, L. B., Stuart-Parrigon, K. L., Koehn, A. J., & Kerns, K. A. (2019). Parent-child attachment and children's experience and regulation of emotion: A meta-analytic review. *Emotion, 19*(6), 1103–1126. <https://doi.org/10.1037/em00000504>
- Coyne, J. C. (1976). Depression and the response of others. *Journal of Abnormal Psychology, 85*(2), 186–193. <https://doi.org/10.1037/0021-843X.85.2.186>
- Cuijpers, P., Harter, M., Miguel, C., Ciharova, M., & Karyotaki, E. (2023). Five decades of research on psychological treatments of depression: A historical and meta-analytic overview. *The American psychologist, 10.1037/amp0001250*. Advance online publication. <https://doi.org/10.1037/amp0001250>
- Daros, A. R., Wardell, J. D., & Quilty, L. C. (2023). Multilevel associations of emotion regulation strategy use during psychotherapy for depression: A longitudinal study. *Journal of affective disorders, 338*, 107–118. <https://doi.org/10.1016/j.jad.2023.06.014>
- First, M. B., Williams, J. B., Karg, R. S., & Spitzer, R. L. (2015). Structured clinical interview for DSM-5—Research version (SCID-5 for DSM-5, research version; SCID-5-RV). Arlington, VA: American Psychiatric Association, (1–94).
- Flett, G. L., Madorsky, D., Hewitt, P. L., & Heisel, M. J. (2002). Perfectionism cognitions, rumination, and psychological distress. *Journal of Rational-Emotive and Cognitive Behavior Therapy, 20*(1), 33–47. <https://doi.org/10.1023/A:1015128904007>
- Gianfredi, V., Beran, M., Koster, A., Eussen, S. J., Odone, A., Signorelli, C., Schaper, N. C., Köhler, S., Bosma, H., Dagnelie, P. C., Stehouwer, C. D., Wesselius, A., Amerio, A., Brinkhues, S., Dukers-Muijers, N., & Schram, M. T. (2021). Association between social network characteristics and prevalent and incident depression: The Maastricht Study. *Journal of affective disorders, 293*, 338–346. <https://doi.org/10.1016/j.jad.2021.06.046>
- Gross, J.J., & John, O.P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships, and well-being. *Journal of Personality and Social Psychology, 85*, 348–362. <https://doi.org/10.1037/0022-3514.85.2.348>
- Hammen C. (2005). Stress and depression. *Annual Review of Clinical Psychology, 1*, 293–319. <https://doi.org/10.1146/annurev.clinpsy.1.102803.143938>
- Hasin, D. S., Sarvet, A. L., Meyers, J. L., Saha, T. D., Ruan, W. J., Stohl, M., & Grant, B. F. (2018). Epidemiology of adult DSM-5 Major Depressive Disorder and its specifiers in the United States. *JAMA Psychiatry, 75*(4), 336–346. <https://doi.org/10.1001/jamapsychiatry.2017.4602>
- Joormann, J. (2010). Cognitive inhibition and emotion regulation in depression. *Current Directions in Psychological Science, 19*(3), 161–166. <https://doi.org/10.1177/0963721410370293>
- Joormann, J., & Stanton, C. H. (2016). Examining emotion regulation in depression: A review and future directions. *Behaviour research and therapy, 86*, 35–49. <https://doi.org/10.1016/j.brat.2016.07.007>
- Koenig, H. G., Westlund, R. E., George, L. K., Hughes, D. C., Blazer, D. G., & Hybels, C. (1993). Abbreviating the Duke Social Support Index for use in chronically ill elderly individuals. *Psychosomatics, 34*(1), 61–69. [https://doi.org/10.1016/S0033-3182\(93\)71928-3](https://doi.org/10.1016/S0033-3182(93)71928-3)
- Lawrence, T. I., Mcfield, A. A., Byrne, M. M., S Tarver, S., & Stewart, T. K. (2022). Depression and Substance Use as Consequences of Exposure to Family Violence: A Moderation Mediation and Self-Medication Hypothesis Study. *Journal of child & adolescent trauma, 16*(1), 69–79. <https://doi.org/10.1007/s40653-022-00464-3>
- Lemmens, L. H., Arntz, A., Peeters, F. P., Hollon, S. D., Roefs, A., & Huibers, M. J. (2011). Effectiveness, relapse prevention and mechanisms of change of cognitive therapy vs. interpersonal therapy for depression: Study protocol for a randomised controlled trial. *Trials, 12*, 150. <https://doi.org/10.1186/1745-6215-12-150>
- Leskela, U., Ryttsala, H., Komulainen, E., Melartin, T., Sokero, P., Lestela-Mielonen, P., & Isometsa, E. (2006). The influence of adversity and perceived social support on the outcome of major depressive disorder in subjects with different levels of depressive symptoms. *Psychological Medicine, 36*(6), 779–788. <https://doi.org/10.1017/S0033291706007276>
- Liu, D. Y., Strube, M. J., & Thompson, R. J. (2024). Do emotion regulation difficulties in depression extend to social context? Everyday interpersonal emotion regulation in current and remitted major depressive disorder. *Journal of psychopathology and clinical science, 133*(1), 61–75. <https://doi.org/10.1037/abn0000877>
- Marcus, D. K., & Nardone, M. E. (1992). Depression and interpersonal rejection. *Clinical Psychology Review, 12*(4), 433–449. [https://doi.org/10.1016/0272-7358\(92\)90126-S](https://doi.org/10.1016/0272-7358(92)90126-S)
- Morawetz, C., Berboth, S., & Bode, S. (2021). With a little help from my friends: The effect of social proximity on emotion regulation-related brain activity. *NeuroImage, 230*, 117817.

- <https://doi.org/10.1016/j.neuroimage.2021.117817>
- Nolen-Hoeksema, S., & Morrow, J. (1991). A prospective study of depression and posttraumatic stress symptoms after a natural disaster: The 1989 Loma Prieta Earthquake. *Journal of Personality and Social Psychology*, 61(1), 115–121. <https://doi.org/10.1037//0022-3514.61.1.115>
- Nolen-Hoeksema, S., Wisco, B. E., & Lyubomirsky, S. (2008). Rethinking rumination. *Perspectives on Psychological Science*, 3(5), 400–424. <https://doi.org/10.1111%2Fj.1745-6924.2008.00088.x>
- Patridge, E. F., & Bardyn, T. P. (2018). Research Electronic Data Capture (REDCap). *Journal of the Medical Library Association : JMLA*, 106(1), 142–144. <https://doi.org/10.5195/jmla.2018.319>
- Paykel, E. S. (2008). Partial remission, residual symptoms, and relapse in depression. *Dialogues in Clinical Neuroscience*, 10(4), 431–437. <https://doi.org/10.31887/DCNS.2008.10.4/espaykel>
- Penninx, B. W., Milaneschi, Y., Lamers, F., & Vogelzangs, N. (2013). Understanding the somatic consequences of depression: Biological mechanisms and the role of depression symptom profile. *BMC Medicine*, 11, 129. <https://doi.org/10.1186/1741-7015-11-129>
- Pérez-Sola, V., Roca, M., Alonso, J., Gabilondo, A., Hernando, T., Sicras-Mainar, A., Sicras-Navarro, A., Herrera, B., & Vieta, E. (2021). Economic impact of treatment-resistant depression: A retrospective observational study. *Journal of affective disorders*, 295, 578–586. <https://doi.org/10.1016/j.jad.2021.08.036>
- Russo, M., Mahon, K., & Burdick, K. E. (2015). Measuring cognitive function in MDD: Emerging assessment tools. *Depression and Anxiety*, 32(4), 262–269. <https://doi.org/10.1002/da.22297>
- Sherbourne, C. D., Hays, R. D., & Wells, K. B. (1995). Personal and psychosocial risk factors for physical and mental health outcomes and course of depression among depressed patients. *Journal of Consulting and Clinical Psychology*, 63(3), 345–355. <https://doi.org/10.1037//0022-006x.63.3.345>
- Schnittker J. (2007). Look (closely) at all the lonely people: Age and the social psychology of social support. *Journal of Aging and Health*, 19(4), 659–682. <https://doi.org/10.1177/0898264307301178>
- Smoski, M. J., LaBar, K. S., & Steffens, D. C. (2014). Relative effectiveness of reappraisal and distraction in regulating emotion in late-life depression. *The American Journal of Geriatric Psychiatry: Official Journal of the American Association for Geriatric Psychiatry*, 22(9), 898–907. <https://doi.org/10.1016/j.jagp.2013.01.070>
- Solomonov, N., Bress, J. N., Sirey, J. A., Gunning, F. M., Flückiger, C., Raue, P. J., Arean, P. A., & Alexopoulos, G. S. (2019). Engagement in socially and interpersonally rewarding activities as a predictor of outcome in "engage" behavioral activation therapy for late-life depression. *The American Journal of Geriatric Psychiatry*, 27(6), 571–578. <https://doi.org/10.1016/j.jagp.2018.12.033>
- Spasojevic, J., & Alloy, L. B. (2001). Rumination as a common mechanism relating depressive risk factors to depression. *Emotion*, 1(1), 25–37. <https://doi.org/10.1037/1528-3542.1.1.25>

Author Note

The authors report no conflicts with any product mentioned or concept discussed in this article. Research reported in this publication was supported by the National Institute of Mental Health of the National Institutes of Health under award number R01 MH113238. Data are available through NIMH Data Archive #2663, “Neurobehavioral Mechanisms of Emotion Regulation in Depression across the Adult Lifespan.”

Correspondence concerning this article should be addressed to Alyssa M. Nelson, Duke University, 113 Edens Drive, Duke University Box # 99217, Durham NC 27708. Email: Amn50@duke.edu

“I SUPPORT THEM, BUT...”: PERCEPTIONS OF TRANSGENDER ATHLETES DEPEND ON ATHLETE GENDER AND SKILL LEVEL

MADDEN G. RITZER, MERRY J. SLEIGH, AND DONNA W. NELSON

WINTHROP UNIVERSITY

Abstract – We investigated perceptions of transgender athletes using an experimental design that manipulated descriptions of athletes’ gender and ability level. We also examined whether characteristics of the perceiver would influence attitudes toward transgender athletes. Participants were provided with a narrative describing either a transgender male or female athlete who played on a college basketball team at either a low or high skill level. Participants provided their perceptions and responded to scales to assess their empathy, belief in a just world, and demographics. Participants expressed general support for transgender athletes and an awareness that these athletes face prejudice and discrimination. Matching our hypothesis, the transgender man with a high level of ability elicited the most support, while the transgender woman with a high level of ability elicited the lowest levels of support. Qualitative data revealed more mixed feelings than the quantitative data, with a focus on biological advantages that accompany male puberty. Participants who believed in a just world, were politically conservative, were lower in empathy, and older reported less support for transgender athletes than their counterparts. These findings suggest that perceptions of transgender athletes are based on the athletes’ gender identity and performance, as well as characteristics of the perceiver.

The inclusion of transgender athletes in competitive sports is increasing in frequency and has resulted in social controversy (Reynolds & Hamidian, 2021; Safer, 2022). In 2020, the first legislation that banned transgender girls from school-sponsored female sports teams was passed in the United States (Martin & Rahilly, 2023). By 2023, 21 states established laws that prevent transgender students from participating in school sports that are incongruent with their sex at birth (Jakubowska, 2024). Attitudes surrounding the issue are complex and influential in driving policies; however, research on the topic is limited (Avalos et al., 2022; Kavoura & Kokkonen, 2021).

A common argument against transgender athletes competing on teams that match their gender identity is that men have a biological advantage over women due to their physiology and higher testosterone levels (Martin & Rahilly, 2023; Pigozzi et al., 2022). This gender disparity becomes more pronounced following puberty, as many of the changes resulting from male hormones persist over time (Hilton & Lundberg, 2021; Jakubowska, 2024). For example, following puberty, a typical male body will have greater cardiorespiratory capacity and muscle mass than a typical female body

(Heather, 2022). Consequently, policies often factor in the developmental stage when the public gender transition takes place, the length of time an individual has undergone hormone replacement, or the length of time an individual has undergone suppression treatment (Jakubowska, 2024). As one example, the National Collegiate Athletic Association currently requires transgender athletes to provide documentation of testosterone levels prior to and during their sports season.

Because the male body is perceived to have an advantage over the female body, the extension of this argument is that banning transgender women, who have gone through male puberty, physically protects cisgender women and promotes fairness in competition (Martin & Rahilly, 2023). This argument relies heavily on the dichotomy reflected in sex-based chromosomes (Jakubowska, 2024), as well as performance records that consistently show superior performance by male athletes compared to female athletes in identical Olympic events (Hissa & Hissa, 2023). A counterargument against the emphasis on biology is an emphasis on sociocultural factors that contribute to identity and athleticism (Jakubowska, 2024). A related argument against

dichotomous sex categories is that levels and fluctuations of naturally occurring hormones can vary widely across developmental time and social situations (Parks, 2016).

Research on content analysis of legislative discourse uncovered persistent themes on both sides of the argument. Arguments against transgender inclusion aligned more closely with politically conservative ideology and reflected values of elitism, tradition, fairness, safety and equal opportunity (Flores et al., 2020; Martin & Rahilly, 2023; Jakubowska, 2024). Arguments supporting transgender inclusion aligned more closely with politically liberal ideology and focus more on human rights and gender fluidity (Moyer, 2022; Posbergh, 2023).

Although support for transgender athletes is increasing, the overriding public perception remains negative, particularly in certain societal groups (Avalos et al., 2022). Women and individuals who hold more flexible beliefs about gender roles tend to be more supportive of transgender athletes than men and individuals who hold traditional gender role beliefs (Flores et al., 2020; Goldbach et al., 2022). Regardless of whether the fan is male or female, sports fans are less supportive of transgender athletes than the general public (Flores et al., 2020). In general, transgender women experience more prejudice and discrimination than transgender men (Read et al., 2020; Worthen, 2013), and this pattern appears to hold true for athletes as well. For example, transgender female athletes are subjected to greater scrutiny and regulation than are transgender male athletes (Fischer & McClearen, 2020). This scrutiny could be due to the fact that even among athletes who support the inclusion of their transgender peers, there is some acknowledgement that there is a competitive advantage for transgender women (Devine, 2022).

Our study responded to the need for additional research on this topic by examining attitudes toward transgender individuals using an experimental design that manipulated a description of an athlete by varying the target athletes' gender and ability level. We were interested in whether additional characteristics of the perceiver would influence perceptions. Previous research on athletes in general demonstrates that empathetic coaches elicit higher performance (Krykun et al., 2023; Sear, 2023), which has led to a call for further investigation into the role of empathy in the world of sports. We also examined belief in a just world as a predictor of perceptions, another factor that has yet to be applied to transgender athletes. Belief in a just world is associated with ideology linked to negative attitudes toward transgender athletes, such as equal opportunity,

fairness, and political conservatism (Clifton & Kerry, 2022; Sappington et al., 2019). These links suggested that belief in a just world would likely predict less support of transgender athletes. We hypothesized that

1. participants would respond most negatively to the transgender female athlete performing at a high ability level than to the other transgender athletes.
2. higher levels of empathy would predict more positive attitudes toward transgender athletes.
3. a stronger adherence to belief in a just world, for self or others, would predict more negative attitudes toward transgender athletes.
4. identifying as more politically liberal would predict more positive attitudes toward transgender athletes.

We also conducted exploratory analyses on whether participants' age, education level, and sports participation related to their perceptions of transgender athletes.

Method

Participants

Participants were young adults recruited through college classrooms, social media, and word-of-mouth. We had 108 participants complete our survey; however, three participants were excluded because they failed to respond correctly to our validity check items. Thus, our final sample was 105 young adults. One participant responded to all survey items except demographic information. The mean age of participants was 27.72 (SD = 13.39). When asked to identify their gender identity, 89 participants were cisgender women, 8 were cisgender men, and 7 categorized themselves as non-binary or "other." When asked to identify their race/ethnicity, 67 participants were White, 26 were Black, 5 were Hispanic/Latinos, and the remainder were Multiracial/Other. When asked to identify their sexual orientation, 79 participants were heterosexual, 16 were bisexual, 1 was homosexual, and 8 were "other." We asked participants about their sports participation; 46 participants had never participated in sports or participated only as a child, 44 had participated during high school, and 14 had participated beyond the high school level. When asked about their educational attainment, 8 participants reported only a high school degree, 53 had some college education, 27 had a college degree, 11 had some graduate training, and 5 attained a terminal graduate degree. We asked participants what best described their political ideology; 8 participants reported being very conservative, 16 as conservative, 12 as slightly conservative, 20 as slightly liberal, 38 as liberal, and 10 as very liberal.

In the online data collection system, participants were randomly divided into four conditions with 26

participants in the transgender man/low ability condition, 25 participants in the transgender man/high ability condition, 27 participants in the transgender woman/low ability condition, and 24 participants in the transgender woman/high ability condition. This project was approved by the university's Institutional Review Board, and participants agreed to voluntarily access the survey in order to participate.

Materials and Procedure

Participants first responded to the 16-item Belief in a Just World Scale, which consists of two sub-scales (Lucas et al., 2011). One sub-scale assessed the extent to which individuals trust that people are treated fairly in the world and they deserve the outcomes that they receive. A sample item is "I feel that people generally earn the rewards and punishments that they get in this world." The second sub-scale assessed the extent to which individuals trust that they are personally treated fairly in the world and deserve the outcomes that they receive. A sample item is "I feel that I generally earn the rewards and punishments that I get in this world." Responses were made on a 5-point scale where 1 represented "strongly disagree" and 5 represented "strongly agree." The published reliabilities for the other-focused scale range from .79 to .93, and for the self-focused scale from .88 to .94 (Lucas et al., 2011). We achieved a Cronbach's alpha of .88 for the other-focused scale and .90 for the self-focused scale.

Participants were then asked to respond to the 8-item Empathy Quotient scale (Loewen et al., 2010), which measured individuals' ability to understand and empathize with the feelings of others. A sample item is "I find it easy to put myself in somebody else's shoes." Responses were made on a 5-point scale where 1 represented "strongly disagree" and 5 represented "strongly agree." The published reliability for this scale is .76 (Loewen et al., 2010), and we achieved a Cronbach's alpha of .82.

Next, participants were randomly assigned to one of four experimental conditions. In each condition, participants read a story about a transgender athlete who was on a college basketball team. The stories varied by the gender and performance level of the athlete resulting in four conditions: Transgender man with low ability on the team; Transgender man with high ability on the team; Transgender woman with low ability on the team; Transgender woman with high ability on the team. The scenarios can be seen in Appendix A. Immediately after reading the story, participants encountered three validity check items to ascertain if they read the story. They were asked the name of the athlete at birth, the sport the athlete played and the skill level of the athlete. In order to

remain in the study, participants had to answer at least two of the questions correctly, and one of the correct answers had to be the skill level of the athlete.

Participants then encountered 11 researcher-created items to examine their perceptions of the transgender athletes described in their experimental condition. The items can be seen in Table 1. These items addressed participants' support of the athlete, perception that the athlete had an unfair advantage over others, belief that the athlete was a champion for transgender rights, and agreement that the athlete encountered prejudice and discrimination. We also asked the extent to which participants felt sorry for the athlete's teammates and felt that it was harder to coach transgender athletes. Responses were made on a 5-point scale where 1 represented "strongly disagree" and 5 represented "strongly agree." We then asked participants to evaluate their personal feeling about transgender athletes and to evaluate how the "average adult" felt about transgender athletes. Responses were made on a 5-point continuum that ranged from "very negatively" to "very positively."

Participants responded to six demographic items that included age, gender, race, sexual orientation, highest level of education, and political ideology. Last, we provided the following statement to participants: "This study is investigating the issue of transgender adults participating in sports that match their current identity rather than their assigned birth identity. Do you have any additional thoughts you want to share on this topic?" We provided an open-ended field with unlimited space.

Results

We ran a 2 (gender) X 2 (ability) MANOVA and found a consistent pattern of support for the transgender man with a high ability level compared to the transgender woman with a high ability level. Participants agreed the most that they would be supportive if this athlete were playing on their college team when the athlete was a transgender man with a high ability level; they were least supportive of a transgender woman with a high ability level [$F(1, 101) = 5.88, p = .02$]. Participants agreed the most that the athlete had an unfair advantage over other players on the team when the athlete was a transgender woman with a high ability level; they agreed the least when the athlete was a transgender man with a high ability level [$F(1, 101) = 5.99, p = .02$]. Similarly, participants agreed the most that the athlete was cheating another athlete out of a spot on the team when the athlete was a transgender woman with a high ability level; they agreed the least when the athlete was a transgender man with a high ability level [$F(1, 101) = 4.28, p = .04$]. Participants also agreed most that the teammates have to be very careful what they say around a transgender

Table 1

Means (and Standard Deviations) for Perceptions of Transgender Athlete Questions by Transgender Status (TG) and Ability Level (A)

	TG Man/ Low A	TG Man/ High A	TG Woman/ Low A	TG Woman/ High A
If this were my college, I would be supportive of this athlete participating on the team.	3.46 (1.42)	4.04 (1.27)	3.96 (1.43)	3.21 (1.41)
This athlete has an unfair advantage over the other players on the team.	2.08 (1.32)	1.92 (1.0)	2.30 (1.35)	3.38 (1.38)
This athlete has as much right to be on the team as anyone else.	3.42 (1.58)	3.64 (1.35)	4.0 (1.33)	3.67 (1.24)
Another athlete is being cheated out of a spot on the team because of this transgender athlete.	2.58 (1.45)	2.16 (1.28)	2.07 (1.44)	2.79 (1.35)
I'm sure the transgender athlete encounters prejudice and discrimination.	4.31 (1.12)	4.32 (.90)	4.56 (1.09)	4.62 (.50)
Society should accept and support transgender athletes.	3.50 (1.48)	3.76 (1.39)	4.26 (1.23)	3.88 (1.26)
Situations like these are ruining women's athletics.	2.77 (1.51)	2.40 (1.53)	2.11 (1.31)	2.29 (1.49)
I bet the teammates have to be very careful what they say around the transgender athlete.	3.42 (1.17)	3.24 (1.23)	2.63 (1.57)	3.46 (1.18)
It is harder to coach the transgender athlete than the other athletes.	2.46 (1.30)	2.44 (1.12)	1.85 (.99)	2.21 (1.25)
I feel sorry for the transgender athlete's teammates.	1.96 (1.18)	2.08 (1.12)	1.89 (1.42)	2.50 (1.35)
This athlete is a champion for transgender rights.	2.96 (1.34)	3.68 (1.15)	3.78 (1.34)	3.46 (.98)

woman with a high ability level and a transgender man with a low ability level [$F(1, 101) = 3.82, p = .05$]. The observed pattern shifted when participants were asked whether the athlete was a champion for transgender rights; participants agreed most when the athlete was a transgender woman with a low ability level and agreed least when the athlete was a transgender man with a low ability level [$F(1, 101) = 4.63, p = .03$]. The associated means and standard deviations can be seen in Table 1.

Across conditions, participants expressed agreement that the athlete had as much right to be on the team as anyone else ($M = 3.67, SD = 1.37$), that society should accept and support transgender athletes ($M = 3.84, SD = 1.35$) and that the transgender athletes encountered prejudice and discrimination ($M = 4.45, SD = .94$). Participants slightly disagreed that "situations like these are ruining women's athletics" ($M = 2.40, SD = 1.45$). Participants also slightly disagreed that they felt sorry for the transgender athletes' teammates ($M = 2.12, SD = 1.28$) and that coaching the transgender athlete was

harder than coaching the other teammates ($M = 2.23, SD = 1.17$). We found no main effects for these variables.

We focused on the items on which participants provided similar responses across conditions and then examined whether participant characteristics predicted those perceptions using Pearson's correlations. Empathy did not predict perceptions of the transgender athletes in the experimental stories; however, the higher the participants were in empathy, the more positively they rated their overall feelings about transgender athletes [$r(104) = .20, p = .04$]. This correlation represents a weak effect size.

Participants' belief that the world is just for them personally did not predict their perceptions of athletes. In contrast, participants with a stronger belief in a just world for others agreed less that the transgender athletes had as much right to be on the team as anyone else [$r(105) = -.22, p = .02$] and that the transgender athletes encountered prejudice and discrimination [$r(103) = -.32, p = .001$]. These participants agreed more that they felt

sorry for the transgender athletes' teammates [$r(104) = .22, p = .02$] that coaching the transgender athlete was harder than coaching the other teammates [$r(103) = .24, p = .02$], and that situations like these are ruining women's athletics [$r(103) = .21, p = .03$]. The higher participants belief that the world is just for others, the more negatively they rated their overall feelings about transgender athletes. These correlations represent weak effect sizes.

Political ideology predicted all of the examined variables. The more politically liberal participants were, the more they agreed that the transgender athlete had as much right to be on the team as anyone else [$r(104) = .72, p < .001$], that society should accept and support transgender athletes [$r(103) = .67, p < .001$] and that the transgender athletes encountered prejudice and discrimination [$r(103) = .40, p < .001$]. The more politically liberal participants were, the more they disagreed that "situations like these are ruining women's athletics" [$r(103) = -.61, p < .001$], they felt sorry for the transgender athletes' teammates [$r(104) = -.53, p < .001$] and that coaching the transgender athlete was harder than coaching the other teammates [$r(103) = -.52, p < .001$]. The more politically liberal they were, the more positively they rated their overall feelings about transgender athletes [$r(104) = .75, p < .001$]. These correlations represent moderate to strong effect sizes.

The older participants were, the more they agreed that transgender athletes encounter prejudice and discrimination [$r(99) = -.22, p = .03$]. The older the participant, the less positively they felt about transgender athletes [$r(100) = -.22, p = .03$] and the more they agreed that transgender athletes are harder to coach than their teammates [$r(99) = .26, p < .01$]. The more positively participants rated their own feelings about transgender athletes, the more positively they believed the "average adult" felt about transgender athletes [$r(104) = .27, p < .01$]. These correlations represent weak effect sizes. Education level did not predict perceptions of transgender athletes. The extent of participants' sports participation only predicted one variable. The greater the length of sports' participation, the less participants agreed that transgender athletes face prejudice and discrimination [$r(103) = -.20, p = .05$], reflecting a weak effect size.

Two raters categorized the qualitative comments, with an inter-rater reliability of .96. On the optional open-ended prompt, we had 23 participants respond. Responses were categorized based on their support for transgender athletes as uncertain, supportive, not-supportive, or mixed. Of the responders, seven participants provided mixed feedback, six provided

positive, six provided negative, and four expressed uncertainty. Sample response options can be seen in Appendix B.

Discussion

We hypothesized that participants would respond more negatively to the transgender female athlete especially when she was performing at a higher ability level than her teammates. This prediction was supported. When the athlete was a skilled transgender woman, participants were least supportive, perceived her as having an unfair advantage, agreed that she was cheating another athlete out of a spot on the team, and that her teammates had to be careful about what they could say in her presence. Participants had the most supportive views on these same issues for the transgender man with a high level of skill. These findings reflect the commonly held perspective that a biologically male body has unfair advantages over a biologically female body (Martin & Rahilly, 2023). A recent example is the transgender female, Lia Thomas, a Division I college swimmer who transitioned after going through male puberty; her competitive success elicited a firestorm of controversy related to the extent of her biological advantages (Shermer, 2022). Even athletes who support the inclusion of transgender athletes report agreement that cisgender women are athletically disadvantaged in comparison to transgender women (Devine, 2022). This data lends support to the limited, but growing, body of evidence documenting greater prejudice toward transgender female athletes compared to transgender male athletes. Our study provides additional confirmation that these attitudes are tied to perceptions of performance ability. The transgender female athlete who performed at the highest level elicited the most negative attitudes, reflecting participants' belief that she achieved success due to biological advantages that her teammates did not have. Similarly, the transgender male athlete in our study elicited the most support when he performed at a high skill level despite not having the advantage of going through male puberty.

Our data revealed a slightly different pattern when we asked participants whether the athlete was a champion for transgender rights. Participants agreed most that the transgender woman with low ability was a champion. In other words, this athlete was biologically expected to perform well but was not performing well. Our participants may have perceived this athlete as representing a courage and perseverance in the face of adversity that highlighted the struggle for acceptance that transgender athletes face. Our participants were least impressed with the transgender man with low ability. This situation matched social expectations (biological

disadvantage leading to lower performance), and champion status is not granted to individuals who match the norm.

It is worth noting that these differences in support level were relative. Overall, our quantitative data suggested that participants held favorable perceptions of all of the athletes. Across conditions, participants agreed that the athletes had the right to compete, should be supported, and were similar to other athletes on the team. Participants also agreed that the transgender athletes dealt with prejudice and discrimination. In other words, our participants reflect the growing support for transgender athletes seen in the public at large (Avalos et al., 2022). In addition, previous researchers documented that women tend to be more supportive of transgender athletes than men (Flores et al., 2020; Goldbach et al., 2022), and our sample was primarily women.

A second goal of our study was to examine whether perceiver characteristics influenced perceptions of transgender athletes. We hypothesized that greater empathy would predict more positive attitudes toward transgender athletes and found limited support for this prediction. Participants responded similarly to the transgender athletes in the experimental stories regardless of their empathy level. The only difference we found was when we directly asked participants how they personally felt about transgender athletes in general. The more empathetic the participant, the more positively they reported feeling. In other words, empathetic participants did not respond more favorably to transgender athletes in our scenarios than did less empathetic participants; however, more empathetic participants believed that they had the most favorable attitudes toward transgender athletes. One interpretation might be that empathetic individuals are not more behaviorally or attitudinally positive toward transgender athletes, they simply assume that they are because of their self-perception of being empathetic people. Another possible explanation is that empathetic individuals have a sincere concern for transgender athletes; however, in specific situations, their feelings of fairness are more salient in driving their reactions.

In contrast to empathy, belief in a just world emerged as a significant predictor of attitudes. Failing to support our hypothesis, we found that belief in a personally just world was not predictive of attitudes. However, matching our hypothesis, a stronger adherence to belief in a just world for others predicted more negative attitudes toward transgender athletes. Participants who agreed more that the world was just (for others) did not support the athletes' right to be on the team, disagreed that the athlete encountered prejudice,

felt sorry for the teammates and coach, and believed that these situations are ruining women's athletics. They also directly reported that they had more negative feelings toward transgender athletes. These correlations were weak but revealed a consistent pattern. Individuals who believe the world is just tend to support the inherent fairness of the natural order and structure of the world. These individuals may have seen the transgender athletes as inappropriately modifying that natural order. The fact that attitudes about justice were more predictive than empathy supports our earlier argument that feelings of concern for these transgender athletes might exist apart from feelings of fairness and attitude agreement.

Belief in a just world is linked to conservative political ideology (Clifton & Kerry, 2022; Sappington et al., 2019), and our data revealed that political conservatism was a similar and even stronger predictor of negative attitudes toward transgender athletes. Matching our hypothesis, the more politically liberal participants were, the more they supported transgender athletes, viewed them as similar to their teammates, and felt they experienced prejudice. Our data is consistent with limited research on this topic and meets the need for additional, confirmatory evidence that political ideology that is less tied to tradition is linked to more positive attitudes (e.g., Avalos et al., 2022; Kavoura & Kokkonen, 2021).

Because data on transgender athletes is limited, we also explored whether participants' age, education level, and sports participation related to their perceptions of transgender athletes. Education level and sports participation were not predictive, with one exception. The longer individuals had played sports, the less they believed that transgender athletes face prejudice and discrimination. This outcome could reflect a protectiveness of athletes wanting other athletes or their sports to be perceived favorably. Another explanation is that more serious athletes may place more weight on other athletes' ability level instead of aspects of the other athletes' identities. In other words, athletes may prioritize a useful teammate over other factors.

Older participants agreed that transgender athletes faced prejudice and discrimination, possibly because they also agreed that they felt more negatively about transgender athletes. One factor to consider is that older adults are more likely to hold traditional or conservative views (e.g., Peterson et al., 2020), and these views are associated with more negative attitudes toward transgender individuals. The presence of transgender athletes is a relatively new social issue, and thus, young adults, like those in our study, may be more familiar with the topic and less guarded in their attitudes.

Our qualitative data reflected much of the quantitative data; however, when participants were given the opportunity to explain their feelings, the findings revealed complexity and contradictions. Over half of the responses were either not supportive of or offered mixed support for transgender athletes competing based on their gender identity. A common response started with a globally supportive comment followed by “but...” and a qualifying comment. Even among those participants who said they generally supported transgender athletes an emergent theme was a focus on the biological advantages of men over women. These comments mirror the public conversation during policy development (e.g., Fisher & McClearen, 2020), as well as the belief, even among athletes, that transgender women enjoy a competitive advantage (Devine, 2022). It might be noteworthy that the vast majority of participants chose not to respond to the prompt, which may reflect a lack of motivation or possibly a social desirability issue where participants were reluctant to explain their opinions. The missing responses might also be a sign of conflicting opinions that were difficult to explain. Of participants who responded, 18% provided responses that cited confusion or misunderstandings. The inconsistency that we found between professed support alongside perceptions of possible unfairness is not unique to our study. Goldbach et al. (2022) surveyed college athletes across the United States and found similar misunderstandings and conflicting viewpoints about transgender athletes. Devine (2022) interviewed female Olympians and, despite their general support for transgender athletes, the Olympians reported a fear of asking questions or saying the wrong thing in the presence of their transgender peers.

The public's level of support is worthy of examination as it influences policies and funding (Martin & Rahilly, 2023; Moyer, 2022; Posbergh, 2023). Social support also affects the transgender athletes themselves. Pockets of vulnerability for transgender athletes include misuse of pronouns, discomfort in the locker room, and lack of mental health resources (Morris & Van Raalte, 2016). Transgender athletes report high levels of stress and fear of self-disclosure (Herrick et al., 2020; Walker et al., 2018). In a recent meta-analysis, Chan et al. (2024) reported that transgender athletes who face discrimination disproportionately present with mental health issues.

Our study had limitations that can be addressed by future researchers. Our participants were primarily women and politically liberal, which may have been a contributing factor to the generally positive attitudes we reported. We did not find outcomes predicted by education level, which may be an accurate finding or may

reflect the fact that most of our participants had a college or higher degree. In addition to sampling broader demographics, future researchers may want to focus on athletes' perceptions of transgender athletes with varied levels of ability. Another avenue of research might be to explore whether the transgender athlete's affiliation with an individual versus team sport might impact perceptions of the athlete.

In sum, we found that participants expressed general support for transgender athletes and an awareness that these athletes face prejudice and discrimination. The transgender man with a high level of ability elicited the most support, while the transgender woman with a high level of ability elicited the lowest levels of support. As supported by our qualitative data, these findings reflect a focus on the biological advantages that accompany male puberty. Participants who believed in a just world, along with those who were politically conservative, lower in empathy, and older, were less supportive than their counterparts. These findings suggest that perceptions of transgender athletes are based on the athletes' gender identity and performance, as well as characteristics of the perceiver. An enhanced understanding of the public's perception may be useful as transgender athletes are currently at the center of controversy, and corresponding policies are in flux

References

- Avalos, S., MacKenzie, K. & Monk-Turner, E. (2022). Out of the closet and onto the Olympic floor: A qualitative look at social media user's perceptions of transgender Olympic athletes. *Gender Issues*, 39, 437-454. <https://doi.org/10.1007/s12147-022-09299-6>
- Chan, A. S. W., Choong, A., Phang, K. C., Leung, L. M., Tang, P. M. K. & Yan, E. (2024). Societal discrimination and mental health among transgender athletes: A systematic review and meta-analysis. *BMC Psychology*, 12, 24. <https://doi.org/10.1186/s40359-023-01493-9>
- Clifton, J. D. W. & Kerry, N. (2022). Belief in a dangerous world does not explain substantial variance in political attitudes, but other world beliefs do. *Social Psychological and Personality Science*, 14, 515-525. <https://doi.org/10.1177/1948550622119324>
- Devine, C. (2022). Female Olympians' voices: Female sports categories and International Olympic Committee transgender guidelines. *International Review for the Sociology of Sport*, 57, 335-361. <https://doi.org/10.1177/10126902211021559>
- Fischer, M. & McClearen, J. (2020). Transgender athletes and the queer art of athletic failure. *Communication & Sport*, 8, 147-167. <https://doi.org/10.1177/2167479518823207>

- Flores, A. R., Haider-Market, D. P., Lewis, D. C., Miller, P. R., Tadlock, B. L., & Taylor, J. K. (2020). Public attitudes about transgender participation in sports: The roles of gender, gender identity conformity, and sports fandom. *Sex Roles*, 83, 382–398. <https://doi.org/10.1007/s11199-019-01114-z>
- Goldbach, C., Chambers-Baltz, S., Feeser, K., Milton, D. C., McDurmon, P. & Knutson, D. (2022). Transgender inclusion in competitive sport: Athletes' attitudes toward transgender athlete participation in intercollegiate sport. *International Journal of Sport and Exercise Psychology*, 22, 802-824. <https://doi.org/10.1080/1612197X.2022.2161109>
- Heather, A. K. (2022). Transwoman elite athletes: Their extra percentage relative to female physiology. *International Journal of Environmental Research and Public Health*, 19, 9103. <https://doi.org/10.3390/ijerph19159103>
- Herrick, S. C., Rocchi, M. A. & Couture, A. L. (2020). A case study exploring the experiences of a transgender athlete in a synchronized skating, a subdiscipline of figure skating. *Journal of Sport and Social Issues*, 44, 421-449. <https://doi.org/10.1177/0193723520919816>
- Hilton, E. N. & Lundberg, T. R. (2021). Transgender women in the female category of sport: Perspectives on testosterone suppression and performance advantage. *Sports Medicine*, 51, 199–214. <https://doi.org/10.1007/s40279-020-01389-3>
- Hissa, M. R. N. & Hissa, M. N. (2023). The influence on sports policies on the right to fair competition for women. *International Journal of Human Rights in Healthcare*, 16(2), 120-127. <https://doi.org/10.1108/IJHRH-04-2021-0084>
- Jakubowska, H. (2024). Who counts as a woman? A critical discourse analysis of petitions against the participation of transgender athletes in women's sport. *International Review for the Sociology of Sport*, 59, 203-221. <https://doi.org/10.1177/10126902231194570>
- Kavoura, A. & Kokkonen, M. (2021). What do we know about the sporting experiences of gender and sexual minority athletes and coaches? A scoping review. *International Review of Sport and Exercise Psychology*, 14, 1-27. <https://doi.org/10.1080/1750984X.2020.1723123>
- Krykun, O., Voronova, V. & Fedorchuk, S. (2023). The role of empathy in the coach-athlete relationship in cheerleading. *European Psychiatry*, 66, S1017.
- Loewen, P., Lyle, G. & Nachshen, J. (2010). An eight-item form of the Empathy Quotient (EQ) and an application to charitable giving. Retrieved from https://www.researchgate.net/publication/254000149_An_eight-item_form_of_the_Empathy_Quotient_EQ_and_an_application_to_charitable_giving
- Lucas, T., Zhdanova, L., & Alexander, S. (2011). Procedural and distributive justice beliefs for self and others: Assessment of a four-factor individual differences model. *Journal of Individual Differences*, 32, 14–25. <https://doi.org/10.1027/1614-0001/a000032>
- Martin, K. & Rahilly, E. (2023). Value frames in discourse supporting transgender athlete bans. *Discourse and Society*, 34, 732-751. <https://doi.org/10.1177/09579265231187058>
- Morris, J. F. & Van Raalte, J. L. (2016). Transgender and gender nonconforming athletes: Creating safe spaces for all. *Journal of Sport Psychology in Action*, 7, 121-132. <https://doi.org/10.1080/21520704.2016.1184732>
- Moyer, V. (2022) Revising trans-exclusionary narratives in women's sports activism. Who are the "women" of women's athletics? In: Magrath R (ed.) *Athletes' Activism: Contemporary Perspectives*. Abingdon: Routledge, pp. 88–98.
- Parks, L. P. (2016). *Sex Testing: Gender Policing in Women's Sports*. Chicago, Illinois: University of Illinois Press.
- Peterson, J. C., Smith, K. B. & Hibbing, J. R. (2020). Do people really become more conservative as they age? *The Journal of Politics*, 82, 600-611. <https://doi.org/10.1086/706889>
- Pigozzi, F., Bigard, X., Steinacker, J., et al. (2022). Joint position statement of the International Federation of Sports Medicine (FIMS) and European Federation of Sports Medicine Associations (EFSMA) on the IOC framework on fairness, inclusion and nondiscrimination based on gender identity and sex variations. *BMJ Open Sport & Exercise Medicine*, 8, e001273. <https://doi.org/10.1136/bmjsem-2021-001273>
- Posbergh, A. (2023). Contradiction of cohesion? Tracing questions of protection and fairness in scientifically driven elite sport policies. *Sociology of Sport Journal*. Epub ahead of print 14. <https://doi.org/10.1123/ssj.2022-0198>
- Read, J., Sargeant, C. & Wright, S. (2020). What beliefs influence children and young people's attitudes towards the transgender population? *Educational & Child Psychology*, 37, 11– 36. <https://doi.org/10.53841/bpsecp.2020.37.1.11>
- Reynolds, A. & Hamidian, J. A. (2021). Transgender athletes in sports competitions: How policy measures can be more inclusive and fairer to all. *Frontiers in Sports and Active Living*, 3, 704178. <https://doi.org/10.3389/fspor.2021.704178>
- Safer, J. D. (2022). Fairness for transgender people in sport. *Journal of the Endocrine Society*, 6,

- bvac035.
<https://doi.org/10.1210/jendso/bvac035>
- Sappington, R., Keum, B. T. & Hoffman, M. A. (2019).
 “Arrogant, ungrateful, anti-American degenerates”: Development and initial validation of the Attitudes Toward Athlete Activism Questionnaire (ATAAQ). *Psychology of Sport and Exercise*, 45, 101552.
<https://doi.org/10.1016/j.psychsport.2019.101552>
- Sear, P. (2023). *Empathic Leadership: Lessons from Elite Sport*. New York, NY: Routledge.
- Shermer, M. (2022, July 22). *Trans athletes & conflicting rights*. Skeptic.
https://www.skeptic.com/reading_room/trans-athletes-conflicting-rights/
- Walker, A., Sexton, L., Valcore, J. L., Sumner, J. & Wodda, A. (2018). Transitioning to social justice: Transgender and nonbinary individuals. In C. Roberson (Ed.), *Routledge Handbook of Social, Economic and Criminal Justice* (pp. 220–233). Routledge.
- Worthen, M. (2013). An argument for separate analyses of attitudes toward lesbian, gay, bisexual men, bisexual women, mtf and ftm transgender individuals. *Sex Roles*, 68, 703-723.
<https://doi.org/10.1007/s11199-012-0155-1>

Appendix A

Experimental Conditions Varied by Gender Status (TG) and Ability Level (A)

TG Man Narrative

Mike was born a female. His given birth name was Molly.

Even before puberty started, Mike recognized that he did not feel like other teen-age girls. He questioned his identity on a daily basis. Despite these feelings, he was a star on the high school girls' basketball team and on female travel teams for basketball. In college, Mike started to see a psychologist and eventually decided to transition to being a man. He took estrogen and drugs to suppress testosterone and other male hormones. Two years after Mike transitioned, he tried out for the men's basketball team on his college campus.

TG Man/ Low A

He made the team but was not one of the better players on the team. He spent most of the season sitting on the bench. Mike was happy to be on the team, but he felt discouraged by sitting on the bench and by the rude comments he sometimes heard other players say behind his back.

TG Man/ High A

He made the team and worked his way to being a starting guard. He got more playing time than some of the players who had been on the team before he arrived. Mike was happy to be on the team, but he felt discouraged by the rude comments he sometimes heard other players say behind his back.

TG Woman Narrative

Molly was born a male. Her given birth name was Mike.

Even before puberty started, Molly recognized that she did not feel like other teen-age boys. She questioned her identity on a daily basis. Despite these feelings, she was a star on the high school boys' basketball team and on male travel teams for basketball. In college, Molly started to see a psychologist and eventually decided to transition to being a woman. She took testosterone and drugs to suppress estrogen and other female hormones. Two years after Molly transitioned, she tried out for the women's basketball team on her college campus.

TG Woman/ Low A

She made the team but was not one of the better players on the team. She spent most of the season sitting on the bench. Molly was happy to be on the team, but she felt discouraged by sitting on the bench and by the rude comments she sometimes heard other players say behind her back.

TG Woman/ High A

She made the team and worked her way to being a starting guard. She got more playing time than some of the players who had been on the team before she arrived. Molly was happy to be on the team, but she felt discouraged by the rude comments she sometimes heard other players say behind her back.

Appendix B

Sample Qualitative Comments to Open-Ended Prompt

Mixed Support ($n = 7$)

I support transgender athletes however there is a biological advantage that has to be acknowledged. Biological males have more muscle mass than females and thus have physical advantage that should be acknowledged.

I feel like I 100% support transgender athletes. My only issue is that I don't know the make up of how fair it actually is. All I see is stuff on Facebook about how it is so unfair for a transgender woman to compete in women's sports. If I knew more, I would have better answers.

Supportive ($n = 6$)

They are humans. They have a right to play sports. People saying they shouldn't do not care about human beings and suck.

I just think that people should respect other people's identities, whether they are cisgender or transgender because it doesn't effect [sic] anybody other than the person themselves. In other words, you do you.

Not Supportive ($n = 6$)

It's a difficult topic because men's bodies are built different than women's. It seems unfair for a man's body to be able to participate in a woman's sport.

I think we should be kind to transgender individuals just like any other person no matter how they identify or the choices they make. However, there is a biological difference no matter how they choose to identify and I feel strongly that biology should determine if they play on the men or women's team. I don't think they shouldn't be allowed to play, but I think based on their genetics and birth given gender, they should play on that team to keep the game "fair."

Uncertain ($n = 4$)

I feel there needs to be more education concerning the differences in transgender people's bodies when they have or have not gone through the puberty of their assigned sex. I feel conflicted.

I am confused.

EMPATHY AND PERCEPTIONS OF FEMALE SPEAKERS VARYING IN CONFIDENCE

MERIDEE A. RITZER AND DONNA W. NELSON

WINTHROP UNIVERSITY

Abstract – We manipulated the confidence of a female speaker and assessed listeners’ characteristics and perceptions of the speaker. We hypothesized that the speaker would be perceived more positively and information recalled more accurately when the speaker was confident. We also hypothesized that listeners high in empathy would evaluate the low confident speaker less harshly. Participants were randomly assigned to one of two experimental conditions and watched a 3-minute video of a female speaker describing a research study. The speaker presented as either high confidence or low confidence. We assessed participants’ perceptions of the speaker, recall, confidence in recall, cognitive empathy, and affective empathy. We achieved mixed support for our hypotheses. As predicted, the confident speaker was viewed more positively and less negatively; however, she did not elicit better recall. Listeners’ empathy did not predict their perceptions with one exception. Listeners higher in affective empathy evaluated the low confidence speaker most negatively and the high confidence speaker least negatively. Men had more accurate recall of the presented information than did women, which likely is a result of men having more prior research experience. These findings support previous research that demonstrate the salience of speaker confidence, which may be even more influential than listener characteristics in driving perceptions.

Conveying confidence is a valuable characteristic in social contexts. An individual who exhibits confidence typically experiences higher social status, is seen as more trustworthy, has greater persuasive power, and is perceived as more knowledgeable (Roche et al., 2023). One way to convey confidence is through speech. Confident speakers are perceived as more committed and believable (Jiang et al., 2020; Mori & Pell, 2019). Listeners are more willing to trust and learn from speakers who convey confidence (Birch et al., 2020; Pozzi & Mazzarella, 2023). The “confidence heuristic” is a term used to reflect the reliably demonstrated fact that listeners’ prefer and attribute more positive characteristics to confident, versus uncertain, speakers (Birch et al., 2020; Thomas & McFadyen, 1995).

Speaker confidence is conveyed to listeners through both verbal and non-verbal cues. For example, stating information in a factual manner and avoiding meaningless hand gestures convey confidence, while a slumped posture and distressed facial expression convey a lack of confidence (Pozzi & Mazarella, 2023). Tentative wording such as “I guess” are even more powerful signals than are non-verbal cues that a speaker lacks confidence (Tenney et al., 2019). Mori and Pell (2019) asked observers to guess speakers’ confidence levels, while the

speakers engaged in tasks that induced either high or low confidence. Even with the audio muted, observers were accurate at identifying speaker confidence. One biological theory to explain the preference for confident speakers is that the vocal characteristics associated with confidence lead to deeper cognitive processing (Guyer et al., 2021). Jiang and Pell (2015) found that confident speakers elicited stronger event-related potentials in listeners’ brains than did uncertain speakers.

In general, female speakers are perceived as less confident than male speakers (Roche et al., 2019, 2022). Female speakers also carry a heavier burden than male speakers, because the manner in which they convey information is more influential in determining listener perceptions. Roche et al. (2022) demonstrated that female speakers are viewed as more uncertain than male speakers, and this difference becomes even more pronounced when female speakers use rising intonation (i.e., a questioning tone) at the end of their sentences. Similarly, male and female speakers are perceived as equally agreeable unless the female speaker pauses before an affirmative statement, in which case her agreeableness rating declines (Roberts & Norris, 2016). Men maintain their ratings of intelligence regardless of whether they speak with vocal fry (also known as creaky voice);

however, women's intelligence ratings decrease if they speak with vocal fry (Taylor et al., 2022). Women may have some intuitive understanding of this gender bias, because women are less likely to give a public presentation when provided the opportunity than are men, even when incentives are offered (DePaola et al., 2021).

Listener characteristics can also influence perceptions of speakers. Much of the research in this area focuses on the linguistic background of the listener and whether it matches, or does not match, the linguistic presentation of the speaker. For example, listeners' familiarity with non-native speech influences their perceptions of non-native speakers (Woolridge et al., 2023). Outside of this particular research framework, there is limited and inconsistent evidence about whether listener characteristics affect perceptions to the extent that speaker characteristics do. Hancock and Pool (2017) recruited cisgender men, cisgender women and transgender women, of varied sexual orientations, to evaluate auditory recordings made by cisgender men, cisgender women, and transgender women. The only significant consequence of listeners' demographics was that heterosexual listeners, compared to non-heterosexual listeners, perceived the transgender women as slightly less feminine. In other words, listeners' gender and sexual orientation were not salient variables in driving their perceptions of a speaker. More recent researchers found no effect of listener gender on attributions made about speakers' gender (Brown et al., 2021b; Hancock et al., 2023). As further evidence that individual listener characteristics might exert minimal influence, researchers have observed that listeners often display a high level of agreement when evaluating speakers. This consensus among listeners persists even when their perceptions are inaccurate, suggesting that listeners rely on shared biases or common evaluative criteria (e.g., Mahrholz et al., 2018; Welch et al., 2021).

One exception is that an extensive body of research addresses the listener characteristic of empathy. However, the focus has been on the importance of active and empathetic listening in interpersonal situation such as between doctor and patient (e.g., Junaid et al., 2022), between mental health clinician and client (e.g., Aadam et al., 2024), or in social dyads (e.g., Brown et al., 2021a). Empathy can exist as either a stable trait or elicited, and both facilitate these social interactions (Jankowiak-Siuda et al., 2011). Individuals with either type of empathy show enhanced responsiveness to speakers, particularly when the topic has an emotional component. For example, listening to vocal expressions of pain result in increased activity in areas of the brain associated with affective and

cognitive aspects of empathy (Mauchand et al., 2023). In a modern twist, Carolus et al. (2021) found that observers reported greater empathy for an Amazon Echo, a voice-activated virtual assistant device, when a research confederate spoke to it rudely versus neutrally.

Despite the recognition that empathy impacts listeners' physiology and emotions, limited research has assessed whether listeners' empathy predicts perceptions of speaker characteristics. Our study brought together these two lines of research. We examined perceptions of a young adult female speaker who conveyed varying degrees of confidence while addressing a young adult audience. We also examined the influence of listeners' ability to recognize the feelings of others (cognitive empathy) as well as their ability to feel the feelings of others (affective empathy) on these perceptions. We hypothesized that:

1. compared to the low confidence (LC) condition, participants in the high confidence (HC) condition would agree more with positive attributes for the speaker and agree less with negative attributes for the speaker.

2. compared to the low confidence (LC) condition, participants in the high confidence (HC) condition would have more accurate recall of presented information.

These predictions align with the "confidence heuristic" suggesting that listeners are more inclined to trust and absorb information from speakers who exhibit confidence, as well as recent research suggesting that confident speakers elicit deeper cognitive processing in listeners. We also hypothesized that:

3. compared to listeners low in empathy, listeners high in either affective or cognitive empathy would evaluate the LC speaker less harshly.

This prediction is based on the idea that individuals with higher empathy are more attuned to and understanding of others' emotional states. This heightened empathy may lead them to be more forgiving and supportive, resulting in more consistent judgments of the speaker irrespective of confidence condition.

Method

Participants

Participants ($n = 57$) were currently enrolled college students recruited through an undergraduate classroom. The sample had a mean age of 18.52 ($SD = .84$). Among the participants, 32 identified as Caucasian, 17 as African American, 5 as multi-racial, and 2 as Hispanic/Latino. In terms of gender, 42 participants identified as women, 10 as men, 2 as non-binary, and 2 as unspecified "other." Regarding sexual orientation, 31 participants reported being heterosexual, 17 identified as

bisexual, 6 as homosexual, and 2 identified as unspecified "other." A single participant responded to the survey but did not provide any demographic data. This project was approved by the university's Institutional Review Board. Students were offered extra credit for participation and could select an alternate extra credit assignment if they chose not to participate. Thus, all participation was voluntary.

Materials and Procedures

To create two randomized groups, students in the original classroom were asked to count off one and two. One of the groups was then relocated to a different classroom across the hallway. The classrooms were selected because they were in the same hallway and had the same video equipment and seat configuration. The researchers confirmed that the sound levels and picture quality were equivalent in the two classrooms.

Participants were asked to make sure nothing was on their desk or in their hands. They were then directed to watch a three-minute video displayed on a large overhead screen at the front of the classroom. One of the groups viewed a peer-aged, female speaker describing the methodology and results of a research study she had conducted. She spoke confidently with no verbal crutches or hesitations, creating our High Confidence condition (HC; $n = 28$). See the appendix for the script. The other group of participants watched a video of the same female speaker following the same script; however, in this Low Confidence condition (LC; $n = 29$), the speaker inserted pauses, oral stumbles, and apologies into the script. These modifications added approximately 12 seconds to the length of the video. To ensure consistency, the two videos were recorded consecutively on the same day with the same camera, lighting, background, and outfit.

Immediately after the video ended, participants were provided with a link to an online Qualtrics survey and accessed it through their cell phones. We provided the following instructions, "We are interested in your personal opinions of the video you just watched. Please give us your honest opinion of how well the speaker matches the following descriptions." Participants encountered six positive characteristics (intelligent, confident, trustworthy, professional, impressive, interesting) and five negative characteristics (inexperienced, incompetent, unethical, hard to understand, cringy). The presentation of the characteristics was randomized such that negative and positive characteristics were not grouped together. All participants viewed the characteristics in the same randomized order. Participants were asked to rate the extent to which each one described the speaker with 1

representing "strongly disagree" and 5 representing "strongly agree." These items were inspired by the Muenster Epistemic Trustworthiness Inventory (Hendriks et al., 2015). Participants received a mean score for the positive adjectives and a mean score for the negative adjectives, with a possible range of one to five. We achieved a Cronbach's Alpha of .85 for the positive traits and .80 for the negative traits. We asked participants to make ratings of one additional descriptor (attractive) to assess whether the physical perceptions of the speaker influenced judgments of her ability.

Next, we provided an open-ended field and asked participants what the speaker did best and how the speaker could improve. Participants were then provided with the statement, "I think I could have done a better job conveying the information." To assess participants' confidence in the accuracy of their memory recall, we provided the statement, "I accurately remember what the speaker talked about in this video." They responded with their level of agreement using a 5-point scale where 1 represented "strongly disagree" and 5 represented "strongly agree."

Participants then encountered a five-item quiz to assess how much information they accurately recalled from the video. One item assessed the dependent variable in the study, while the others assessed memory of the study's hypotheses and results. Participants could earn a score from zero to five, with a higher score indicating more accurate recall of the video's content.

Participants then encountered the 20-item Perth Empathy Scale (Brett et al., 2023), which assesses the trait of empathy. This scale can be divided into two subscales. The cognitive empathy subscale measures the ability to recognize the feelings of others. A sample item was "Just by seeing or hearing someone, I know if they are feeling sad." The affective empathy subscale measures the tendency to feel the feelings of others. A sample item was, "When I see or hear someone who is sad, it makes me feel sad too." The published reliability for cognitive empathy is .92, and for affective empathy is .80 (Brett et al., 2023). We achieved a Cronbach's Alpha of .95 for cognitive empathy and .86 for affective empathy.

We asked participants four demographic questions, including age, gender, race/ethnicity, and sexual orientation. In order to further understand our participants, we also asked them to quantify how much experience they personally had conducting scientific research on a five-point scale where 1 represented "never" and 5 represented "extensive."

Results

We divided the sample into high and low cognitive empathy by using a median split with cognitive

empathy scores. We ran a 2 (condition) X 2 (cognitive empathy) MANOVA. There were no significant interactions in how listeners ascribed positive traits [$F(3,53) = .20, p = .66, ns$] or negative traits [$F(3,53) = .32, p = .57, ns$] to the speaker. There were also no interactions related to listeners' recall accuracy [$F(3,53) = .008, p = .93, ns$], confidence in recall accuracy [$F(3,53) = 2.40, p = .13, ns$], or belief they could have done a better job [$F(3,53) = .06, p = .81, ns$]. We found main effects for condition. Compared to listeners who saw the LC speaker, those who saw the HC speaker expressed higher agreement for positive traits [$F(3,53) = 44.21, p < .001$], lower agreement for negative traits [$F(3,53) = 34.22, p < .001$], and less agreement that they could have done a better job [$F(3,53) = 11.87, p = .001$]. There was no main effect for condition related to recall accuracy, confidence in recall accuracy, or attractiveness of the speaker. The comparison between the HC and LC

conditions for individual trait variables and recall accuracy can be seen in Table 1. We found no main effects for cognitive empathy on any of the variables.

We then divided the sample into high and low affective empathy by using a median split with affective empathy scores. We ran a 2(condition) X 2 (affective empathy) MANOVA. We found a significant interaction for negative traits, with listeners high in affective empathy agreeing more with negative traits for the LC speaker and agreeing less for the HC speaker [$F(3,53) = 5.04, p = .03$]. The interactions were not significant for positive traits [$F(3,53) = 1.65, p = .21, ns$], recall accuracy [$F(3,53) = 3.86, p = .06, ns$], confidence in recall accuracy [$F(3,53) = .01, p = .94, ns$], or belief they could have done a better job [$F(3,53) = .02, p = .88, ns$]. There were no main effects for affective empathy. The data related to affective and cognitive empathy can be seen in Table 2.

Table 1

Means and Standard Deviations for Speaker Attributes and Recall Accuracy in HC and LC Conditions

	High Confidence Speaker	Low Confidence Speaker
Intelligent	4.29 (.90)*	3.83 (.89)
Confident	4.50 (.58)**	2.24 (1.30)
Trustworthy	4.04 (.69)**	3.24 (1.15)
Professional	4.29 (.60)**	2.59 (1.40)
Impressive	3.93 (.66)**	2.76 (1.06)
Interesting	3.89 (.96)*	3.14 (1.33)
Attractive	2.79 (1.10)	2.72 (.96)
Inexperienced	2.18 (1.02)	3.62 (1.27)**
Incompetent	1.68 (.82)	2.59 (1.27)**
Unethical	1.89 (1.10)	2.34 (1.23)
Hard to Understand	1.61 (.88)	3.72 (1.22)**
Cringy	2.61 (1.10)	3.34 (1.29)*
Recall Accuracy	2.04 (.88)	1.86 (.74)

* $p < .05$

** $p < .01$

Table 2

Means and Standard Deviations for Dependent Variables Across Conditions and Affective Empathy

	High Confidence Speaker		Low Confidence Speaker	
	Negative Traits	Positive Traits	Negative Traits	Positive Traits
Low Affective Empathy	2.35 (.66)	3.94 (.35)	3.19 (.78)	2.93 (.42)
High Affective Empathy	1.63 (.60)	4.40 (.38)	3.42 (1.0)	2.92 (1.11)
Low Cognitive Empathy	2.13 (.81)	4.06 (.38)	3.32 (.87)	2.89 (.51)
High Cognitive Empathy	1.88 (.61)	4.27 (.46)	3.32 (.96)	2.94 (1.05)

Because we did not find categorical differences related to accuracy or confidence in recall accuracy, we examined the data file as a whole. The more confidence participants had in the accuracy of their recall, the: fewer negative traits they ascribed to the speaker [$r(57) = -.42$, $p < .001$] and the more positive traits they ascribed [$r(57) = .33$, $p = .01$]. The more confidence participants had in the accuracy of their recall, the more prior research experience they had [$r(56) = .31$, $p = .02$]. These represent weak to moderate correlations.

We compared participants who identified as men with those who identified as women using an independent samples t -test. Compared to women, men more accurately recalled information from the two videos [$t(50) = 2.22$, $p = .02$]. There was a trend for men to express more confidence in the accuracy of their recall than women, although it fell just short of significant [$t(50) = 1.57$, $p = .06$, ns]. Men reported more prior research experience than did women [$t(50) = 1.94$, $p = .03$]. We found no gender difference in the attractiveness rating for the speakers [$t(50) = -1.61$, $p = .06$, ns]; in fact, female participants rated the female speaker slightly (but not significantly) more attractive than the male participants.

We compared African American and Caucasian adults using an independent t -test. We found no differences in the accuracy of their memory recall [$t(47) = .82$, $p = .21$, ns], confidence in memory recall [$t(47) = .53$, $p = .30$, ns], or prior research experience [$t(47) = -1.36$, $p = .09$, ns]. Age did not predict any variables of interest.

Two raters categorized the qualitative comments, with an inter-rater reliability of .97. A summary of the data can be seen in Table 3. First, the raters examined the patterns for the two conditions on what the speaker did best. For the HC speaker ($n = 28$), 42.8% of compliments related to “explained research well” (e.g., communicate their data and hypothesis clearly and effectively”), and 25% related to “clear message and speaking” (e.g., “annunciate and clearly communicate points”). For the LC speaker responses ($n = 29$), 58.6% related to “explained research well” (e.g., “giving data to support her hypothesis and explaining findings”), and the next highest category was 13.7% of participants providing a criticism of the speaker (e.g., “I don’t know, everything seemed terrible.”).

The raters then categorized the qualitative comments related to what could be improved. For the HC speaker ($n = 28$), 25% of the suggestions related to “increased eye contact” (e.g., “she should make more eye contact”) followed by 17.8% for “provide more detail” (e.g., “include percentages and number within results”) and “nothing” (e.g., “no idea”). For the LC speaker responses ($n = 29$), 27.6% related to “be more confident” (e.g., “the speaker can improve by being a bit more confident in what they are talking about”), and the next highest category was 20.7% of participants pointing out one of the low confidence behaviors (e.g., “stuttering, stopping to say I’m [sic] sorry, hand gestures”). This qualitative data serves as a validity check, providing evidence that participants differentiated between the high and low confidence conditions.

Table 3

Percentage of Participants Providing Responses in Each Rater-Created Category of Qualitative Responses

	High Confidence Speaker	Low Confidence Speaker
Did Best	Explained research well (42.9%) Clear message/speech (25.5%) Confident (10.7%) Voice tone/pace (10.7%) Concise/Focused (7.1%) Eye contact (3.7%)	Explained research well (58.6%) Nothing/provided criticism (13.8%) Corrected mistakes (10.3%) Friendly face/smile (10.3%) Loud speaking tone (6.9%)
Could Improve	Increase eye contact (25%) Nothing (17.9%) Provide more detail (17.9%) Be more engaging (14.3%) Talk slower (10.7%) Relax/ be less stiff (10.7%) Be less confident (3.5%)	Be more confident (27.6%) Avoid verbal crutches (20.7%) Know information better (17.2%) Be clearer (17.2%) Relax/ calm down (13.8%) Increase eye contact (3.4%)

Discussion

We hypothesized that compared to the low confidence condition, participants in the high confidence condition would agree more with positive attributes for the speaker and agree less with negative attributes for the speaker. Our data supported these hypotheses. As predicted, the high confidence speaker received significantly higher positive ratings and lower negative ratings. The qualitative comments from participants matched this pattern. The favorable impressions of the high confidence speaker match previous research documenting that listeners attribute a wide range of positive attributes to confident, versus uncertain, speakers (e.g., Birch et al., 2020; Jiang et al., 2020; Pozzi & Mazzarella, 2023).

We also hypothesized that compared to the low confidence condition, participants in the high confidence condition would have more accurate recall of presented information. This hypothesis was not supported. The confidence of the speaker did not predict listeners' recall accuracy or confidence in recall accuracy. Previous researchers have shown that confident speakers are more persuasive (e.g., Guyer et al., 2021); however, our speaker was not attempting to persuade; instead, she was attempting to inform. Being more persuasive might differ from being a better teacher. Future researchers may want to explore this idea further. An alternate explanation for speaker confidence not predicting recall is that our participants had low recall overall. Their inability to accurately respond to questions about the speech's content may reflect low motivation or lack of interest. These participants were young undergraduate students listening to a talk about research. They may not have seen the personal relevance of the topic. In support of this argument, the participants reported low levels of personal research experience. Some of the qualitative comments indicated that the audience felt the speech could have been more engaging (i.e., less boring).

In terms of recall accuracy and recall confidence, it is worth noting that listener characteristics were also not highly predictive. Listeners' empathy level, age, and race were unrelated to their recall. Compared to women, men had more accurate recall; however, men also reported more research experience than women. Thus, men's better recall may reflect their prior experience with the topic more so than their gender. (It is worth noting that perceptions of speaker attractiveness do not explain this gender difference.) Accuracy did not predict how listeners viewed the speaker, but listeners who had more confidence in their accuracy evaluated the speakers more positively. This finding might reflect affect infusion, where students who felt they personally learned

something experienced positive emotions that they then generalized to the speaker.

We also hypothesized that greater cognitive empathy, the ability to understand others' feelings, would result in listeners' evaluating the low confidence speaker less harshly. This prediction was not supported. Instead, we found that listener's cognitive empathy did not influence their attribution of positive or negative traits to either speaker; the salient factor was the confidence of the speaker. Regardless of cognitive empathy level, the high confidence speaker was viewed more favorably. A deeper intellectual understanding of how the speaker was feeling was not strong enough to override the impressions created by the speech itself. The fact that listeners' cognitive empathy did not predict perceptions of the speakers matches previous researchers' argument that individual listener characteristics exert minimal influence (Brown et al., 2021b; Hancock et al., 2023; Marholz et al., 2018; Welch et al., 2021).

In contrast to cognitive empathy and failing to support our hypothesis, listeners who were higher in affective empathy, the ability to feel what others' feel, judged the low confidence speaker most negatively and the high confidence speaker least negatively. One explanation might be that personally relating to (i.e., taking on) the feeling of low confidence elicited unhappy feelings that led to harsher ratings, while personally relating to the feeling of high confidence elicited happy feelings that led to less harsh ratings. Another possibility is that affective empathy predicts emotion dysregulation (Thompson et al., 2022). The listeners with these more rapidly fluctuating emotions may have been more likely to use extreme negative ratings. In support of this argument, affective empathy impacted listeners' responses to negative attributes about the speakers but not their responses to positive attributes of the speakers. Again, it is noteworthy, and consistent with previous research, that listeners' characteristics exerted minimal influence on perceptions of the speakers.

To enhance the generalizability of our findings, future researchers could investigate whether these outcomes hold true for male versus female speakers or across a variety of speakers who differ in age, race, and cultural background. Additionally, a more nuanced understanding of variables that impact listeners' perceptions of speakers could be obtained by extending our protocol to examine the effect of nonverbal cues, including gestures, facial expressions, and body language. Future researchers might also examine how listeners' personal interest in the topic interacts with the studied variables. Jablon-Roberts and McCracken (2022) reported that undergraduate students have

overwhelmingly favorable impressions of guest speakers in the college classroom, so replicating this type of study in person rather than via video may render different results. Finally, future research could investigate the effects of speaker confidence in different contexts, such as work presentations, political speeches, and informal conversations, using a larger, more diverse sample.

In sum, we found that a confident female speaker was viewed more positively than an uncertain female speaker; however, confidence did not impact how much listeners recalled. Listeners' empathy did not predict their perceptions with one exception. Listeners higher in affective empathy evaluated the low confidence speaker most negatively and the high confidence speaker least negatively. Age, race, gender, and ethnicity did not emerge as highly predictive listener characteristics. These findings suggest that the confidence of female speakers may be even more influential than listener characteristics.

References

- Aadam, B., Poon, A. W. C. & Fernandez, E. (2024). Listening in mental health clinical practice. *The British Journal of Social Work*, 54, 246-266. <https://doi.org/10.1093/bjsw/bcad193>
- Birch, S. A., Severson, R. L., & Baimel, A. (2020). Children's understanding of when a person's confidence and hesitancy is a cue to their credibility. *Plos One*, 15, e0227026. <https://doi.org/10.1371/journal.pone.0227026>
- Brett, J. D., Becerra, R., Maybery, M. T., & Preece, D. A. (2023). The psychometric assessment of empathy: Development and validation of the Perth Empathy Scale. *Assessment*, 30(4), 1140-1156. <https://doi.org/10.1177/10731911221086987>
- Brown, C. L., West, T. V., Sanchez, A. H., & Mendes, W. B. (2021a). Emotional empathy in the social regulation of distress: A dyadic approach. *Personality and Social Psychology Bulletin*, 47, 1004-1019. <https://doi.org/10.1177/0146167220953987>
- Brown, K. M., Dahl, K. L., Cler, G. J. & Stepp, C. E. (2021b). Listener age and gender diversity: Effects on voice-based perception of gender. *Journal of Voice: Official Journal of the Voice Foundation*, 35(5), 739-745. <https://doi.org/10.1016/j.jvoice.2020.02.004>
- Carolus, A., Wienrich, C., Törke, A., Friedel, T., Schwietering, C. & Sperzel, M. (2021) 'Alexa, I feel for you!' Observers' empathetic reactions towards a conversational agent. *Frontiers in Computer Science*, 3, e682982. <https://doi.org/10.3389/fcomp.2021.682982>
- DePaola, M., Lombardo, R., Pupo, V. & Scappa, V. (2021). Do women shy away from public speaking? A field experiment. *Labour Economics*, 70, 102001. <https://doi.org/10.1016/j.labeco.2021.102001>
- Guyer, J. J., Briñol, P., Vaughan-Johnston, T. I., Fabrigar, L. R., Moreno, L., & Petty, R. E. (2021). Paralinguistic features communicated through voice can affect appraisals of confidence and evaluative judgments. *Journal of Nonverbal Behavior*, 45(4):479-504. <https://doi.org/10.1007/s10919-021-00374-2>
- Hancock, A. B., Hao, G., Ni, A., Liu, H. & Johnson, L.W. (2023). Gender attributions by cisgender and gender diverse listeners rating vowels, reading, and monologues. *Journal of Voice: Official Journal of the Voice Foundation*. Electronic publication. <https://doi.org/10.1016/j.jvoice.2023.09.011>
- Hancock, A. & Pool, S. F. (2017). Influence of listener characteristics on perceptions of sex and gender. *Journal of Language and Social Psychology*, 36(5). <https://doi.org/10.1177/0261927X17704460>
- Hendriks, F., Kienhues, D., & Bromme, R. (2015). Measuring laypeople's trust in experts in a digital age: The Muenster Epistemic Trustworthiness Inventory (METI). *PLOS One*, 10(10), e0139309. <https://doi.org/10.1371/journal.pone.0139309>
- Jablon-Roberts, S. & McCracken, A. (2022). Undergraduate student perceptions of industry guest speakers in the college classroom. *Journal of the Scholarship of Teaching and Learning*, 22, 76-88. <https://doi.org/10.14434/josotl.v22i3.32317>
- Jankowiak-Siuda, K., Rymarczyk, K., & Grabowska, A. (2011) How we empathize with others: A neurobiological perspective. *Medical Science Monitor*, 17, RA18-24. <https://doi.org/10.12659/msm.881324>
- Jiang, X., Gossack-Keenan, K. & Pell, M. (2020). To believe or not believe: How voice and accent information in speech alter listener impressions of trust. *The Quarterly journal of Experimental Psychology*. <https://doi.org/10.1177/1747021819865833>
- Jiang, X. & Pell, M. D. (2015). On how the brain decodes vocal cues about speaker confidence. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior*, 66, 9-34. <https://doi.org/10.1016/j.cortex.2015.02.002>
- Junaid, S.M., Mahboob, U., Nadeem, N., Lqbal, N., Khuwaja, A.M. & Bangash, M. (2022). How to teach 'listening for learning' in a clinical context? *Journal of Ayub Medical College*, 34(2), 392-393.
- Mahrholz, G., Belin, P. & McAleer, P. (2018). Judgements of a speaker's personality are correlated across

- different content and stimulus types. *PLoS ONE*, 13(10): e0204991.
<https://doi.org/10.1371/journal.pone.0204991>
- Mauchand, M., Armony, J. L. & Pell, M. D. (2023). The vocal side of empathy: Neural correlates of pain perception in spoken complaints. *Social, Cognitive and Affective Neuroscience*, 19, nsado75. <https://doi.org/10.1093/scan/nsado75>
- Mazzarella, D., Reinecke, R., Noveck, I., & Mercier, H. (2018). Saying, presupposing and implicating: How pragmatics modulates commitment. *Journal of Pragmatics*, 133, 15–27.
<https://doi.org/10.1016/j.pragma.2018.05.009>
- Mori, Y & Pell, M. D. (2019). The look of (un)confidence: Visual markers for inferring speaker confidence in speech. *Frontiers in Communication*, 4.
<https://doi.org/10.3389/fcomm.2019.00063>
- Pozzi, M., & Mazzarella, D. (2023). Speaker trustworthiness: Shall confidence match evidence? *Philosophical Psychology*, 37(1), 102–125.
<https://doi.org/10.1080/09515089.2023.2193220>
- Roberts, F., & Norris, A. (2016). Gendered expectations for “agreeableness” in response to requests and opinions. *Communication Research Reports*, 33, 16–23.
<https://doi.org/10.1080/08824096.2015.1117437>
- Roche, J. M., Asaro, K., Morris, B. J. & Morgan, S. D. (2023). Gender stereotypes and social perception of vocal confidence is mitigated by salience of socio-indexical cues to gender. *Frontiers in Psychology*, 14.
<https://doi.org/10.3389/fpsyg.2023.1125164>
- Roche, J. M., Morgan, S. D., Brannick, S. N., & Bryndel, K. (2019). Acoustic correlates of female confidence: A production and comprehension study. *Journal of the Acoustical Society of America*, 145, 3410–3426.
<https://doi.org/10.1121/1.5109661>
- Roche, J. M., Morgan, S. D., & Fisk, S. (2022). Gender stereotypes drive perceptual differences of vocal confidence. *Journal of the Acoustical Society of America*, 151, 3031–3042.
<https://doi.org/10.1121/10.0010382>
- Taylor, B., Wheeler-Hegland, K. & Logan, K. J. (2022). Impact of vocal fry and speaker gender on listener perceptions of speaker personal attributes. *Journal of Voice: Official Journal of the Voice Foundation*. S0892-1997(22)00292-2.
<https://doi.org/10.1016/j.jvoice.2022.09.018>
- Tenney, E. R., Meikle, N. L., Hunsaker, D., Moore, D. A., & Anderson, C. (2019). Is overconfidence a social liability? The effect of verbal versus nonverbal expressions of confidence. *Journal of Personality and Social Psychology*, 116(3), 396–415.
<https://doi.org/10.1037/pspi0000150>
- Thomas J. P., & McFadyen R. G. (1995). The confidence heuristic: A game-theoretic analysis. *Journal of Economic Psychology*, 16, 97–113.
[https://doi.org/10.1016/0167-4870\(94\)00032-6](https://doi.org/10.1016/0167-4870(94)00032-6)
- Thompson, N. M., van Reekum, C. M. & Chakrabarti, B. (2022). Cognitive and affective empathy relate differentially to emotion regulation. *Affective Science*, 3, 118–134.
<https://doi.org/10.1007/s42761-021-00062-w>
- Welch, B., van Mersbergen, M. R. & Helou, L. B. (2021). Listeners’ perceptions of speaker personality traits based on speech. *Journal of Speech, Language, and Hearing Research*, 64, 4762–4771. https://doi.org/10.1044/2021_JSLHR-20-00582
- Woolridge, L. R., Leach, A.-M., Blake, C., & Elliott, E. (2024). Do accents speak louder than words? Perceptions of linguistic speech characteristics on deception detection. *Journal of Language and Social Psychology*, 43(2), 195–223.
<https://doi.org/10.1177/0261927X231209428>

Appendix

Script

I'm going to tell you about a research project that I recently did examining how college students perceived their high school teachers versus how they perceived their college teachers. The way that students perceive their teachers influences how well they get along with them. This is important because the more students get along with their teachers, the more they learn, the better grades they get, and the better behaved they are in class. Most of the research has focused on K-12 classrooms. My study added to what was previously known because I focused on college students.

I hypothesized that college students would report more positive interactions with their college teachers compared to their high school teachers. I also decided to look at academic entitlement. Students high in academic entitlement think that learning is the responsibility of the teacher, not their responsibility. They expect a lot of assistance and blame others when they fail.

I hypothesized that students who felt a lot of academic entitlement would have more negative interactions with all of their teachers.

I used survey methodology. First, I asked participants to think only about their high school teachers. I asked them to tell me what their teachers were like, how they interacted with them, and how they behaved in their classrooms. Then, I asked participants to think about their college teachers and answer the same questions.

When I analyzed the data, I found that overall, students viewed teachers positively; however, students attributed more positive traits to and were more willing to communicate with college teachers. College teachers were viewed as more professional and more sensitive to diversity. Students also said they lied and cheated more in high school than they did in college.

These results probably reflect students' different maturity levels and goals. It also matches my hypothesis!

My other prediction was that an attitude of entitlement would interfere with student-teacher relationships. This hypothesis was also correct! The more entitled students were, the less positive things they had to say about their high school and college teachers. They were also more likely to cheat and to complain about teachers to other teachers.

Last, I examined demographic characteristics. White students reported feeling more comfortable around high school teachers, while Black students were more likely to talk with their college teachers. This might have something to do with those earlier results I told you about, where college teachers were considered to be more sensitive to diversity. Men and women said the same thing about their high school teachers. However, in college, men were more likely to talk to their teachers, especially if they had something to complain about. I compared students who reported being heterosexual with those in other categories. Non-heterosexual students described a more informal interaction style with their high school and college teachers... for example, joking around and gossiping.

Thank you for allowing me to share my research findings. I hope you found it interesting and will maybe feel inspired to do research of your own...if you haven't already.

ETHNIC DIFFERENCES IN SEEKING HELP AFTER EXPOSURE TO DOMESTIC VIOLENCE

ANDREA SANTIBANEZ AND CARMEN FARRELL

UNIVERSITY OF SOUTH CAROLINA BEAUFORT

Abstract – Domestic violence (DV) is highly prevalent worldwide and the United States is no exception. DV is something that can happen in ethnic minority communities and people who need help may not seek it. The purpose of the present study was to better understand ethnic differences in rates of DV and rates of seeking help after exposure to DV in the United States. Additionally, citizenship status was explored as another factor that might impact these rates. The national dataset used for this project was the National Crime Victimization Survey from 2021 with 8,982 participants. The researchers hypothesized that non-minority individuals would have higher rates of experiencing DV in this dataset, but they would also have higher rates of seeking professional help after exposure. They also hypothesized that non-citizens would seek less professional help than citizens. The results showed that there were no statistically significant differences across ethnicity or citizenship in rates of DV or in rates of seeking professional help after exposure, which does not support the hypotheses. However, these results point to how significant underreporting can be, even in a national dataset with a plethora of resources and funding. Implications for future research in this area are discussed.

Keywords: domestic violence, help-seeking, barriers, and citizenship

Domestic violence (DV) is a critical issue in the United States and around the world. On average, every year in the United States, there are over 10 million people who experience DV (Huecker et al., 2023). Therefore, it is important to know how we can best help victims and prevent future victimization from occurring. Additionally, the U.S. is made up of many ethnic groups, yet not all ethnic groups may have equal access to care such as financial assistance, medical help, legal assistance, living arrangements, and counseling services after experiencing DV. For example, research on DV has been conducted and researchers have found that there is minimal effort in helping victims who are minorities, as well as insignificant progress in this area over time (Reina & Lohman, 2015). For this reason, it is important to analyze whether ethnic differences exist in experiencing DV and in help-seeking behaviors after exposure to know what groups need more support in the U.S.

Help-Seeking Behaviors – Formal and Informal

Help-seeking behaviors for DV in the extant literature are usually divided into formal and informal help-seeking. Formal help-seeking behaviors may include help from doctors, mental health professionals, and law enforcement (Satyen et al., 2019). Seeking help formally

is associated with lower levels of psychological distress (Cuevas et al., 2014). In one study, it was reported that Caucasians sought help by seeking mental health counselors or psychologists (Satyen et al., 2019). Compared to European Americans, Latin Americans are less likely to seek professional mental health services (Gul et al., 2021). It is difficult to say that one group is more or less likely to use a type of formal help-seeking over the other when results are inconsistent across studies. However, research shows a trend of Caucasian victims seeking more help from mental health professionals than other ethnic groups (Gul et al., 2021; Satyen et al., 2019).

In addition to formal help seeking, there is also informal help-seeking which typically includes help from family and friends (Satyen et al., 2019). Previous research has found that when seeking help, cultural differences are evident. For example, Latina victims have been found to seek more informal help than formal (Cuevas et al., 2014). Compared to Caucasian women, it has been reported that Hispanic women seek less help from family and friends (Satyen et al., 2019). The extant literature has yet to confirm consistent differences in help-seeking behavior after experiencing DV (Satyen et al., 2019).

Barriers to Help-Seeking – Citizenship and Language

The most prominent barriers for women seeking help after exposure to DV were discrimination, documentation status, and financial challenges (Reina et al., 2014; Reina & Lohman, 2015). Research has reflected that some women may be at greater risk of experiencing DV when their abuser uses their citizenship against them and the women fear being reported to Homeland Security (Messing et al., 2022). Participants who were undocumented described that they were afraid of being deported due to their documentation status or the uncertainty of their residency (Alvarez et al., 2016; Reina et al., 2014; Salcido & Adelman, 2004). Not only was the fear within the victims themselves, but they also discussed that their partner or abuser threatened them with deportation.

Not only is citizenship status one of the biggest barriers to seeking help after DV, but researchers have found that language plays a significant role as well. Victims have noted that the representation of employees from organizations designed to help victims of DV is not proportionate to the number of minority victims there are and employees who are said to speak the victims' language are not always well-trained interpreters (Messing et al., 2022). It is critical to have both language and culturally competent employees in DV services, hotlines, and other organizations to best assist victims who do not fit Eurocentric or English norms.

Barriers to Help-Seeking – Cultural Beliefs

Cultural and individual beliefs may also interfere with or influence a victim's help-seeking behaviors. In the Latino community, gender norms are common such as *Machismo* which involves certain characteristics like aggressiveness, being the family provider, having control over wife and kids, and no display of emotions among men (Gul et al., 2021). In the Asian community, some Asian women may have deeper struggles with seeking help because there is often a belief that physical abuse from a man is acceptable (Satyen et al., 2019). In Reina and colleagues' study (2014), more than half of the women reported that their cultural norms in marriage influenced whether they sought help based on what their culture reflected as acceptable. Additionally, in a country like the United States that is a melting pot of ethnic backgrounds, differences also exist in assimilation to U.S. culture versus the culture of their ethnic background. More specifically, if one is more acculturated to U.S. society, researchers have found that they may be more likely to seek formal help than Latinos who keep their cultural values strong (Sabina et al., 2012). Therefore, it is important to understand how cultural beliefs may

influence whether victims seek help and what ways are best to support victims with different beliefs.

Barriers to Help-Seeking – Other Factors

Other factors also exist as barriers to seeking help for DV. Immigrant victims in the U.S. may face issues that others do not have to face, such as increased difficulty in obtaining a place to stay and finding a job because of their documentation status, which may increase their chances of returning to their abuser (Wachter et al., 2021). Some immigrants also arrive exposed to poverty or lower educational standards (Reina & Lohman, 2015). Other victims feel shame and guilt in reporting their abuse due to societal stigma surrounding DV and its victims (Reina et al., 2014). Therefore, immigrants have numerous additional factors that may make their path to seeking help more difficult. Other barriers to help-seeking found in the literature include the fear of violence persisting or worsening, the environment their children would be in (e.g. separation from the parent who is the victim), the misinformation on the help available, as well as the guilt and shame that some victims may feel for communicating their victimization (Reina et al., 2014; Satyen et al., 2019). Experiencing DV is such a complex and multifaceted issue that it is important to take these factors into consideration and try to understand what these differences look like in a national sample.

Current Study

For the present study, the researchers utilized a national dataset to examine ethnic and citizenship status differences in help-seeking behaviors after exposure to DV. It was hypothesized that non-minority individuals will have higher rates of experiencing DV in this dataset, but also have higher rates of seeking professional help after exposure, due to access to resources. Non-minority individuals are those who report themselves to be Caucasian. Citizenship status was also hypothesized to impact how often people seek help, such that non-citizens would seek help less than citizens. People from minority backgrounds likely have problems with citizenship status as well as language barriers that non-minorities in the United States do not have to consider. The researchers believed this dataset would reflect higher DV rates for non-minorities, but this may be explained through minority rates being underreported. With the usage of a national dataset, the goal of this research was to take the first step in understanding what the numbers look like on a national scale, in hopes of advocating for support of minoritized groups and their extra need to get the proper care through organizational support and assistance from the government and authorities.

Method

Materials

The national dataset used for this project was the National Crime Victimization Survey (United States, Bureau of Justice Statistics, 2022) from 2021, which was the most recent year that is accessible via the Inter-university Consortium for Political and Social Research. There were 8,982 participants from this national sample. One hundred fifty-seven participants from the National Crime Victimization Survey (NCVS) reported experiencing DV, as defined in the results section below. The NCVS was conducted through the U.S. Census Bureau and was a population survey. Once participant consent was obtained, the first interviews were done at home. Participants then had the option to do follow-up interviews in person or over the phone. The dataset covered why the crime was or was not reported as well as asked questions about help they sought after experiencing the crime, and how the crime impacted their life afterwards (United States Bureau of Justice Statistics, 2022).

Participants

The participants' ages for the present study ranged from 18 to 82 with the mean age being 41. There were 28 men and 129 women of our subsample who experienced DV. The participants' household income ranged from less than \$5,000 to \$200,000 or more annually. Of the participants, 151 were citizens of the U.S.

Table 1
Demographic Characteristics

Demographic	N	%
Gender		
Female	129	82.16
Male	28	17.83
Marital status		
Married	38	24.2
Widowed	7	4.5
Divorced	23	14.6
Separated	9	5.7
Never Married	79	50.3
Ethnicity		
White	95	60.5
Black	20	12.7
Asian	4	2.5
Hispanic	28	17.8
Native	1	.6
Mixed	9	5.7
Educational Attainment		
High school degree or less	30	19.35
Some college or an associate's degree	62	40
Bachelor's degree or higher	57	36.77

(including being naturalized and living in the U.S. territories) while 5 were not. See Table 1 for a breakdown of all demographics.

Results

The first relationship analyzed in the study was between White participants and minorities and their victimization by DV out of the total national sample of 8,982. A chi-square test of independence was performed to look at whether there was a link between being a White person or a minority and experiencing DV. The relation between these variables was nonsignificant, $X^2(1, N = 8982) = 1.39, p = .239$. This result tells us that there was not a significant difference in minority status and rates of experiencing DV.

Prior to the rest of the analyses, a composite variable for DV was created. For that variable, aspects of DV included from the dataset were harassed/abusive language, sexual contact with and without force, threat of rape or rape, and sexual/physical abuse. Individuals who experienced any of these types of DV were coded with a score of 1 and only their data were utilized for all the following analyses. There were 157 participants who fit these criteria.

In order to compare the help-seeking behaviors of participants exposed to DV, chi-square tests of independence were utilized. First, ethnic differences in help-seeking behaviors were explored. Some of the ethnicities did not have enough participants to run a chi-square analysis, so only White people, Black people, and Hispanic people were used in this analysis. A chi-square test of independence was performed to examine the link between ethnicity and seeking professional help for feelings experienced as a victim of a DV crime. The relation between these variables was nonsignificant, $X^2(2, N = 96) = 5.87, p = .053$. This result tells us that there was not a significant difference in the rate of seeking help across ethnicities. See Table 2 for breakdown of counts.

Table 2
Counts for Ethnicity and Victims Seeking Professional Help

	Did you seek professional help?		
	Yes	No	Total
Ethnicity			
White	30	38	68
Black	2	12	14
Hispanic	8	6	14
Total	40	56	

From there, we collapsed across ethnic groups to compare a minority status group to White people. A chi-square test of independence was performed to look at whether there was a link between minority status and seeking professional help after experiencing DV. The relation between these variables was nonsignificant, $X^2(1, N = 102) = 2.06, p = .15$. This result tells us that there was not a significant difference in the rate of seeking help between White people and minorities.

Next, help-seeking behaviors were examined between citizens and non-citizens. A chi-square test of independence was performed to explore associations between citizenship status and seeking professional help for feelings experienced as a victim of crime. The relation between these variables was nonsignificant, $X^2(1, N = 102) = .66, p = .417$. This tells us that there was not a significant difference in citizenship status and seeking professional help for feelings experienced as a victim of crime. See Table 3 for breakdown of counts.

Table 3
Counts for Citizenship Status and Victims Seeking Professional Help

	<u>Did you seek professional help?</u>		Total
	Yes	No	
<u>Citizenship Status</u>			
Citizen	40	59	99
Non-Citizen	0	3	3
Total	40	62	

Discussion

The researchers hypothesized that non-minority individuals would have higher rates of experiencing DV in the dataset and they would also have higher rates of help-seeking behaviors after exposure to DV. However, when comparing White people and minorities, both groups had fewer people experiencing DV than those who did not experience it; thus, there was not a significant difference in group rates of experiencing DV. Next, comparing White people and minorities in seeking help, for both groups, fewer people sought professional help than those who did not in each group, with no significant difference found. Even when broken down by the type of help sought, none of the differences were significant.

These results contribute to existing literature on immigrant and minority women and their help-seeking behaviors after exposure to DV, which reflect some inconsistencies. For example, Gul and colleagues (2021)

reported that some studies reflected Latin victims with more cases of DV than Caucasians, while other studies showed similar or fewer cases across Latinas and Caucasians. However, what most articles reflect is that immigrants or minority individuals face many barriers to obtaining optimal assistance after exposure to DV such as language barriers and fear of deportation (Alvarez & Fedock, 2018; Gul et al., 2021; Messing et al., 2022; Ragavan et al., 2020; Reina et al., 2014; Reina & Lohman, 2015; Rodriguez et al., 2018; Satyen et al., 2019). Despite nonsignificant results, the present study sheds critical light on the issue of reporting and the need for further sensitivity in exploration and surveys of DV exposure, as well as help-seeking behaviors afterwards.

The researchers also hypothesized that non-citizens would seek professional help less often than citizens in this national data set. A chi-square analysis was conducted based on this hypothesis to look at whether there was a relationship between citizenship status and seeking professional help. Again, these results were insignificant. However, it is important to note that none of the non-citizens sought professional help. There were only 5 non-citizens in our sample, but it would be interesting to see if there could be a significant relationship between citizenship status and help-seeking with a larger sample.

Limitations

The present study examined important differences but did not find any significance, which may be due to several limitations. Out of the 8,982 participants available from the National Crime Victimization Survey, only 157 participants qualified to have experienced DV according to how we defined it. Even when utilizing a national dataset with a plethora of resources for data collection including methodology that provided translators, the rates of DV in this sample are significantly lower than the national statistics reported by the government and other advocacy agencies. Therefore, when comparing ethnic groups, some groups had a small number of representatives. For example, White people had 68, Black and Hispanic people each had 14, and our remaining groups such as Asian, Mixed, and Native people had 3 participants or less.

Part of the reason the current sample is small may be due to underreporting and/or social desirability. Underreporting is something that may happen in any study that involves self-reports. It has been found that DV may be underreported when based on a population survey (World Health Organization, 2010; Satyen et al., 2019). The NCVS is a population survey. Additionally, the experience of being a victim of DV is challenging to discuss, and the prevalence rates may be hidden or not

reflect the number of actual victims (World Health Organization, 2010). Researchers have shown that some victims feel guilt, shame, and/or fear, are misinformed, or may feel that their violence is normal (Reina et al., 2014; Satyen et al., 2019). Similarly, the sample was limited in terms of citizenship, such that there was not a comparable number of citizens ($N = 151$) in comparison to non-citizens ($N = 5$). Victims of DV may not trust law enforcement and they may also fear that contacting them may result in deportation (Rodriguez et al., 2018).

Another influencing factor is cultural differences. As the survey included people of different ethnicities, these folks may perceive and understand DV differently. As discussed in the introduction section, gender norms such as *Machismo* are common in the Latin community which may lead to folks being more accepting of aggressive behavior (Gul et al., 2021). Similarly, some Asian women may believe that physical abuse is acceptable in a relationship (Satyen et al., 2019). This is why it is difficult to know how many victims of DV there are as definitions vary.

Not only may understandings of DV vary by ethnicity but they may also vary based on geographical location within the United States. It has been discussed that DV in adolescents may be more likely to occur in the South than in other regions of the U.S. (McDonell et al., 2010). In this same region, where honor beliefs are present, individuals are more likely to view a woman's infidelity as negative while viewing a man's aggressive behavior toward his wife as positive (Gul et al., 2021). Therefore, individuals in the South may perceive DV differently than people from other locations because of the cultural views surrounding marital dynamics.

Future Directions

Based on the results of this study, it is evident that future research must be done. Many questions remain unanswered. As mentioned above, we primarily compared citizens to non-citizens and ethnicities in general. We did not compare these groups by separating women and men which could be explored in future research. We also did not compare the rates of victimization as well as help-seeking by state as we simply looked at the national statistics.

Additionally, language was not an issue in this sample because the administrators made sure the participants had all the support they needed if they did not speak English to complete the study. In future studies, it would be good to also compare English speakers to non-English speakers because it is evident that language may impede an immigrant or Latina victim's ability to seek help (Reina et al., 2014; Reina & Lohman, 2015). Also, we only analyzed professional help-

seeking and not informal help-seeking such as seeking support from friends or family. Most studies reflect that seeking informal help after DV exposure is most common among minorities, Hispanic people, or Immigrants, while professional/formal help-seeking such as seeking medical help is more common among non-minorities (Cuevas et al., 2014; Gul et al., 2021; Sabina et al., 2014; Satyen et al., 2019).

Conclusion

Although the results were nonsignificant, the present study sheds critical light on limitations that exist in researching DV exposure and help-seeking behaviors, even when utilizing a national sample. Satyen and colleagues (2019) discussed that there are inconsistent findings across cultures regarding help-seeking behaviors after experiencing DV, and the present study supported that with insignificant findings. Therefore, research needs to continue to expand in this area and create environments where victims feel safe reporting their experiences and seeking the help they may so desperately need.

References

- Alvarez, C., & Fedock, G. (2018). Addressing intimate partner violence with Latina women: A call for research. *Trauma, Violence, & Abuse*, 19(4), 488-493. <https://doi.org/10.1177/1524838016669508>
- Alvarez C.P., Davidson P.M., Fleming C., & Glass N.E (2016) Elements of effective interventions for addressing intimate partner violence in Latina women: A systematic review. *PLoS ONE*. 11(8): e0160518. <https://doi.org/10.1371/journal.pone.0160518>
- Cuevas, C. A., Bell, K. A., & Sabina, C. (2014). Victimization, psychological distress, and help-seeking: Disentangling the relationship for Latina victims. *Psychology of Violence*, 4(2), 196-209. <https://doi.org/10.1037/a0035819>
- Gul, P., Cross, S. E., & Uskul, A. K. (2021). Implications of culture of honor theory and research for practitioners and prevention researchers. *American Psychologist*, 76(3), 502-515. <https://doi.org/10.1037/amp0000653>
- Huecker M.R., King K.C., & Jordan G.A., & Smock, W. (2023) Domestic violence. In *StatPearls*. <https://www.ncbi.nlm.nih.gov/books/NBK499891/>
- McDonell, J., Ott, J., & Mitchell, M. (2010). Predicting dating violence victimization and perpetration among middle and high school students in a rural southern community. *Children and Youth Services Review*, 32(10), 1458-1463. <https://doi.org/10.1016/j.childyouth.2010.07.001>.

- Messing, J., Wachter, K., AbiNader, M., Ward-Lasher, A., Njie-Carr, V., Sabri, B., Murray, S., Noor-Oshiro, A., & Campbell, J. (2022). "We have to build trust": Intimate partner violence risk assessment with immigrant and refugee survivors. *Social Work Research*, 46(1), 53–64. <https://doi.org/10.1093/swr/svab030>
- Ragavan, M. I., Thomas, K. A., Fulambarker, A., Zaricor, J., Goodman, L. A., & Bair-Merritt, M. H. (2020). Exploring the needs and lived experiences of racial and ethnic minority domestic violence survivors through community-based participatory research: A systematic review. *Trauma, Violence, & Abuse*, 21(5), 946–963. <https://doi.org/10.1177/1524838018813204>
- Reina, A.S., & Lohman, B.J. (2015) Barriers preventing Latina immigrants from seeking advocacy services for domestic violence victims: A qualitative analysis. *J Fam Viol* 30, 479–488 <https://doi.org/10.1007/s10896-015-9696-8>
- Reina, A. S., Lohman, B. J., & Maldonado, M. M. (2014). "He said they'd deport me": Factors influencing domestic violence help-seeking practices among Latina immigrants. *Journal of Interpersonal Violence*, 29(4), 593–615. <https://doi.org/10.1177/0886260513505214>
- Rodriguez, R., Macias, R. L., Perez-Garcia, R., Landeros, G., & Martinez, A. (2018). Action research at the intersection of structural and family violence in an immigrant Latino community: A youth-led study. *Journal of Family Violence*, 33(8), 587–596. <https://doi.org/10.1007/s10896-018-9990-3>
- Sabina, C., Cuevas, C. A., & Schally, J. L. (2012). Help-seeking in a national sample of victimized Latino women: the influence of victimization types. *Journal of interpersonal violence*, 27(1), 40–61. <https://doi.org/10.1177/0886260511416460>
- Sabina, C., Cuevas, C. A., & Lannen, E. (2014). The likelihood of Latino women to seek help in response to interpersonal victimization: An examination of individual, interpersonal and sociocultural influences. *Psychosocial Intervention*, 23(2), 95–103. <https://doi.org/10.1016/j.psi.2014.07.005>
- Salcido, O. & Adelman, M. (2004). "He has me tied with the blessed and damned papers": undocumented-immigrant battered women in Phoenix, Arizona. *Human Organization*. 63(2): 162–172. <https://doi.org/10.17730/humo.63.2.v5w7812lpxextpbw>
- Satyen, L., Rogic, A.C. & Supol, M. (2019). Intimate partner violence and help-seeking behavior: A systematic review of cross-cultural differences. *J Immigrant Minority Health* 21, 879–892 <https://doi.org/10.1007/s10903-018-0803-9>
- United States Bureau of Justice Statistics. (2022). *National crime victimization survey*, (Version 20) [Data set]. Inter-university Consortium for Political and Social Research. <https://doi.org/10.3886/ICPSR38429.v1>
- Wachter, K., Cook Heffron, L., & Dalpe, J. (2021). "We weren't ready": Provider perspectives on addressing intimate partner violence among refugees and immigrants in the United States. *Journal of Family Violence*, 37(2), 235–246. <https://doi.org/10.1007/s10896-021-00285-2>
- World Health Organization. (2010). Preventing intimate partner and sexual violence against women: Taking action and generating evidence. <https://www.who.int/publications/i/item/preventing-intimate-partner-and-sexual-violence-against-women-taking-action-and-generating-evidence>

VALIDITY EVIDENCE FOR A COMPUTER-ADAPTIVE ASSESSMENT OF READING ACHIEVEMENT AT THE SECONDARY LEVEL

CATHERINE WINDHAM AND GIANCARLO A. ANSELMO

WINTHROP UNIVERSITY

Abstract – There is some evidence that benchmark assessments are valid predictors of students' performance on high-stakes assessments. However, there is less research on this relationship at the middle school level, particularly when looking at computer-adaptive benchmark assessments. As such, this study analyzed the criterion validity of the computer-adaptive Star Reading assessment (STAR-R) when given three times a year and then compared to a statewide high-stakes assessment. Further, this study also tested the divergent validity of the STAR-R by examining if the STAR-R would produce significantly different mean scores when comparing students labeled academically gifted, students in regular education, and those in special education. Participants included sixth ($N = 218$) and seventh ($N = 270$) grade students from a rural middle school in the southeastern region of the United States. Correlations were strong between the STAR-R and the high-stakes assessment across the year for both grades, suggesting that the STAR-R may hold value as a predictor for high-stakes assessment performance as early as nine months in advance. Further, the divergent validity of the Star reading assessment is supported by significant differences in performance on both the Star reading assessment and the high-stakes assessment between students in regular education and those in special education or gifted programs. Future research directions and implications of the research results are discussed.

Assessments are deeply embedded in school systems and serve a variety of purposes. Due to their broad number of purposes, they are divided into several descriptive categories. A widespread category of assessment, the benchmark assessment, is used to monitor student progress for purposes such as identifying students who may need an academic intervention and deciding how to modify instruction for teachers (Abrams et al., 2012). Benchmark assessments are also being frequently used to predict student performance on high-stakes assessments (Shapiro & Gebhardt, 2012). High-stakes assessments are assessments with significantly impactful outcomes, including statewide assessments that may dictate a school's funding or assessments with scores that determine whether a student is qualified to move to the next grade level or for acceptance into a given university (Jones & Ennes, 2018).

Computer Adaptive Testing

One form of benchmark assessment used in schools is the computer-adaptive test (CAT). CATs were developed from Alfred Binet's (1916) research on adaptive testing and were later expanded upon by David

J. Weiss as technology advanced (Weiss, 1973). CATs differ from traditional fixed-item pencil-and-paper tests in that they are administered on a computer or similar device. Unlike a fixed-item test, CATs utilize a technology that modifies the questions a test-taker sees based on how they answered previous questions. Questions get harder as the test-taker answers them correctly. Questions get easier as the test-taker answers them incorrectly.

Research has shown that this personalization provides stronger correlations to a high-stakes assessment when compared to curriculum-based measurement, another category of benchmark assessment (Shapiro & Gebhardt, 2012). This finding suggests that CATs are beneficial to school systems by providing a testing option that may have higher predictive validity than other testing methods.

Research has explored the impact of CAT in schools. In a study with students in grades three, five, seven, and nine, Martin and Lazendic (2018) found that CAT provided several positive benefits when compared to computer fixed-item testing, especially in the ninth-grade

students. The researchers report students having both increased motivation and engagement while testing.

Earlier research by Ortner et al. (2014) examined the impact of CAT versus fixed-item testing on motivation in secondary students. The researchers measured motivation by testing fear of failure, subjective probability of success, and flow, a mental state of enjoyment that comes from a match of one's abilities and the task they are completing (Csikszentmihalyi, & LeFevre, 1989). Subjective probability of success was found to be equal for both high and low-achieving students who completed the CAT, supporting the idea that CATs adequately tailor items to the test-taker. Whereas there was no difference in flow between the fixed-item and CAT groups, flow was predicted by test performance in the fixed-item group and was not predicted in the CAT group. In other words, higher-achieving students reported more flow in the fixed-item group. Additionally, even though fear of failure was higher in the CAT group, motivation was more consistent overall. Researchers suggested that these last two findings may support CAT as a fairer testing method than fixed-item testing because students may be more successful when they are more motivated, so less-motivated students could be at a disadvantage. The studies from Martin and Lazendic (2018) and Ortner et al. (2014) suggest that the use of CATs in schools might be beneficial for the students not just in how reliable or valid the test might be, but in how they could increase and provide consistent motivation.

Though there is growing research on the use of CATs in academic settings, further research needs to be done to support and expand upon the results of prior research (Shapiro & Gebhardt, 2012; Ortner et al., 2014). This study intends to add to the body of literature about CAT in academic settings.

Previous Literature on the Predictability of High-Stakes Assessment Scores

While distinct from CATs, curriculum-based measurements (CBMs) are a specific type of assessment that have a purpose similar to benchmark CATs. There is a wide body of knowledge supporting the predictive validity of traditional CBMs on high-stakes assessments in both elementary school and middle school.

For example, Whitley (2019) looked at two types of CBMs, Oral Reading Fluency (ORF) and Maze to determine how each assessment predicts performance on high-stakes assessments for math and reading. Maze consists of a grade-level reading passage from 150 to 400 words, where the first and last sentences are intact, but every seventh word is deleted from the rest of the passage. Three-word choices are provided for each blank,

and the student selects the correct word while reading silently. The measure is then timed for 3 minutes. Participants included 197 fifth- and sixth-grade students from three schools. The researcher found that the relationships between both CBMs and high-stakes assessments were positive and moderately strong ($r = .53 - .67$, depending on the assessments examined). Likewise, Espin et al. (2010) found reading aloud and Maze CBMs had a strong positive relationship ($r = .75 - .81$) with a middle school high-stakes reading assessment from 236 eighth-grade students.

While CBMs have similarities with computer-adaptive benchmark assessments, there are also distinct differences. Therefore, the predictive validity of CBMs is not directly transferable to the predictive validity of CATs. Shapiro and Gebhardt (2012) studied the differences in the predictive validity of CBMs versus CATs in mathematics using the computation (*M-COMP*) and concept/application (*M-CAP*) Math CBM probes published by AIMSweb and the STAR math CAT assessment. Researchers looked at the relationships between these assessments and a high-stakes statewide assessment. Participants included 352 first- through fourth-grade students at two elementary schools. Results from grades three and four were correlated with the high-stakes assessment scores. Correlations were low to moderate for the *M-COMP* and *MCAP* ($r =$ means of $.29$ and $.36$, respectively), and moderate for Star Math (averaging $r = .6$).

A more recent study supported the predictive validity of Star Math. Turner et al. (2022) found that the Star Math assessment had a strong positive relationship ($r = .76 - .79$) with a high-stakes statewide assessment. This study examined 669 third- through fifth-graders from one elementary school.

Examining literacy instead of math, Clemens et al. (2015) focused on the predictive validity of the Star Early Literacy (SEL) assessment for students in kindergarten. The assessment was given at three points (fall, winter, and spring) in students' kindergarten year. Researchers examined students' reading abilities at the end of their kindergarten year and again at the end of their first-grade year through a number of measures, including a reading CBM. Moderate correlations were found between the SEL and the measures used to determine reading abilities at the end of students' kindergarten and first-grade years ($r = .39 - .61$ and $r = .27 - .60$, respectively). The researchers ultimately found that the SEL scores held the most predictive validity for both kindergarten and first grade when combined with the paper-based assessments also administered to measure reading abilities.

Research Questions

A strong body of literature has supported the predictive validity of CBMs for high-stakes assessments in elementary and middle school, especially in reading. Some research exists regarding the predictive validity of computer-adaptive benchmark tests at the elementary school level. However, little research has been done regarding the predictive validity of computer-adaptive benchmark assessments at the middle school level, leaving room for questions concerning how educators can use CAT scores to best help middle school students. As far as we know, no researchers have examined the correlations between CAT reading scores and high-stakes assessment scores, particularly at multiple points in a year. Further, there is little mention of the divergent validity of computer-adaptive assessments amongst groups of students with different learning needs. Also, current research is missing on the stability of CAT reading scores over a summer break from school. This research seeks to fill these gaps by asking three questions:

- 1) How does the criterion validity of the Star Reading assessment (STAR-R) for the North Carolina End-of-Grade Reading assessment (NC-EOG-R) vary for sixth- and seventh-grade students when administered at different points throughout the academic year (fall, winter, and spring)?
- 2) Do STAR-R scores differentiate between regular education students, those labeled academically gifted, and those receiving special education services?
- 3) Is there a significant difference between the end-of-year scores in sixth grade and the beginning of the year scores in seventh grade?

Participants

Participants were sixth- (*N* = 218) and seventh- (*N* = 260) grade students from one rural middle school in the southeastern region of the United States. In this school’s district, 51.5% of students received free or reduced lunch. Table 1 shows available student demographics. The vast majority of students had all three STAR-R scores and an EOG reading score. Of those who were missing a STAR-R score, most were only missing one out of the three (24 sixth-grade students and 20 seventh-grade students). A small number of students were missing two out of three STAR-R scores (three

Table 1
Student Demographics

Variable	Sixth Grade		Seventh Grade	
	<i>N</i>	%	<i>N</i>	%
Total	218	100	260	100
White	125	57.3	159	61.2
African American	58	26.6	64	24.6
Hispanic	21	9.6	18	6.9
Multiracial	13	6.0	19	7.3
Asian	1	0.5	0	0
Female	110	50.5	124	47.7
Male	108	49.5	136	52.3
ELL Students	6	2.8	7	2.7
Students with IEPs	24	11	41	15.8
Gifted Students	34	15.6	39	15

Note. IEP = Individualized Education Plan; ELL = English Language Learner.

sixth-grade students and 10 seventh-grade students). One seventh-grade student was missing an EOG score, and one seventh-grade student did not have any Star scores.

Measures

Star-Reading

The STAR-R is a CAT administered on a computer with access to the internet. It measures the reading skills of students K-12. Its goal is to provide information about students’ reading abilities that can be used to modify instruction in the classroom and compare to national norms (Renaissance, 2023b). The STAR-R has 34 items and takes an average of 20 minutes. Because it is computer-adaptive, items shown to students are modified based on the previous answers given in a test. Students receive three types of scores. Their criterion-referenced score denotes whether a student is considered “proficient.” Their norm-referenced score shows how students compare to other students at the same point in time. It is shown as a “percentile rank” and as a “grade equivalent” score. The scaled score enables an individual to see a student’s progress over time. For STAR-R, the scale ranges from 600-1400.

Renaissance, the publisher of the STAR assessments, provides validity scores of the STAR-R for various reading achievement assessments (Renaissance, 2023b). Overall, the range of predictive validity for grades one through six was .69 - .72 (average $r = .71$). For grades seven through twelve, it was .72 - .87 (average $r = .80$). The range of mean correlations for within-grade concurrent validity for grades one through six was .72 - .80 (average $r = .74$) and for grades seven through twelve was .65 - .76 (average $r = .72$).

Renaissance also provides reliability scores for the STAR-R, reporting that the overall reliability of the scaled score to be .98. The coefficient for both grade six and grade seven was .95 (Renaissance, 2023b). In addition, Renaissance (2023a) has reported criterion validity for the STAR-R with the NC EOG-R. STAR-R scores were strongly positively correlated with NC-EOG-R scores across grades three through eight (averaging $r = .78$).

North Carolina End-of-Grade Reading Assessment

The North Carolina End-of-Grade Test in Reading is a statewide standardized assessment administered yearly for students in grades three through eight. It has 48 items in grades three to five and fifty-two items for grades six to eight (North Carolina Department of Public Instruction (NCDPI, 2023)). It takes an estimated two hours to complete. It is administered online. Students' performance is grouped into one of four levels: "Not Proficient," "Level 3," "Level 4," and "Level 5" (NCDPI, 2022). Students receive a scaled score that adjusts based on grade level. The score that determines one's level changes each year with a student's grade.

The NC EOG assessment is deemed reliable and valid as documented by NCDPI. The reliability estimates range from .90 to .93 across grade levels, and the criterion-related validity ranges from .61 to .66 across grade levels (NCDPI, 2016).

Procedures

STAR-R scores were independently collected and stored by school personnel prior to this study. Students took the STAR-R in August of 2022 (Fall), January of 2023 (Winter), and May of 2023 (Spring). All examiners were trained on STAR-R administration before giving the assessment. STAR-R was group-administered on computers and was automatically scored by the computer as students completed the assessment.

This study was deemed exempt from Institutional Review Board (IRB) approval by the authors' university's IRB because the study did not meet the federal definition of research involving human subjects because the data analyzed was not individually

identifiable and was archival. The researchers received consent from the school district administration to conduct a secondary data analysis of the STAR-R and NC-EOG-R scores for all students. Students' STAR-R scores from the fall, winter, and spring of the academic year were examined against their NC-EOG-R scores. Bivariate correlations were performed using SPSS to test the criterion validity of the STAR-R for the NC-EOG-R. Independent ANOVAs were performed to examine the mean differences between regular education students, students labeled academically gifted, and students receiving special education services. An independent-samples t -test was completed to determine if significant differences existed between students' end-of-year STAR-R scores and their beginning of the year STAR-R scores in the next grade level.

Results

The first research question in this study examined the criterion validity of the STAR-R for the NC-EOG-R at three points in the academic year using bivariate correlations. All correlations between the STAR-R scores and NC-EOG-R scores were strong and positive, with a p -value of $< .001$.

In sixth grade, the correlation between each STAR-R score and the NC-EOG-R ranged from .82 to .83. The fall and spring STAR-R scores had stronger correlations with the NC-EOG-R, $r(202) = .83$ and $r(209) = .83$, respectively. Table 2 reports descriptive statistics for sixth-grade STAR-R and NC-EOG-R scores, as well as correlations between the STAR-R scores and the NC-EOG-R scores. Additionally, we looked at R^2 for these correlations to find the amount of variance explained by STAR-R scores for the NC-EOG-R scores. For sixth grade, STAR-R scores explained between 67% and 69% of the variance in NC-EOG-R scores, depending on what point in the year the STAR-R was taken.

In seventh grade, the correlations were slightly stronger, ranging from .83 to .85. The winter STAR-R score demonstrated the strongest correlation with the NC-EOG-R, $r(231) = .85$. Table 3 shows descriptive statistics for seventh-grade STAR-R and NC-EOG-R scores, as well as correlations between the STAR-R scores and the NC-EOG-R scores. As with sixth grade, we used the correlations above to calculate how much variance in NC-EOG-R scores was explained by STAR-R scores. The variance explained ranged from 68% to 72%, depending on the time of year the STAR-R was taken.

Table 2

Sixth Grade: Descriptive Statistics and Correlations Between Star Scores and EOG Scores

Variable		1	2	3	4
1. EOG Reading	Pearson correlation	--			
	<i>n</i>	218			
2. Fall STAR-R	Pearson correlation	.83	--		
	<i>n</i>	204	204		
3. Winter STAR-R	Pearson correlation	.82	.87	--	
	<i>n</i>	209	195	209	
4. Spring STAR-R	Pearson correlation	.83	.87	.84	--
	<i>n</i>	211	197	205	211
<i>M</i>	--	552.4	616.7	643.4	685.2
<i>SD</i>	--	9.4	211.3	242.5	239.5

All correlations are significant at $p < .001$ (2-tailed).

Table 3

Seventh Grade: Descriptive Statistics and Correlations Between Star Scores and EOG Scores

Variable		1	2	3	4
1. EOG Reading	Pearson correlation	--			
	<i>n</i>	259			
2. Fall STAR-R	Pearson correlation	.83	--		
	<i>n</i>	243	244		
3. Winter STAR-R	Pearson correlation	.85	.86	--	
	<i>n</i>	233	228	234	
4. Spring STAR-R	Pearson correlation	.84	.86	.85	--
	<i>n</i>	252	238	231	253
<i>M</i>	--	551.3	679.2	671.7	718.7
<i>SD</i>	--	9.4	246.9	258.5	255.3

All correlations are significant at $p < .001$ (2-tailed).

Table 4
Sixth Grade: Mean Differences and ANOVA Results

Dependent variable	$F(2, 200 - 208)$	p	η^2	Special Education		Regular Education		Academically Gifted	
				M	SD	M	SD	M	SD
BOY Star-R	88.628	<.001	.469	319.09	135.92	604.51	154.20	870.44	168.67
MOY Star-R	87.388	<.001	.459	327.17	146.05	626.01	187.60	953.27	159.34
EOY Star-R	77.839	<.001	.428	339.26	143.20	678.27	184.53	950.74	193.00

Note: Means in a row are all significantly different from one another

Table 5
Seventh Grade: Mean Differences and ANOVA Results

Dependent variable	$F(2, 201 - 250)$	p	η^2	Special Education		Regular Education		Academically Gifted	
				M	SD	M	SD	M	SD
BOY Star-R	90.57	<.001	.429	347.24	161.35	682.27	190.73	939.78	191.89
MOY Star-R	95.89	<.001	.454	338.78	159.65	667.19	200.86	979.70	174.51
EOY Star-R	103.21	<.001	.452	365.76	186.96	732.19	189.19	979.41	194.24

Note: Means in a row are all significantly different from one another

Table 6
Independent-Samples t -test Comparing End of Year Sixth Grade STAR-R Scores to Beginning of Year Seventh Grade STAR-R Scores

Variable	End of Sixth Grade		Beginning of Seventh Grade		t	df	p	Cohen's d
	M	SD	M	SD				
STAR-R	685.22	239.52	679.18	246.85	-.264	453	.39	-.025

*Equal variances not assumed.

In seventh grade, the correlations were slightly stronger, ranging from .83 to .85. The winter STAR-R score demonstrated the strongest correlation with the NC-EOG-R, $r(231) = .85$. Table 3 shows descriptive statistics for seventh-grade STAR-R and NC-EOG-R scores, as well as correlations between the STAR-R scores and the NC-EOG-R scores. As with sixth grade, we used the correlations above to calculate how much variance in NC-EOG-R scores was explained by STAR-R scores. The variance explained ranged from 68% to 72%, depending on the time of year the STAR-R was taken.

The second research question addressed the mean differences in scores between subgroups of students using one-way ANOVAs. All mean differences were significant, and all p -values were $< .001$.

Comparisons were made among students who receive special education, regular education, and academically gifted students. This comparison was completed for STAR-R data collected at the beginning, middle, and end of the academic year. A one-way analysis of variance indicated a significant difference in STAR-R at the beginning of the year, $F(2, 201) = 88.62, p < .001$. Student-Newman-Keuls (SNK) post hoc test results revealed that the special education group had significantly lower STAR-R scores ($M = 319.09, SD = 135.92$) than the regular education group ($M = 604.51, SD = 154.20$), and the academically gifted group was significantly higher than the regular education group ($M = 870.44, SD = 168.67$). The same pattern and post hoc results were seen at the middle and end of the year, as shown in Table 4.

Similar results were found when looking at scores from seventh-grade students. The beginning of the year data for seventh graders was significant, $F(2, 200) = 60.10, p < .001$. SNK post hoc test results revealed that the special education group had significantly lower STAR-R scores ($M = 347.24, SD = 161.35$) than the regular education group ($M = 682.27, SD = 190.73$), and the academically gifted group was significantly higher than the regular education group ($M = 939.78, SD = 191.89$). The same pattern and similar post hoc results were seen at the middle and end of the year, as shown in Table 5.

The final research question was answered using an independent-samples t -test conducted to compare STAR-R scores for the same group of students at the end of their sixth grade and then again at the beginning of their seventh-grade year. There was no significant difference in the scores from the end of sixth grade ($M = 685.22, SD = 239.52$) to the beginning of seventh grade ($M = 679.18, SD = 246.85$) time frames; $t(453) = -264, p = .39$. Data is summarized in table 6. These results

suggest that students' scores do not change significantly over the summer break.

Discussion

To contribute to the body of literature regarding computer-adaptive tests (CATs) in secondary academic settings, this study examined the criterion validity of one particular CAT, the STAR-R, in comparison to a high-stakes assessment, the NC-EOG-R, at the middle school level. The correlations from this study were stronger than those in some prior studies, where researchers found correlations between CATs and high-stakes assessments to be moderate (Clemens et al., 2015; Shapiro & Gebhardt, 2012). However, the strength of the correlations between the STAR-R and the NC-EOG-R in both sixth and seventh grade were consistent with some recent findings (Turner et al., 2022) and the publisher's own reports (Renaissance 2023a, 2023b), which found the correlations between CATs and high-stakes assessments to be strong. These correlations suggest that STAR-R has high concurrent and predictive validity when the criterion measure is a statewide assessment.

With these strong correlations at all three points in the academic year, these results suggest that CATs, such as the STAR-R, may be a valuable tool in identifying students needing an academic intervention or enrichment as early as nine months prior to the high-stakes assessment. At the beginning of sixth grade and the beginning of seventh grade, STAR-R accounted for almost 70% of the variation in NC-EOG-R scores. The strong concurrent and predictive validity coefficients for STAR-R at the sixth and seventh grade level at the start of the year ($r = .83$ and $.83$) extends the current literature and provides further evidence that STAR-R is appropriate for measuring reading achievement at the secondary level.

To focus on the extent to which STAR-R can differentiate between groups of students with different learning needs (i.e., those receiving special education services and those considered academically gifted), this study utilized ANOVAs to compare mean differences between these groups of students. In both sixth and seventh grade, there was a significant mean difference between the groups. On average, academically gifted students scored significantly higher than those in regular education, and students receiving special education services scored significantly lower than those in regular education. These findings are in line with expectations. These results further support the validity of STAR-R in its ability to discriminate between groups of students.

Future research is needed to validate and expand upon the results of this study. Replications are suggested to examine the stability of these findings across a variety

of settings and among different populations in different states. It is also important for this research to be conducted across the United States, as statewide assessments are designed based on standards selected by each state. As shown by the findings of Kingsbury et al. (2003), state assessments often have many limitations not limited to their psychometric qualities. Significant differences between test content across states hinder using a state assessment as a reliable and valid measure of achievement. Despite these issues, these assessments continue to be required by federal legislation (Every Student Succeeds Act, 2015) and remain an important standard for evaluating students, teachers, and school districts.

Also, replication of the study using a nationally normed test such as the NAEP has the potential to provide further information about the criterion validity of STAR-R. Comparing to the national test would yield valuable information because it removes the variability that exists between states and their reading assessments. This study should also be replicated with a larger sample size and a more diverse population. Further, studies can also examine the significance of the relationships adjacent to those explored in this study, such as the relationships between cut scores for CAT benchmarks and high-stakes assessments and the relationship between scores and cut scores for CAT benchmarks and high-stakes assessments in subjects other than reading.

Limitations

Readers should consider two limitations before drawing conclusions about this study's results. These data were gathered from one rural school in North Carolina; thus, generalizability to populations beyond this sample might be limited. A second limitation of this study concerns the archival nature of the data. Because this was a secondary data analysis, researchers had no data regarding the fidelity of the administration procedures.

Implications for Practice

Despite the limitations discussed above, this study provided significant results that may be valuable in academic settings. Some research has shown that CATs may have some positive benefits in academic settings (Martin & Lazendic, 2018; Ortner et al., 2014), and further support of the validity of the STAR-R suggests that educators can utilize these benefits without sacrificing validity in their benchmark assessments. The strength of this relationship at the beginning of the year, in particular, suggests that interventions can reliably begin earlier than one might initially think. This is further supported by the data showing that students' scores do not change drastically over the summer. These

findings suggest schools might consider defining intervention/instructional groups and the students who fit best into those groups at the end of the year instead of waiting till the start of the next school year. This would allow schools to begin targeted instruction earlier in the school year.

Further, this study demonstrates that CAT benchmark assessments like the STAR-R are valid at the middle school level. This supports STAR-R as an assessment that has the potential to be used consistently throughout much of a student's grade-school experience, making it easier to track progress and decreasing the opportunity for complications that come with the learning curve associated with navigating various new assessments.

These data are also important for schools because a significant benefit of utilizing a CAT is the ease of administration, allowing schools to have an entire class assessed at one time. Because the validity is high, CAT assessments appear to be a great starting point for a multi-gating benchmark approach where traditional CBM could be best used as a part of a survey-level assessment for students at the secondary level who do poorly on STAR-R. This can be completed as part of the multidisciplinary problem-solving process for groups of students who perform poorly on STAR-R. For intervention planning, further information (oral reading fluency as an example) can be collected on students who perform below established benchmarks. At the middle school level, it is also more feasible and practical to administer traditional CBM to small groups of students and a CAT like the STAR-R to the entire school.

References

- Abrams, L., Varier, D., & McMillan, J.H. (2012). Formative assessment and benchmark testing: Phase 2. *Metropolitan Educational Research Consortium*.
https://scholarscompass.vcu.edu/cgi/viewcontent.cgi?article=1014&context=merc_pubs.
- Binet, A., & Simon, T. (1916). *The development of intelligence in children: The Binet-Simon scale* (E. Kite, Trans.). Williams & Wilkins. (Original work published 1905)
- Clemens, N. H., Hagan-Burke, S., Luo, W., Cerda, C., Blakely, A., Frosch, J., Gamez-Patience, B., & Jones, M. (2015). The predictive validity of a computer-adaptive assessment of kindergarten and first-grade reading skills. *School Psychology Review*, 44(1), 76-97.
<https://link.gale.com/apps/doc/A435358919/AONE?u=winthropudl&sid=bookmark-AONE&xid=dficabc3>

- Cohen, J. (2013). *Statistical power analysis for the behavioral sciences* (Rev. ed.). Academic Press, 25-27. (Original work published 1977)
- Csikszentmihalyi, M., & LeFevre, J. (1989). Optimal experience in work and leisure. *Journal of Personality and Social Psychology*, 56(5), 815–822. <https://doi.org/10.1037/0022-3514.56.5.815>
- Espin, C., Wallace, T., Lembke, E., Campbell, H., & Long, J. D. (2010). Creating a progress-monitoring system in reading for middle-school students: Tracking progress toward meeting high-stakes standards. *Learning Disabilities Research and Practice*, 25(2), 60–75. <https://doi.org/10.1111/j.1540-5826.2010.00304.x>
- Every Student Succeeds Act, Pub. L. No. 114-354 (2015). Retrieved from <https://www.congress.gov/114/plaws/publ95/PLAW-114publ95.pdf>
- Jones, M. G., & Ennes, M. (2018). High-stakes testing. *Oxford Bibliographies*. <https://doi.org/10.1093/obo/9780199756810-0200>.
- Kingsbury, G. G., Olson, A., Cronin, J., Hauser, C., & Houser, R. (2003). *The state of state standards: Research investigating proficiency levels in fourteen states*. Lake Oswego, OR: Northwest Evaluation Association.
- North Carolina Department of Public Instruction. (2016). Data results. Raleigh, NC: North Carolina Department of Public Instruction. Retrieved from <http://www.ncpublicschools.org/accountability/reporting/>
- Martin, A. J., & Lazendic, G. (2018). Computer-adaptive testing: Implications for students' achievement, motivation, engagement, and subjective test experience. *Journal of Educational Psychology*, 110(1), 27–45. <https://doi.org/10.1037/edu0000205>.
- North Carolina Department of Public Instruction. (2022). North Carolina end-of-grade tests of reading grades 3–8 (edition 5) and beginning-of-grade 3 reading test. <https://www.dpi.nc.gov/documents/accountability/testing/achievelevels/eog-reading-grades-3-8-and-beginning-grade-3-reading-test-achievement-level-ranges-and-descriptors>.
- North Carolina Department of Public Instruction. (2023). End-of-grade reading tests at grades 3–8: North Carolina test specifications. <https://www.dpi.nc.gov/documents/accountability/testing/eog/eog-reading-grades-3-8-test-specifications>.
- Ortner, T. M., Weißkopf, E., & Koch, T. (2014). I will probably fail: Higher ability students' motivational experiences during adaptive achievement testing. *European Journal of Psychological Assessment*, 30(1), 48–56. <https://doi.org/10.1027/1015-5759/a000168>
- Renaissance. (2023a). Pathway to proficiency: Linking star reading and star math to the North Carolina end-of-grade (EOG) tests. <https://renaissance.widen.net/view/pdf/vvkvovtimg/R45814.pdf?t.download=true&u=zceria>.
- Shapiro, E. S., & Gebhardt, S. N. (2012). Comparing computer-adaptive and curriculum-based measurement methods of assessment. *School Psychology Review*, 41(3), 295–305. <https://link.gale.com/apps/doc/A309458376/PPPC?u=winthropudl&sid=bookmark-PPPC&xid=d5fc4fb6>
- Turner, M. I., Van Norman, E. R., & Hojnoski, R. L. (2022). An independent evaluation of the diagnostic accuracy of a computer adaptive test to predict proficiency on an end of year high-stakes assessment. *Journal of Psychoeducational Assessment*, 40(7), 911–916. <https://doi.org/10.1177/07342829221110676>
- Whitley, S. (2019). Oral reading fluency and MAZE selection for predicting 5th and sixth grade students' reading and math achievement on a high stakes summative assessment. *Reading Improvement*, 56(1), 24–35. <https://link.gale.com/apps/doc/A584328146/AONE?u=winthropudl&sid=bookmark-AONE&xid=047021fa>.
- Weiss, D. J. (1973). *The stratified adaptive computerized ability test* (Research Report 73-3). Minneapolis: University of Minnesota, Department of Psychology, Psychometric Methods Program, Computerized Adaptive Testing Laboratory. <https://eric.ed.gov/?id=ED084301>.

PSYCHEDELIC AND DISSOCIATIVE-ASSISTED THERAPY, CURRENT RESEARCH AND FUTURE POTENTIAL

ALEXANDER M. SINGELIS

CALIFORNIA STATE UNIVERSITY CHICO

Abstract – Psychedelic-assisted therapy (PAT) involving classical psychedelics such as lysergic acid diethylamide (LSD), psilocybin, and ayahuasca, and therapy with dissociatives such as 3,4 methylenedioxymethamphetamine (MDMA) and ketamine, are emerging as effective novel treatment methods for psychopathology. These compounds are of note because while their application as a psychopharmacological agent is novel, many have been studied for decades in other contexts, and so their associated risks are better understood than other newly developed pharmaceuticals. Emerging research is showing that compared to their potential benefits, the short-term risks of taking these substances are negligible, and long-term aversive effects are virtually unheard of. In patients with treatment-resistant affective disorders, PAT is proving to be an effective novel treatment, with a rapid onset and long-term duration of positive effects. Few of these disorders have had designated new treatment methods developed within the preceding decades. The mystical, neurological, and psychological effects of these compounds, in addition to their seemingly high efficacy and long-enduring effects are what make them promising novel options for treating individuals suffering from psychiatric comorbidities. These treatments provide the potential for palliative psychological relief to individuals dealing with psychiatric ailments without having to numb them from their emotional expression. More research needs to be done to understand the potential of these treatment methods, the population groups that could benefit, and how they affect people's receptiveness to other treatments post-treatment with adjunctive psychedelics or dissociatives. Present research warrants additional focus on the elevated ethical concerns present in psychedelic and dissociative research, such as the limitation of informed consent and the exploitation of client patient relationship.

Keywords: psychedelics, psychiatric disorders, treatment-resistant, long-term effects

Psychiatric disorders such as major depressive disorder and post-traumatic stress disorder (PTSD) are incredibly damaging to public health, and if left untreated can be lethal (World Health Organization, 2022). For individuals suffering from severe cases of affective disorders, anxiety disorders, substance use disorders, eating disorders, and traumatic disorders, their condition(s) can produce challenges in multiple complex and interacting aspects of their quality of life, physiological, and psychosocial health (World Health Organization, 2022). But while clinical understanding of these conditions and their effects on personal functioning has become much broader in the past ten years, few novel treatment methods have been developed (Ceskova & Silhan, 2018).

Medical science has been steadily progressing and our understanding of the brain and neurological

function is growing due to advancements in imaging technologies. Yet mental health practitioners are still relying on updated versions of treatment methods that are decades old (Ceskova & Silhan, 2018). These treatment methods often require weeks to months of repeated administration to achieve desired effects, and their efficacy is questionable (Gomez et al., 2018). In the case of typical pharmacological agents, common chronic adverse physiological and psychological side-effects are well established (Edinoff et al., 2021). For those who do not respond to these first line treatments, there are few other options currently available that do not involve considerable somatic intervention such as deep brain stimulation, and electroconvulsive therapy. As a result, there exists an urgent need to develop new treatment methods that address existing weaknesses. PAT and dissociatives seem to be of heightened interest in this

context, partly because there are so few novel treatment methods available for treatment-resistant patients (Howes et al., 2022). Furthermore, for those struggling with interacting comorbid psychiatric conditions one single course of treatment often does not provide adequate benefits, and so more intense programs involving multiple treatment approaches are required (Kelly et al., 2012). While mid-twentieth century trials of psychedelics produced evidence for the treatment of psychiatric conditions (Aday et al., 2019), public opinion of psychedelic usage became increasingly unfavorable, including their use in a scientific context. Consequently, research into treatment methods involving psychedelics or pharmacological agents with acute psychoactive effects was stunted and has only begun to re-emerge in a serious capacity within the past two decades (Tupper et al., 2015). The research that is emerging surrounding psychotherapy with adjunctive treatment with drugs such as classical psychedelics and dissociatives is of particular interest as it addresses some crucial weak points in most modern frontline treatment methods: *a)* the long latency period between the beginning of treatment and emergence of positive therapeutic results, *b)* the lack of effective treatment options for those unresponsive to modern forms of treatment, and *c)* the lack of effective treatment options for those suffering from comorbid and interacting psychiatric conditions.

Classical Psychedelics

Classification, Subjective Profile, and Potential Mechanisms

The group classical psychedelics include drugs such as LSD, psilocybin-containing mushrooms, and N,N-diethyltryptamine (DMT). They are some of the most frequently studied psychedelics in a therapeutic context. Their mechanisms of action, efficacy in therapeutic treatment, similar low levels of adverse effects, and comparable changes in perception and subjective experience make them able to be grouped together and loosely generalized into the group of classical psychedelics (Leger & Unterwald, 2021). As a group of drugs which produce an altered state of consciousness, they best represent the therapeutic effects and pharmacological characteristics of psychedelics. All act as potent agonists of 5-HT serotonin receptors, particularly 5-HT_{2A} receptors, which have been implicated in the regulation of depressive behaviors (Nautiyal & Hen, 2017). Since the serotonin system plays a large role in mood regulation (Berger et al., 2009), this could help to explain the acute changes in mood and sense of well-being people often experience during a psychedelic trip, as well as longer-term changes in emotion and affect expression. Compared to

dissociatives, classical psychedelics more readily produce vivid perceptual hallucinations such as visual and auditory changes. In addition, they make users more sensitive to sensory information such as light, sound, and touch and provoke feelings and perceptions of abstract mysticism, spiritual, and transformative subjective experiences. Patients often report that their experience taking classical psychedelics had a significant positive impact on them, listing it as one of the most profound and memorable experiences of their lives (Agin-Liebes et al., 2020; Andersen et al., 2020) when allowed to elaborate on treatment experience and effects during follow-up.

Classical psychedelics appear to increase neuroplasticity, synaptogenesis, and spinogenesis. The 5-HT_{2A} serotonin receptor, with which psychedelics readily bind, has been implicated in the mediation of cortical neuroplasticity (Vargas et al., 2023). A broader potential mechanism is neuroplasticity through the stimulation of increased levels of brain-derived neurotrophic factor (BDNF). As an endogenous compound it is important for brain function and synapse formation, particularly in memory-focused areas such as the hippocampus (McPhee et al., 2020) where the creation of new synapses or long-term potentiation of existing synapses is vital to maintaining proper function. Classical psychedelics appear to have a particularly high binding affinity to receptors which trigger the release of neurotrophic factors, far exceeding that of other antidepressants (Moliner et al., 2023). A second mechanism could be alterations in cortical glutamate levels, which, as a secondary function, cause the release of neurotrophic factors, such as BDNF (Muttoni et al., 2019). Changes in neural networks and inter-neural communication could explain some of the long-term therapeutic effects of psychedelic-assisted therapy. For instance, increased top-down regulation by the prefrontal cortex of the default mode network, a series of regions and pathways implicated in automaticity, and the limbic system, which plays a major role in emotion, seems to reduce habitual negative rumination and autonomic threat response in patients with affective disorders (Carhart-Harris et al., 2017). Increased top-down regulation allows greater mediation by the prefrontal cortex of specific behavior and thought-related brain regions, possibly explaining the myriad of constructive changes detected in patients treated with classical psychedelics in a therapeutic setting. The increased density of dendritic spines produced by classical psychedelics (Ly et al., 2018; Shao et al., 2021) could be responsible for the increased usage of endogenous neurotransmitters. As a result, classical psychedelics

could achieve a similar outcome as SSRIs of increasing monoamine usage, albeit through a different mechanism of action. An additional result of classical psychedelic administration is a lasting increase in global brain integration and resting-state functional connectivity (Daws et al., 2022), along with diminished activity in stress response-related areas such as the amygdala (Carhart-Harris et al., 2017; Muttoni et al., 2019) possibly through the increased top-down regulation by the prefrontal cortex.

Current Research, Common Methods and Features

Common features of classical psychedelic research are *a)* various validated psychopathology scale measurements administered in a pretest-posttest design, *b)* some number of psychotherapy sessions before, between, and after instances of psychedelic intervention to build rapport, establish pathological baselines, and integrate features of the psychedelic experience, and *c)* a comfortable space with eyeshades and calming music provided where the psychedelic experience takes place in the presence of professional clinicians (Andersen et al., 2020; Davis et al., 2020; Goodwin et al., 2022; Griffiths et al., 2016; Leger & Unterwald, 2021; Shnayder et al., 2023). These last two features are vital to patient well-being. With all psychoactive drugs, set (personality, present mindset, and expectations) and setting (environmental conditions, such as physical surroundings and the presence of other people) are important in the drug-taking experience, and are essential with drugs such as psychedelics, which can readily produce a wide variety of vivid hallucinations and acute changes in emotional and psychological state.

To minimize adverse effects and maximize therapeutic results, the set and setting were consistently designed to be comfortable and familiar, with features fairly consistent across studies to provide a stable environment for the psychedelic administration session to take place. In order to facilitate positive therapeutic outcomes, patients are frequently encouraged to engage in relaxed introspection and personal exploration of the psychedelic experience. Therapists are in the room or available nearby to guide the patient through any discomfort that might occur during treatment but are usually instructed to not interfere with or try to manipulate the patient's experience. The goal is to allow the psychedelic experience to take place naturally without strict direction from clinicians. When examining the patient's subjective experiences during psychedelic intervention, multiple studies note the positive correlation between the patient's peak perception of a deep and transcendent mystical experience and the

effectiveness of treatment quantified as the effect size between instances of measurement (Andersen et al., 2020; Griffiths et al., 2016), forming a correlation which could explain some of the dose-dependent efficacy common in the current literature. Consequently, there could exist an overlap in the mechanisms that produce an altered state of consciousness and those that facilitate positive therapeutic effects. When psychedelic treatment is over, integrative psychotherapy sessions take place with the goals of discussing the subjective perceptions of the psychedelic experience, as well as any emotional and pathological changes being experienced post-psychedelic intervention. Studies vary in duration and number of therapy and psychedelic sessions, but most conclude with some talk therapy sessions and follow-ups to discuss lasting changes and establish the progression and longevity of treatment results.

Efficacy and Treatment Outcomes

Determining exact efficacy from modern studies of classical psychedelic's therapeutic profile can be challenging due to factors such as small subject groups, variability in treatment design and therapy type, lack of or unclear blinding procedures, narrow subject recruitment criteria, and lack of a control condition. However, many clinical and experimental studies (Davis et al., 2020; Goodwin et al., 2022; Griffiths et al., 2016; Shnayder et al., 2023), in addition to modern meta and systematic-analyses (Andersen et al., 2020; Haikazian et al., 2023; Ko et al., 2023; Leger & Unterwald, 2021; Muttoni et al., 2019; Schimmers et al., 2022), find strong therapeutic results from baseline measurements in multiple areas of psychiatric dysfunction. In a large-scale multi-center randomized double-blind trial by Goodwin et al. (2022), 233 participants with treatment-resistant depression were recruited into a study assessing the antidepressant effects of COMP360, a synthetic psilocybin formulation. Results demonstrated that the highest dosage group of 25mg (n=79) in conjunction with psychological support significantly improved clinician and self-assessed ratings of depression at week three, and reductions from baseline depression levels measured via self-assessment remained by week twelve.

Clinicians working in these studies frequently report marked improvements in patients' psychological well-being, emotional profile, and behavior, including in studies where clinicians rating participants were unaware of the substance administered (Griffiths et al., 2016). Studies frequently rely on clinician and/or self-administered scale assessments. Patients, particularly those facing a diagnosis of terminal or life-threatening illness, often report and demonstrate significant positive changes in more general aspects of psychological health.

Subjects show improvements in constructs such as optimism, spirituality, positive self-image, sense of connection with others and their community, life purpose, sense of peace, positive life-outlook, mortality acceptance, and overall quality of life (Agin-Liebes et al., 2020; Griffiths et al., 2016). One participant noted, “It has given me a different perspective on my life and has helped me to move on with my life and not focus on the possibility of cancer recurring. I try not to hold onto or stress unimportant things” (Agin-Liebes et al., 2020, 162)

Psychedelic-assisted therapies utilizing classical psychedelics produced rapid-acting and sustained positive therapeutic effects in those suffering from affective disorders and substance use disorders (Fuentes et al., 2020) and nicotine addiction. Scores on depressive and anxiety measurement inventories, rates of substance abstinence, personal function, and overall quality of life consistently demonstrate some significant level of dose-dependent efficacy. One meta-analysis of seven studies by Perez et al. (2023) found a significant correlation between the dosage given and level of therapeutic effects for the treatment of depression, where fifty percent of effects were observed at 10.13 mg/70 kg, and ninety five percent of effects were observed at 41.14mg/70 kg. Most significant is the consistent findings of the sustained duration of positive therapeutic effects resulting from a single or multiple psychedelic interventions in studies that have long-term repeated measurement protocols. Although there are still questions about the ideal number of drug interventions, a meta-analysis done by Leger and Unterwald (2021) suggests that repeated psychedelic treatment sessions produce significantly greater efficacy than single-dose treatment.

Efficacy and duration of effects appear strong even in patients characterized as being treatment-resistant (Goodwin et al., 2022), who are a commonly selected cohort in studies examining PAT. Many patients continue to demonstrate large effect sizes from baseline measurements up to and over six months post-treatment. A meta-analysis by Andersen and colleagues (2020) found that psychedelic-assisted therapy was associated with clinically significant reductions in heavy alcohol usage at all assessment points over a 32-week follow-up period. Reductions in depressive symptom scores of around 40-60% were observed six months post-treatment. A follow-up by Agin-Liebes et al. (2020) of cancer patients who received psilocybin-assisted therapy found sustained significant therapeutic effects at the 3.2 and 4.5-year marks. These patients continued to have reduced scores in areas such as anxiety, depression, hopelessness, and demoralization and overwhelmingly reported personally meaningful and positive changes.

Other studies have found positive results with anxiety disorders and disordered eating after a single or multiple treatment sessions with classical psychedelics (Leger & Unterwald, 2021; Peck et al., 2023).

The sustained therapeutic effects of classical psychedelics and PAT after a single or small number of treatment sessions more broadly contribute to the interpretation that there exist underlying biological mechanisms and neural changes occurring as a result of PAT methods. These mechanisms likely extend beyond acute changes such as monoamine levels in the brain, the typical mechanism of modern pharmacological intervention. A trial by Carhart-Harris et al. (2021) suggests that two treatments with psilocybin-assisted therapy could be more effective than a continuous daily administration of the modern SSRI escitalopram. However, their trial failed to achieve statistical significance in part due to limited sample size. Longevity of effects may be mediated by an interaction between the neurological effects of psychedelics, combined with integrative therapy sessions. Neurological alterations might have enough of an effect on mood and motivation to enhance therapeutic engagement to a level that would be expected from non-treatment-resistant patients. In addition, the subjective alterations produced by psychedelics might allow for a deeper and richer exploration of patient experiences to facilitate optimal treatment outcomes. If the positive therapeutic effects of psychedelic therapy persist for the duration that the current research suggests, it is likely that neurological alterations contribute to creating and sustaining long-term improvements in mental health.

Adverse Effects

Adverse effects from classical psychedelics appear to be of low concern in a professional therapeutic context as psychedelics appear to be well tolerated by patients (Andersen et al., 2020; Griffiths et al., 2016; Haikazian et al., 2023; Ko et al., 2023; Muttoni et al., 2019; Perez et al., 2023; Schimmers et al., 2022; Shnayder et al., 2023). Psychedelics are not seriously reinforcing or addictive (Johnson et al., 2018) in a controlled therapeutic setting. Commonly reported side effects are fatigue, nausea, headaches, perceived changes in body temperature, elevated heart rate and blood pressure, and a heightened sense of psychological discomfort, abnormality, or anxiety (Bender & Hellerstein, 2022). Recognizing that anxiety during treatment could be unavoidable in achieving positive therapeutic effects is important. Some level of discomfort might be inevitable for patients during the process of facing and processing their psychological issues, and so might not qualify as an adverse effect by definition.

Moreover, these symptoms can be dealt with and managed by therapists and staff present during experimental treatment and in follow-up integrative sessions. Side effects are typically transient, subsiding by the end of or shortly after psychedelic treatment.

Considering the potential benefits of adjunctive therapeutic treatment with psychedelics, the present evidence suggests an acceptable risk-benefit profile. More serious adverse effects such as self-harm, suicidality, and sustained changes in psychological well-being, cognition, and physical health are extremely rare, according to a meta-analysis of adverse effects by Bender and Hellerstein (2022). In the few cases where severe adverse effects were found, they were determined to not be a result of psychedelic treatment. However, this must be taken into context because those who are considered high risk, such as those with active psychotic conditions or are considered at risk of self-harm and suicidal ideation are generally prevented from inclusion in study participation through strict screening and recruitment criteria.

Dissociatives

Classification, Subjective Profile, and Potential Mechanisms

Dissociatives differ from their classical psychedelic counterparts by not as readily producing vivid hallucinations and abstract mystical experiences. While some hallucinations still might occur, they are not nearly as consistent and abstract as with classical psychedelics. Dissociatives do still enhance perceptual sensitivity to light, sound, and touch and similarly affect neurotransmitter systems such as serotonin, dopamine, and norepinephrine (Mitchell et al., 2021; Sessa et al., 2019; Tedesco et al., 2021). But unlike classical psychedelics, dissociatives are characterized by producing subjective experiences where the user feels as if they are disconnected from their body and are not present.

A prominent dissociative, MDMA, is known for enhancing feelings of social connection, compassion, mindfulness, and empathy, as a result promoting deep emotional explorations between oneself and others (Tedesco et al., 2021). One theorized neurological mechanism is MDMA's positive effect on oxytocin signaling (Kirkpatrick et al., 2014), a neuropeptide/peptide hormone associated with supporting ingroup bonding. An increase in neural oxytocin concentration could help make the purposeful recollection and communication of traumatic memories during psychotherapy sessions more tolerable. As a result, these prosocial characteristics are what make MDMA such a prime candidate for treating anxiety and traumatic disorders, particularly in treatment-resistant individuals, because MDMA has the potential to inhibit

impulsive fear response and diminish the “characteristic feelings of social detachment and hypervigilance in PTSD [patients]” (Hoskins et al., 2021, 2). Consequently, MDMA promotes a cognitive and socioemotional state conducive to patient-therapist communication, facilitating engagement with traumatic memories and encouraging post-traumatic growth and achievement of fear extinction (Gorman et al., 2020; Hoskins et al., 2021; Mitchell et al., 2021; Tedesco et al., 2021).

Ketamine is the drug that least resembles the psychoactive and therapeutic profile of the other drugs described. Neurological mechanisms of therapeutic effects are believed to be a modulation of glutamate system signaling through the disinhibition of cortical GABAergic interneurons by antagonizing their presynaptic N-methyl-D-aspartate receptors (Abdallah et al., 2016; Browne & Lucki, 2013; Floriano et al., 2023; Lazarevic et al., 2021; Matveychuk et al., 2020). Disinhibition is the inhibition of neurons which inhibit other neurons, thus enabling the increased activation of those neurons. An additional mechanism is an increase in dendritic spine growth (Lazarevic et al., 2021; Wu et al., 2021) due to an increase in BDNF as a result of increased glutamate signaling in the prefrontal cortex (Abdallah et al., 2016).

Current Research, Common Methods and Features

MDMA and ketamine therapies are further along in the treatment development process than are classical psychedelics. Ketamine is approved for usage by practitioners both for its traditional usage as an anesthetic and its new antidepressant usage at licensed clinics. In 2017, MDMA was given breakthrough therapy designation for the treatment of PTSD by the Food and Drug Administration to accelerate treatment development. As such, studies with MDMA are homing in on key treatment areas and study design. Ketamine is most prominently used for its antidepressant effects in those who are resistant to treatment. Typically given either intranasally or intravenously for fast drug absorption, the therapeutic effects are rapid. Less common in ketamine studies and treatment are regular integrative therapy sessions. Ketamine treatment seems to lean more on the purely neurological effects to facilitate treatment outcomes, rather than using it to amplify the effects of other treatment methods.

Studies on MDMA are more similar to those done with classical psychedelics, with structured pretest-posttest protocols, integrative therapy, and drug intervention making up the course of treatment, with frequent sample sizes of about sixty participants (N=18-92) (Gorman et al., 2020; Mitchell et al., 2021; Nicholas

et al., 2022; Ot'alora G et al., 2018; van der Kolk et al., 2024; Wolfson et al., 2020). Many modern studies of MDMA-assisted therapy make use of "A Manual for MDMA-Assisted Therapy in the Treatment of PTSD" developed by the Multidisciplinary Association for Psychedelic Studies. The manual outlines various aspects of therapist selection, creation of a safe physical setting, supportive protocols, preparation, aspects of study and therapy conduction, integrative follow-up design, and other design features to facilitate study methodology that can be more easily compared (Multidisciplinary Association for Psychedelic Studies et al., 2017). Patients in experimental groups often undergo multiple sessions of MDMA treatment, sometimes in two closely occurring dosages. The second dosage being supplemental, normally contingent on patient well-being and subjective state (Tedesco et al., 2021).

The purpose of many MDMA trials with treatment-resistant PTSD patients is to elevate subjects' direct engagement with trauma-focused talk therapy. Since these forms of therapy rely on getting patients to recall traumatic memories to help process them and achieve fear extinction, patients often have trouble actively engaging with these memories without experiencing traumatic anxiety, which may lead to a breakdown in active communication and patient engagement. What MDMA-assisted therapy seems to do is achieve a "window of tolerance" (Mitchell et al., 2021, 1031), where cognitive and emotional engagement with traumatic memories becomes more manageable. The window of tolerance might be most open during the active effects of MDMA, but positive changes in emotional processing, fear management, and openness to social interaction that seem to make talk therapy more tolerable appear to persist into the integration sessions weeks after MDMA intervention (van der Kolk et al., 2024). MDMA seems to contribute to the creation of a prime candidate for trauma-focused talk therapy. Patients given MDMA are, in ideal circumstances, engaged emotionally, highly social, and open to recalling and discussing their experiences, but without undergoing a fear response that pushes them away from therapy or causes defensive disengagement during treatment.

Efficacy and Treatment Outcomes

Ketamine appears to have rapid-acting therapeutic efficacy in treating affective disorders, including in treatment-resistant patients (Boudieu et al., 2023; Drozd et al., 2022; Floriano et al., 2023; Nikolin et al., 2023), with efficacy varying marginally depending on the number of interventions, dosage, and route of administration. Depressive symptoms are decreased substantially for a period of a few days, with effects

peaking around the 24-hour mark. However, the antidepressant and anxiolytic effects of ketamine do not appear to persist nearly as long as MDMA and classical psychedelics. Symptoms typically return to near baseline levels within a few weeks post-treatment (Boudieu et al., 2023; Drozd et al., 2022; Matveychuk et al., 2020; Schimmers et al., 2022), necessitating follow-up administration to achieve therapeutic benefits. With ketamine treatment, fewer studies have required repeated integrative psychotherapy in addition to drug administration, which could account for the lack of long-term therapeutic outcomes (Schimmers et al., 2022). A meta-analysis by Drozd et al. (2022) examined studies where ketamine was combined with some form of psychotherapy to treat any type of physical or psychological ailment. Their findings suggest that adjunctive therapy could increase the duration of ketamine's anxiolytic and antidepressant effects, but currently it is unclear whether therapeutic benefits could be extended to the duration seen in research with MDMA or classical psychedelics. The disparity in sustained long-term effects could be because of the differences in induced neuroplasticity and synaptogenesis between ketamine and classical psychedelics. However, a discrepancy could still be beneficial because not all patients can fully submit to the rigorous and involved course of treatment necessary in therapy with psychedelics or alternative dissociatives. Consequently, some patients could benefit from treatment with a drug such as ketamine that appears to have a less abstract subjective experience and provides strong short-term relief after a single intervention. Additionally, ketamine does not readily produce the swath of socioemotional changes that make MDMA a prime candidate in a dual psychoactive therapy treatment model, where one modality of treatment interacts with and enhances the other.

MDMA-assisted therapy appears to be efficacious for treating PTSD and conditions typically comorbid with traumatic disorders. Many studies have affirmed the substantial effect sizes seen in experimental groups using various versions of the Clinician-Administered PTSD Scale (CAPS), a widely used tool for diagnosing and assessing PTSD (Weathers et al., 2013). Many studies have found MDMA to be extremely beneficial and well tolerated by patients with severe and/or treatment-resistant anxiety or PTSD (Gorman et al., 2020; Hoskins et al., 2021; Mitchell et al., 2021; Ot'alora G et al., 2018; Tedesco et al., 2021). MDMA's emotional and sociability-enhancing properties appear to magnify the benefits of therapy and make active engagement more manageable for patients who must repeatedly recall, discuss, and

recontextualize traumatic events. Moreover, these factors also appear to improve rates of patient dropout (Tedesco et al., 2021), which are notoriously high in traumatic therapies (Mitchell et al., 2021) primarily because they are emotionally demanding. A placebo-controlled phase 3 trial of ninety patients found that MDMA-assisted therapy significantly improved measures of emotional intelligence, self-compassion, and feelings of self-capacity in patients with PTSD (van der Kolk et al., 2024). Such factors are crucial in allowing patients to overcome the negative effects of trauma on emotional openness and self-determination and work with clinicians to help overcome traumatic experiences. MDMA seemingly makes the healing and rebuilding process of therapy for victims of trauma more tolerable, justifying its designation as a breakthrough therapy.

Similar to classical psychedelics, therapy with adjunctive MDMA administration produces long-lasting effects, with most studies finding benefits months to a year after treatment ceases. A long-term follow-up of sixteen patients with chronic treatment-resistant PTSD found that “mean CAPS and IES-R [impact of events scale-revised] scores were not statistically different from their 2-month (short-term) mean scores” (Mithoefer et al., 2013, 31). This long-term follow-up was conducted between 17 to 74 months after the final MDMA session. Most subjects reported specific lasting benefits from study participation, such as general well-being (89%), fewer nightmares/flashbacks/intrusive memories (68%), increased ability to feel emotions (68%), and reduced anxiety (63%).

Adverse Effects

Overall, dissociatives seem to be well tolerated by patients in a controlled therapeutic setting (Azari et al., 2023; Boudieu et al., 2023; Ot'alora G et al., 2018). Most adverse effects are brief and nonsevere. Common experiences are dry mouth, teeth clenching, jaw tightness, dizziness, headaches, fatigue, perceived changes in body temperature, nausea, changes in appetite, and temporary increase in anxiety (Mitchell et al., 2021; Tedesco et al., 2021). Exceedingly rare are reports of serious adverse effects such as increased suicidal ideation.

As with classical psychedelic treatment, it should be expected that elevated levels of emotional vulnerability and mood intensity during dissociative treatment could lead to discomfort for patients. However, as long as treatment and clinician protocols are designed to account for these potential changes, their risks can be mitigated, and emotional vulnerability can be utilized to assist in therapeutic mechanisms where appropriate. In considering the potential benefits of adjunctive

dissociative treatment with therapy, the observed adverse risks do not present abnormal levels of concern to patient wellbeing.

Current Research Limitations and Risks

However, there are limitations on what the current body of research can tell us. The limited sample size may negatively affect the generalizability of the current data. While ketamine is being used more widely, and MDMA research is progressing, our understanding of classical psychedelics in a therapeutic setting is still limited by small sample sizes. Studies do not always include true control groups or complete cessation of other regularly taken pharmaceutical agents. Integrative talk therapy methods are not entirely consistent or well documented across studies, although as the field moves towards more phase 3 trials, improvements in both sample group size and methodological rigor will be seen. Current sample groups and participating clinicians are usually disproportionately white, so more research involving groups from various ethnic, racial, and socioeconomic backgrounds is necessary to bolster research conclusions. There also exists great concern about potential bias in the current research (Hovmand et al., 2023) arising from the possibility of accidental unblinding of participants and clinicians during intervention due to the strong acute perceptual effects of many psychedelics and dissociatives. Blinding participants to their placement in an experimental group is challenging when psychedelics and dissociatives produce such strong subject effects. Nevertheless, this problem could be accounted for with more rigorous blinding procedures, active placebos, improved participant screening, and studies including multiple dosage levels. Another source of bias could be the improper introduction of spiritual or religious language on the part of the clinician during post-intervention therapeutic followup (Johnson, 2020). Just as in non-psychedelic forms of psychotherapy, discussion of these topics can be appropriate, but must be initiated by the participant as to not break with proper therapeutic scope and invalidate potential qualitative insights. Once again more rigorous therapeutic protocols are the clear solution. Established in most studies with varying dosage levels are the dose-dependent therapeutic effects of psychedelic and dissociative therapy, which decreases concern over functional unblinding affecting validity.

Conversely as it pertains to risk, most patients who are considered at high risk of serious adverse complications such as those with a history of psychosis, self-harm, and suicide attempts are screened out of studies with psychedelics or dissociatives. The likelihood of discovering a substantial risk as it pertains to

physiological or psychological complications resulting from a single psychedelic or dissociative intervention is low. These substances have been tested for many decades both in controlled studies looking to investigate potential neurotoxicity and in observational work examining regular recreational drug users. However, there are still some questions as to their risks in patients experiencing disruptive psychiatric issues and their usage in a therapeutic setting in combination with various forms of psychotherapy. While a controlled professional setting makes adverse effects more manageable, having patients continue to focus on uncomfortable or disruptive subjective experiences poses a potential risk, especially if post-treatment follow-ups do not adequately address patients' needs. In addition to the potential physiological risks that must be considered, ethical and therapeutic risks could be of elevated concern to participants during psychedelic and dissociative treatment. This is as a result of the intensity of the treatment, the vulnerable state it can leave participants in during drug intervention creating the opportunity for sexual exploitation by clinicians, the level of expertise needed to perform the therapeutic component of the protocol, and the limitations of informed consent during experimental intervention elevated by the emotional effects of psychedelics and dissociatives (Anderson et al., 2020; Johnson, 2020; Sznitman et al., 2024). Improved informed consent procedures and clinician training procedures, clearer guidelines and protective measures can all address these issues, which require special consideration during the formulation and review of research protocols.

Possibilities for Future Research

Questions remain within the modern psychedelic and dissociative research, one being in establishing prime candidate groups who would benefit from these novel treatment methods. Most clear are candidates who have been diagnosed with comorbid psychiatric problems. The development of a problem such as substance use disorder rarely develops on its own, instead often following some other diagnosis and/or life-altering event. New research should examine how to broadly improve psychiatric function to maximize the therapeutic outcomes in multiple areas of dysfunction. Psychedelic and dissociative treatments will not replace the role of modern pharmaceuticals, such as SSRIs. However, studies could examine the longer-term effects of psychedelic and dissociative therapy on treatment-resistant patients. Even if adjunctive treatment with therapy does not eliminate pathological symptoms in all patients on its own, does it make patients who were previously unresponsive to various treatments more

receptive to those same treatments? Consequently, psychedelic and dissociative therapy would be more akin to surgery, an intense and involved process that both reduces dysfunction directly and primes the patient for a more effective long-term process of healing. Given classical psychedelics' and dissociatives' myriad of neurological alterations, there could exist potential for changes in treatment receptivity for treatment-resistant patients. Studies should focus on controlling blinding methods to diminish expectancy effects. However, does the presence of positive language around psychedelic or dissociative treatment in unblinded clinical trials improve treatment outcomes compared to neutral language? It will become increasingly relevant to control exact language used during the research process as the general public becomes more familiar with modern psychedelic and dissociative research separate from the currently stigmatized view of psychoactive drugs. Do certain types of talk therapy interact differently with particular drugs and patient psychiatric dysfunction? If so, can talk therapy be tailored to create treatment plans and protocols particularly well suited to the needs of each patient? At what point in the number of treatment sessions are diminishing returns seen?

Future studies should make room for deeper qualitative data, particularly the direct reports and testimonies of experimental participants. Their insights into the way psychedelic and dissociative therapy affects the course of their treatment journey are of great value in establishing a clear profile of these substances, refining existing treatment methods, and guiding the objectives of future research. Finally, any forms of future research should seek to strengthen blinding procedures, ethical therapeutic protocols, clinician guidelines, participant screening procedures, informed consent procedures and participant protections. If the area of psychedelic and dissociative science is to move forward and become a truly legitimate field, it can only do so safely and properly if it abides by all of the guidelines and protocols which have made modern scientific research possible. While it can be challenging due to the inherently sensational and intriguing nature of these novel substances, greater focus must be put on ensuring scientific and ethical rigor in exploration of these substances' therapeutic potential. "Despite the association between psychedelic use and ego dissolution, grandiosity can loom large with initial psychedelic experiences, leading even conservative individuals to become wildly enthusiastic about the potentials of psychedelics to heal and transform" (Anderson et al., 2020).

Conclusion

Psychedelic and dissociative therapy are novel approaches to treating psychological ailments. They integrate and build upon existing treatment methods while having the potential to achieve outcomes beyond the capabilities of existing treatments. In addition, they possess potential to provide elevated benefits in terms of onset and duration of therapeutic effects and the breadth of conditions with which they can provide relief. Many studies have found strong efficacy for patients with chronic treatment-resistant psychiatric disorders. Psychedelics and dissociatives seem to be well tolerated in a therapeutic setting with minimal adverse consequences for patient health and function. Many appear to have stimulating effects on neuroplasticity and neural wiring conducive to the treatment of psychiatric disorders. More research is needed to bridge the gaps in present research and should be directed toward the goal of developing practical treatment plans for individuals with comorbid dysfunction and resistance to existing treatment methods. Greater emphasis should be put on ensuring participant safety, proper clinician practice, stronger blinding procedures and methodological rigor, and tempering excited expectations about the future of the field.

References

- Abdallah, C. G., Adams, T. G., Kelmendi, B., Esterlis, I., Sanacora, G., & Krystal, J. H. (2016). Ketamine's Mechanism of Action: A Path to Rapid-Acting Antidepressants. *Depression and Anxiety*, 33, 689-697. <https://doi.org/10.1002/da.22501>
- Aday, J. S., Bloesch, E. K., & Davoli, C. C. (2019). Beyond LSD: A Broader Psychedelic Zeitgeist during the Early to Mid-20th Century. *Journal of Psychoactive Drugs*, 51(3), 210-217. <https://doi.org/10.1080/02791072.2019.1581961>
- Agin-Liebes, G. I., Malone, T., Yalch, M. M., Mennenga, S. E., Ponté, K. L., Guss, J., Bossis, A. P., Grigsby, J., Fischer, S., & Ross, S. (2020). Long-term follow-up of psilocybin-assisted psychotherapy for psychiatric and existential distress in patients with life-threatening cancer. *Journal of Psychopharmacology*, 34(2), 155-166. <https://doi.org/10.1177/0269881119897615>
- Andersen, K. A. A., Carhart-Harris, R., Nutt, D. J., & Erritzoe, D. (2020). Therapeutic effects of classic serotonergic psychedelics: A systematic review of modern-era clinical studies. *Acta Psychiatrica Scandinavica*, 143(2), 101-118. <https://doi.org/10.1111/acps.13249>
- Anderson, B. T., Danforth, A., & Grob, C. (2020). Psychedelic medicine: safety and ethical concerns. *The Lancet Psychiatry*, 7(10), 829-830. [https://doi.org/10.1016/S2215-0366\(20\)30146-2](https://doi.org/10.1016/S2215-0366(20)30146-2)
- Azari, L., Hemati, H., Tavasolian, R., Shahdab, S., Tomlinson, S. M., Babilonia, M. B., Huang, J., Tometich, D. B., Turner, K., Jim, H. S. L., & Tabriz, A. A. (2023). The efficacy and safety of ketamine for depression in patients with cancer: A systematic review. *International Journal of Clinical and Health Psychology*, 24(1), Article 100248. <https://doi.org/10.1016/j.ijchp.2023.100428>
- Bender, D., & Hellerstein, D. J. (2022). Assessing the risk-benefit profile of classical psychedelics: a clinical review of second-wave psychedelic research. *Psychopharmacology*, 239(1), 1907-1932. <https://doi.org/10.1007/s00213-021-06049-6>
- Berger, M., Gray, J. A., & Roth, B. L. (2009). The Expanded Biology of Serotonin. *Annual Review of Medicine*, 60, 335-366. <https://doi.org/10.1146/annurev.med.60.042307.110802>
- Boudieu, L., Mennetrier, M., Llorca, P.-M., & Samalin, L. (2023). The Efficacy and Safety of Intranasal Formulations of Ketamine and Esketamine for the Treatment of Major Depressive Disorder: A Systematic Review. *Pharmaceutics*, 15(12), Article 2773. <https://doi.org/10.3390/pharmaceutics15122773>
- Browne, C. A., & Lucki, I. (2013). Antidepressant effects of ketamine: mechanisms underlying fast-acting novel antidepressants. *Frontiers in Pharmacology*, 4, Article 161. <https://doi.org/10.3389/fphar.2013.00161>
- Carhart-Harris, R., Giribaldi, B., Watts, R., Baker-Jones, M., Murphy-Beiner, A., Murphy, R., Martell, J., Blemings, A., Erritzoe, D., & Nutt, D. J. (2021). Trial of Psilocybin versus Escitalopram for Depression. *The New England Journal of Medicine*, 384(15), 1402-1411. <https://doi.org/10.1056/NEJMoa2032994>
- Carhart-Harris, R. L., Roseman, L., Bolstridge, M., Demetriou, L., Pannekoek, J. N., Wall, M. B., Tanner, M., Kaelen, M., McGonigle, J., Murphy, K., Leech, R., Curran, H. V., & Nutt, D. J. (2017). Psilocybin for treatment-resistant depression: fMRI-measured brain mechanisms. *Scientific Reports*, 7(1), Article 13187. <https://doi.org/10.1038/s41598-017-13282-7>
- Cesková, E., & Silhan, P. (2018). Novel treatment options in depression and psychosis. *Neuropsychiatric Disease and Treatment*, 14, 741-747. <https://doi.org/10.2147/NDT.S157475>
- Davis, A. K., Barrett, F. S., May, D. G., Cosimano, M. P., Sepeda, N. D., Johnson, M. W., Finan, P. H., & Griffiths, R. R. (2020). Effects of Psilocybin-Assisted Therapy on Major Depressive Disorder.

- JAMA Psychiatry*, 78(5), 481-489.
<https://doi.org/10.1001/jamapsychiatry.2020.3285>
- Daws, R. E., Timmermann, C., Giribaldi, B., Sexton, J. D., Wall, M. B., Erritzoe, D., Roseman, L., Nutt, D., & Carhart-Harris, R. (2022). Increased global integration in the brain after psilocybin therapy for depression. *Nature Medicine*, 28, 844-851.
<https://doi.org/10.1038/s41591-022-01744-z>
- Drozdz, S. J., Goel, A., McGarr, M. W., Katz, J., Ritvo, P., Mattina, G. F., Bhat, V., Diep, C., & Ladha, K. S. (2022). Ketamine Assisted Psychotherapy: A Systematic Narrative Review of the Literature. *Journal of Pain Research*, 15, 1691-1706.
<https://doi.org/10.2147/JPR.S360733>
- Edinoff, A. N., Akuly, H. A., Hanna, T. A., Ochoa, C. O., Patti, S. J., Ghaffar, Y. A., Kaye, A. D., Viswanath, O., Urits, I., Boyer, A. G., Cornett, E. M., & Kaye, A. M. (2021). Selective Serotonin Reuptake Inhibitors and Adverse Effects: A Narrative Review. *Neurology International*, 13(3), 387-401.
<https://doi.org/10.3390/neurolint13030038>
- Floriano, I., Silvino, A., & Bernardo, W. M. (2023). The use of esketamine in the treatment of patients with oral antidepressant-resistant depression: systematic review and meta-analysis. *Journal of The Brazilian Medical Association*, 69(6), Article e2023D696.
<https://doi.org/10.1590/1806-9282.2023D696>
- Fuentes, J. J., Fonseca, F., Elices, M., Farre, M., & Torrens, M. (2020). Therapeutic Use of LSD in Psychiatry: A Systematic Review of Randomized-Controlled Clinical Trials. *Frontiers in Psychiatry*, 10, Article 943.
<https://doi.org/10.3389/fpsy.2019.00943>
- Gomez, A. F., Barthel, A. L., & Hofmann, S. G. (2018). Comparing the efficacy of benzodiazepines and serotonergic anti-depressants for adults with generalized anxiety disorder: a meta-analytic review. *Expert Opinion on Pharmacotherapy*, 19(8), 883-894.
<https://doi.org/10.1080/14656566.2018.1472767>
- Goodwin, G. M., Aaronson, S. T., Alvarez, O., Arden, P. C., Baker, A., Bennett, J. C., Bird, C., Blom, R. E., Brennan, C., Brusch, D., Burke, L., Campbell-Coker, K., Carhart-Harris, R., Cattell, J., Daniel, A., DeBattista, C., Dunlop, B. W., Eisen, K., Feifel, D., ... Hoppe, A. I. (2022). Single-Dose Psilocybin for a Treatment-Resistant Episode of Major Depression. *New England Journal of Medicine*, 387(1), 1637-1648.
<https://doi.org/10.1056/NEJMoa2206443>
- Gorman, I., Belser, A. B., Jerome, L., Hennigan, C., Shechet, B., Hamilton, S., Yazar-Klosinski, B., Emerson, A., & Feduccia, A. A. (2020). Posttraumatic Growth After MDMA-Assisted Psychotherapy for Posttraumatic Stress Disorder. *Journal of Traumatic Stress*, 33(2), 161-170.
<https://doi.org/10.1002/jts.22479>
- Griffiths, R. R., Johnson, M. W., Carducci, M. A., Umbricht, A., Richards, W. A., Richards, B. D., Cosimano, M. P., & Klinedinst, M. A. (2016). Psilocybin produces substantial and sustained decreases in depression and anxiety in patients with life-threatening cancer: A randomized double-blind trial. *Journal of Psychopharmacology*, 30(12), 1181-1197.
<https://doi.org/10.1177/026988116675513>
- Haikazian, S., Chen-Li, D. C. J., Johnson, D. E., Fancy, F., Levinta, A., Husain, M. I., Mansur, R. B., McIntyre, R. S., & Rosenblatt, J. D. (2023). Psilocybin-assisted therapy for depression: A systematic review and meta-analysis. *Psychiatry Research*, 329, Article 115531.
<https://doi.org/10.1016/j.psychres.2023.115531>
- Hoskins, M. D., Sinerton, R., Nakamura, A., Underwood, J. F. G., Slater, A., Lewis, C., Roberts, N. P., Bisson, J. I., Lee, M., & Clarke, L. (2021). Pharmacological-assisted Psychotherapy for Post Traumatic Stress Disorder: a systematic review and meta-analysis. *European Journal of Psychotraumatology*, 12(1), Article 1853379.
<https://doi.org/10.1080/20008198.2020.1853379>
- Hovmand, O. R., Poulsen, E. D., Arnfred, S., & Storebø, O. J. (2023). Risk of bias in randomized clinical trials on psychedelic medicine: A systematic review. *Journal of Psychopharmacology*, 37(7), 649-659.
<https://doi.org/10.1177/02698811231180276>
- Howes, O. D., Thase, M. E., & Pillinger, T. (2022). Treatment resistance in psychiatry: state of the art and new directions. *Molecular Psychiatry*, 27, 58-72.
<https://doi.org/10.1038/s41380-021-01200-3>
- Johnson, M. W. (2020). Consciousness, Religion, and Gurus: Pitfalls of Psychedelic Medicine. *ACS Pharmacology & Translational Science*, 4(2), 578-581.
<https://doi.org/10.1021/acspsci.0c00198>
- Johnson, M. W., Griffiths, R. R., Hendricks, P. S., & Henningfield, J. E. (2018). The abuse potential of medical psilocybin according to the 8 factors of the Controlled Substances Act. *Neuropharmacology*, 142, 143-166.
<https://doi.org/10.1016/j.neuropharm.2018.05.012>
- Kelly, T. M., Daley, D. C., & Douaihy, A. B. (2012). Treatment of Substance Abusing Patients with Comorbid Psychiatric Disorders. *Addictive Behaviors*, 37(1), 11-24.
<https://doi.org/10.1016/j.addbeh.2011.09.010>
- Kirkpatrick, M. G., Francis, S. M., Lee, R., de Wit, H., & Jacob, S. (2014). Plasma oxytocin concentrations following MDMA or intranasal oxytocin in

- humans. *Psychoneuroendocrinology*, 46, 23-31.
<https://doi.org/10.1016/j.psyneuen.2014.04.006>
- Ko, K., Kopra, E. I., Cleare, A. J., & Rucker, J. J. (2023). Psychedelic therapy for depressive symptoms: A systematic review and meta-analysis. *Journal of Affective Disorders*, 322, 194-204.
<https://doi.org/10.1016/j.jad.2022.09.168>
- Lazarevic, V., Yang, Y., Flais, I., & Svenningsson, P. (2021). Ketamine decreases neuronally released glutamate via retrograde stimulation of presynaptic adenosine A1 receptors. *Molecular Psychiatry*, 26, 7425-7435.
<https://doi.org/10.1038/s41380-021-01246-3>
- Leger, R. F., & Unterwald, E. M. (2021). Assessing the effects of methodological differences on outcomes in the use of psychedelics in the treatment of anxiety and depressive disorders: A systematic review and meta-analysis. *Journal of Psychopharmacology*, 36(1), 20-30.
<https://doi.org/10.1177/02698811211044688>
- Liao, C., Gregg, I., Davoudian, P. A., Savalia, N. K., Delagarza, K., & Kwan, A. C. (2021). Psilocybin induces rapid and persistent growth of dendritic spines in frontal cortex in vivo. *Neuron*, 109(16), 2535-2544.
<https://doi.org/10.1016/j.neuron.2021.06.008>
- Ly, C., Greb, A. C., Cameron, L. P., Wong, J. M., Barragan, E. V., Wilson, P. C., Burbach, K. F., Zarandi, S. S., Sood, A., Paddy, M. R., Duim, W. C., Dennis, M. Y., McAlister, A. K., Ori-McKenney, K. M., Gray, J. A., & Olson, D. E. (2018). Psychedelics Promote Structural and Functional Neural Plasticity. *Cell Reports*, 23(11), 3170-3182.
<https://doi.org/10.1016/j.celrep.2018.05.022>
- Matveychuk, D., Thomas, R. K., Swainson, J., Khullar, A., MacKay, M.-A., Baker, G. B., & Dursun, S. M. (2020). Ketamine as an antidepressant: overview of its mechanisms of action and potential predictive biomarkers. *Therapeutic Advances in Psychopharmacology*, 10.
<https://doi.org/10.1177/2045125320916657>
- McPhee, G. M., Downey, L. A., & Stough, C. (2020). Neurotrophins as a reliable biomarker for brain function, structure and cognition: A systematic review and meta-analysis. *Neurobiology of Learning and Memory*, 175(1), Article 107298.
<https://doi.org/10.1016/j.nlm.2020.107298>
- Mitchell, J. M., Bogenschutz, M., Lilienstein, A., Harrison, C., Kleiman, S., Parker-Guilbert, K., G, M. O., Garas, W., Paleos, C., Gorman, I., Nicholas, C., Mithoefer, M., Carlin, S., Poulter, B., Mithoefer, A., Quevedo, S., Wells, G., Klaire, S. S., van der Kolk, B., ... Doblin, R. (2021). MDMA-assisted therapy for severe PTSD: a randomized, double-blind, placebo-controlled phase 3 study. *Nature Medicine*, 27(6), 1025-1033.
<https://doi.org/10.1038/s41591-021-01336-3>
- Mithoefer, M. C., Wagner, M. T., Mithoefer, A. T., Jerome, L., Martin, S. F., Yazar-Klosinski, B., Michel, Y., Brewerton, T. D., & Doblin, R. (2013). Durability of improvement in post-traumatic stress disorder symptoms and absence of harmful effects or drug dependency after 3,4-methylenedioxymethamphetamine-assisted psychotherapy: a prospective long-term follow-up study. *Journal of Psychopharmacology*, 27(1), 28-39.
<https://doi.org/10.1177/0269881112456611>
- Moliner, R., Giry, M., Brunello, C. A., Kovaleva, V., Biojone, C., Enkavi, G., Atenucci, L., Kot, E. F., Goncharuk, S. A., Kauinkoski, K., Kuutti, M., Fred, S. M., Elsilä, L. V., Sakson, S., Cannarozzo, C., Diniz, C. R. A. F., Seiffert, N., Rubiolo, A., Haapaniemi, H., ... Castrén, E. (2023). Psychedelics promote plasticity by directly binding to BDNF receptor TrkB. *Nature Neuroscience*, 26(1), 1032-1041.
<https://doi.org/10.1038/s41593-023-01316-5>
- Multidisciplinary Association for Psychedelic Studies, Mithoefer, M., Mithoefer, A., Jerome, L., Ruse, J., Doblin, R., Gibson, E., Ot'alora, M., & Sola, E. (2017). *A Manual for MDMA-Assisted Therapy in the Treatment of PTSD* (8.1 ed.). Multidisciplinary Association for Psychedelic Studies. <https://maps.org/2014/01/27/a-manual-for-mdma-assisted-therapy-in-the-treatment-of-ptsd/>
- Muttoni, S., Maddalena, A., & Christopher, J. (2019). Classical psychedelics for the treatment of depression and anxiety: A systematic review. *Journal of Affective Disorders*, 258, 11-24.
<https://doi.org/10.1016/j.jad.2019.07.076>
- Nautiyal, K. M., & Hen, R. (2017). Serotonin receptors in depression: from A to B. *F1000Research*, 6.
<https://doi.org/10.12688/f1000research.9736.1>
- Nicholas, C. R., Wang, J. B., Coker, A., Mitchell, J. M., Klaire, S. S., Yazar-Klosinski, B., Emerson, A., Brown, R. T., & Doblin, R. (2022). The effects of MDMA-assisted therapy on alcohol and substance use in a phase 3 trial for treatment of severe PTSD. *Drug and Alcohol Dependence*, 233, Article 109356.
<https://doi.org/10.1016/j.drugalcdep.2022.109356>
- Nikolin, S., Rodgers, A., Schwaab, A., Bahji, A., Zarate Jr., C., Vazquez, G., & Loo, C. (2023). Ketamine for the treatment of major depression: a systematic review and meta-analysis. *eClinicalMedicine*, 62, Article 102127.
<https://doi.org/10.1016/j.eclinm.2023.102127>
- Ot'alora G, M., Grigsby, J., Poulter, B., Van Derveer, III, J. W., Giron, S. G., Jerome, L., Feduccia, A. A., Hamilton, S., Yazar-Klosinski, B., Emerson, A.,

- Mithoefer, M. C., & Doblin, R. (2018). 3,4-Methylenedioxymethamphetamine-assisted psychotherapy for treatment of chronic posttraumatic stress disorder: A randomized phase 2 controlled trial. *Journal of Psychopharmacology*, 32(12), 1295-1307. <https://doi.org/10.1177/0269881118806297>
- Peck, S. K., Shao, S., Gruen, T., Yang, K., Babakanian, A., Trim, J., Finn, K. M., & Kaye, W. H. (2023). Psilocybin therapy for females with anorexia nervosa: a phase 1, open-label feasibility study. *Nature Medicine*, 29(8), 1947-1953. <https://doi.org/10.1038/s41591-023-02455-9>
- Perez, N., Langlest, F., Mallet, L., Pieri, M. D., Sentissi, O., Thorens, G., Seragnoli, F., Zullino, D., Kirschner, M., Kaiser, S., Solmi, M., & Sabe, M. (2023). Psilocybin-assisted therapy for depression: A systematic review and dose-response meta-analysis of human studies. *European Neuropsychopharmacology*, 76, 61-76. <https://doi.org/10.1016/j.euroneuro.2023.07.011>
- Schimmers, N., Brecksema, J. J., Smith-Apeldoorn, S. Y., Veraart, J., van den Brink, W., & Schoevers, R. A. (2022). Psychedelics for the treatment of depression, anxiety, and existential distress in patients with a terminal illness: a systematic review. *Psychopharmacology*, 239, 15-33. <https://doi.org/10.1007/s00213-021-06027-y>
- Sessa, B., Higbed, L., & Nutt, D. (2019). A Review of 3,4-methylenedioxymethamphetamine (MDMA)-Assisted Psychotherapy. *Frontiers in Psychiatry*, 10, Article 138. <https://doi.org/10.3389/fpsy.2019.00138>
- Shnayder, S., Ameli, R., Sinaii, N., Berger, A., & Agrawal, M. (2023). Psilocybin-assisted therapy improves psycho-social-spiritual well-being in cancer patients. *Journal of Affective Disorders*, 323(1), 592-597. <https://doi.org/10.1016/j.jad.2022.11.046>
- Sznitman, S. R., Broers, B., Auer, R., & Tal, K. (2024). Taking the potential harms of psychedelic-assisted therapy seriously: How do we prevent or mitigate the risks to vulnerable patients? *International Journal of Drug Policy*, 131, Article 104521. <https://doi.org/10.1016/j.drugpo.2024.104521>
- Tedesco, S., Gajaram, G., Chida, S., Ahmad, A., Pentak, M., Kelada, M., Lewis, L., Krishnan, D., Tran, C., Soetan, O. T., Mukona, L. T., & Jolayemi, A. (2021). The Efficacy of MDMA (3,4-Methylenedioxymethamphetamine) for Post-traumatic Stress Disorder in Humans: A Systematic Review and Meta-Analysis. *The Cureus Journal of Medical Science*, 13(5), Article e15070. <https://doi.org/10.7759/cureus.15070>
- Tupper, K. W., Wood, E., Yensen, R., & Johnson, M. W. (2015). Psychedelic medicine: a re-emerging therapeutic paradigm. *Canadian Medical Association Journal*, 187(14), 1054-1059. <https://doi.org/10.1503/cmaj.141124>
- van der Kolk, B. A., Wang, J. B., Yehuda, R., Bedrosian, L., Coker, A. R., Harrison, C., Mithoefer, M., Yazar-Klosinski, B., Emerson, A., & Doblin, R. (2024). Effects of MDMA-assisted therapy for PTSD on self-experience. *PLOS One*, 19(1), Article e0295926. <https://doi.org/10.1371/journal.pone.0295926>
- Vargas, M. V., Dunlap, L. E., Dong, C., Carter, S. J., Tombari, R. J., Jami, S. A., Cameron, L. P., Patel, S. D., Hennessey, J. J., Saeger, H. N., McCorvy, J. D., Gray, J. A., Tian, L., & Olson, D. E. (2023). Psychedelics promote neuroplasticity through the activation of intracellular 5-HT_{2A} receptors. *Science*, 379, 700-706. <https://doi.org/10.1126/science.adfo435>
- Wolfson, P. E., Andries, J., Feduccia, A. A., Jerome, L., Wang, J. B., Williams, E., Carlin, S. C., Sola, E., Hamilton, S., Yazar-Klosinski, B., Emerson, A., Mithoefer, M. C., & Doblin, R. (2020). MDMA-assisted psychotherapy for treatment of anxiety and other psychological distress related to life-threatening illnesses: a randomized pilot study. *Scientific Reports*, 10, Article 20442. <https://doi.org/10.1038/s41598-020-75706-1>
- World Health Organization. (2022). *Mental disorders*. World Health Organization (WHO). Retrieved October 3, 2024, from <https://www.who.int/news-room/fact-sheets/detail/mental-disorders>
- Wu, M., Minkowicz, S., Dumrongprechachan, V., Hamilton, P., & Kozorovitskiy, Y. (2021). Ketamine Rapidly Enhances Glutamate-Evoked Dendritic Spinogenesis in Medial Prefrontal Cortex Through Dopaminergic Mechanisms. *Biological Psychiatry*, 89(11), 1096-1105. <https://doi.org/10.1016/j.biopsych.2020.12.022>

CALL FOR PAPERS

The *Journal of Psychological Inquiry* encourages undergraduate students to submit manuscripts for publication. Consider the following when you begin to write your manuscript.

- Manuscripts must have an undergraduate student as the primary author. Manuscripts written by anyone who has already graduated from college are acceptable if the work was completed while the primary author was still an undergraduate student. Graduate students or faculty may be co-authors, if their role was one of teacher or mentor rather than equal collaborator.
- Include a sponsoring statement from a faculty supervisor. Faculty sponsors should confirm that they inspected the paper's content, method, adherence to APA style and ethics, grammar, and overall presentation. This sponsoring statement should be uploaded with the manuscript.
- For a manuscript to be considered for publication in JPI, the first author must meet one of the following conditions: a) the primary author has paid a one-time \$30 processing fee, or b) the primary author is or was a student at an institution that has paid an annual \$80 processing fee for unlimited submissions from students who attend that institution.
 - To submit the processing fee, either submit a payment on Venmo.com, or send a check.
 - Venmo payment: submit to @psychinquiry
 - Check: make it payable to the **Great Plains Behavioral Research Association**, and send to the address below:
 - Ken Sobel
Department of Psychology and Counseling
University of Central Arkansas
201 Donaghey Ave.
Conway, AR 72035
- Submit original manuscripts only. Do not submit manuscripts that have been accepted for publication or have been published elsewhere.
- All manuscripts should be formatted in accordance with the latest edition of the APA Publication Manual.
- To submit a manuscript, go to the submission portal at www.editorialmanager.com/jpi
- The reviewing process should ideally require 60 days between submitting a manuscript and receiving a reply from the action editor.
- If a manuscript requires revisions, the author or authors are responsible for making the necessary changes and resubmitting the manuscript to the journal. Manuscripts may need to be revised more than once before being accepted for publication.

The *Journal of Psychological Inquiry* publishes each of the following kinds of articles.

- Empirical studies
- Literature reviews
- Historical articles
- Special features I: Evaluating controversial issues.
 - Two students work together on different facets of the same issue.
 - Select a controversial issue relevant to an area of psychology.
 - Examples:
 - Developmental psychology: Does violence in the media have harmful effects on children?
 - Human sexuality: Are sex and gender categorical or continuous variables?
 - Cognitive psychology: Are repressed memories real?
 - Each student addresses the current empirical research and makes a persuasive case for one side of the argument.
- Special features II: Conducting psychological analyses – Dramatic
 - This type of article is a psychological analysis of a television program or movie.
 - Select an episode from a popular, 30–60 minute television program, or a well-known feature-length film (typically between 90 and 120 minutes long).

- Describe the salient behaviors, activities, and / or interactions of the main characters, and interpret them using psychological concepts and principles.
 - Use appropriate concepts and principles from the research literature.
 - The manuscript should identify the title of the show or film, and for television shows, the name of network and episode.
 - See the APA style guide to find out how to appropriately reference an episode of a television show or movie.
- Special features III: Conducting psychological analyses – Current events
 - This type of article analyzes a current event.
 - Select an event that has garnered widespread coverage in the national media.
 - Analyze the event from one or more areas of psychology.
 - Pay close attention to the people at the center of the event, and to the people who were affected, directly or indirectly, by the event.
 - What were their motivations, expectations, and reactions to the event?
- Special features IV: Teaching techniques
 - The student and faculty mentor should select a teaching technique used by the faculty member that the student found to be particularly helpful in promoting learning.
 - Describe the technique in sufficient detail so other faculty members can replicate the technique in their own teaching.
 - Provide reasons why the student thought the technique worked so well.
 - The faculty member should explain why they developed the technique, and what they hoped to accomplish in terms of learning outcomes.