2025

Volume 29, Number 2, Fall

Issn 1085-6641

Journal of Psychological Inquiry

Journal of Psychological Inquiry
Editorial Board

Journal of
Psychological Inquiry

Volume 29 Number 2 Fall, 2025

Journal Contents

From the Editor's Desk

One of the most vital functions of research is its ability to broaden the limits of scholarly exploration into uncharted intellectual territory. With each publication, research has the capacity to inspire readers to explore novel ideas and consider options that may have seemed impossible.

The articles included in this edition certainly exemplify the impactful and inspirational nature of research. As a graduate student and novice researcher, I found myself captivated by new domains of psychological research that I had never considered prior to reading the selections included in this edition. The level of dedication and care each contributing author has given to their work gives me hope for the ever-expanding future of psychology.

I have an insurmountable amount of gratitude for those who submitted their research for this edition. To have the opportunity to read such enriching submissions has been a privilege this semester and is arguably what makes being a copyeditor so enjoyable. I am incredibly optimistic that our readers will find the included selections as captivating and inspiring as I have.

Katelyn Pack

Katelyn Pack
Graduate Assistant
University of Central Arkansas

Journal of Psychological Inquiry
2025, Vol. 29, No. 2, pp. 5-14

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

PREDICTORS OF SUSCEPTIBILITY TO FALSE ONLINE MENTAL HEALTH INFORMATION

MADDEN G. RITZER AND DONNA W. NELSON

WINTHROP UNIVERSITY

Abstract – We investigated factors that predict susceptibility to false online mental health information. We introduced a fictitious mental health disorder, fear of future anxiety (FOFA) and presented the information to young adults in different online formats and with differing prevalence rates. Contrary to our hypothesis, we did not find evidence supporting the idea that the format would alter beliefs about FOFA. Instead, we observed that young adults were susceptible to the false mental health information presented online, regardless of the format or prevalence rate, particularly among those who relied more heavily on social media for health information. Notably, young adults were more inclined to diagnose others with the false mental health issue than they were to diagnose themselves. Moreover, individuals more inclined to believe they had FOFA tended to be White, women, and non-heterosexual. Consistent with our hypothesis, they also displayed characteristics such as heightened sensitivity to criticism, less resilience, and increased feelings of imposter syndrome. These findings highlight young adults' receptiveness to mental health information online regardless of its validity and provide new insight into how unhealthy personality traits might enhance this receptiveness.

People frequently use online media as a source of health information, with an increasing number of people specifically seeking mental health information (Starvaggi et al., 2024). The use of online resources is particularly high for people in crisis situations and when mental health resources are not physically available (Cohen et al., 2023; Situmorang, 2020). Young adults, because of their high levels of interaction with social media in general, are frequent consumers of online mental health information (Pretorius et al., 2019).

Online mental health information can be very beneficial. For example, the availability of social media content related to mental health facilitates help-seeking and creates a sense of community for young people struggling with mental health issues (Li et al., 2023; Zhang et al., 2024). Struggling individuals can both offer and receive social support from others, or receive support from professional providers, online in a way that might not be possible in the physical world (Naslund et al., 2020). Online information also increases public awareness of disorders and has been shown to reduce the associated stigma (Herrera-Peco et al., 2023; Jankowski et al., 2023).

Although online mental health resources can be beneficial, they also have the potential to be harmful. Individuals struggling with a particular disorder can easily access communities that perpetuate the problem, such as pro-eating disorder sites (Zhang et al., 2024). In an examination of the usefulness of mental health communities, Li et al. (2023) discovered that the quality of help-offering replies was inconsistent and in part depended on how accurately the help-seeking adults were able to present their issue. An additional concern is that viewers can misuse information. After TikTok influencers described having unusual and atypical symptoms of rare mental health disorders, psychologists saw a growing number of young adults present with the same symptoms (Olvera et al., 2021). Similarly, people with some mental health issues, such as borderline personality, may exhibit false symptoms that they find online, compromising their ability to receive an accurate diagnosis and treatment (Bastiaens et al., 2024).

The accuracy of online mental health information varies (Starvaggi et al., 2024; Wilhelm et al., 2022). Online information is presented more frequently by non-expert users than by discipline-specific experts, raising concerns about its validity (Reed, 2023). Content moderators lack clear guidance about when and how to present mental health information, resulting in varied quality of what is approved (Hoops et al., 2023). Even when trained therapists use social media to discuss mental health, there are limited ethical guidelines and oversight (White & Hanley, 2023).

Consumers of online mental health content may not have the expertise to discern its quality for themselves. Heiss et al. (2024) presented adults with

short Tik Tok videos that varied in whether the narrator was an expert or not and whether the information was supported with scientific references. Both types of presenters similarly changed viewers' perceptions, and the use of scientific support had no impact. Twitter researchers found that retweeting posts about mental health research had more to do with whether the tweet contained a hyperlink than with the quality of the information (Madden et al., 2024). Ravi et al. (2022) recently argued that self-diagnostic symptoms provided on social media tend to be general enough that even healthy adults may be led to believe that they have a mental health issue.

To further investigate factors that predict belief in fake mental health information presented online, we created a false mental health disorder, fear of future anxiety (FOFA). This fake disorder was inspired by the rising anxiety rates in young adults in the United States, particularly in groups with instability issues such as being unmarried or lower SES (Goodwin et al., 2020). We presented information about the disorder to young adults online via a self-diagnostic test or in a scientific summary of symptoms. We also provided different estimates of its prevalence during early adulthood as being high (65%) or low (15%). We hypothesized that participants would be more likely to perceive the disorder as real, agree that people have it, feel sympathy for those individuals, and show interest in learning more about it when:

1. they were given the self-diagnostic test instead of the scientific summary, because of the personally relevant nature of the information, and simultaneously provided with a higher estimated prevalence rate, suggesting that the disorder is common in their age group,

2. they experienced the unhealthy psychological traits of: sensitivity to criticism, fear of missing out, imposter syndrome, and low resilience,

3. they reported higher levels of reliance on social media for health information.

We also conducted exploratory analyses to investigate whether demographic characteristics of our participants would predict their belief in the fake disorder.

Method

Participants

This study was approved by the university's Institutional Review Board and used a convenience sample. Participants were recruited through college classrooms, social media, and word-of-mouth. Potential participants were provided with a link to the online study, where they encountered additional information about the study and the informed consent. Each participant had to indicate, in the online system, that they agreed to participate before accessing the survey.

We had a total of 105 participants complete our survey; however, one did not correctly respond to the manipulation check item. Thus, the final sample consisted of 104 young adults with a mean age of 19.44 (*SD* = 2.46). Among them, 74 were women, 24 were men, and six categorized themselves as non-binary or "other." Regarding race, 57 participants were White, 30 were Black, and the remainder identified as Multiracial/Other (seven participants), Asian (five participants), or Latinx/Hispanic (four participants). In terms of sexual orientation, 61 participants were heterosexual, 23 were bisexual, eight were gay or lesbian, eight identified as "other," and four did not provide a response.

Materials and Procedure

First, participants responded to four demographic items to assess age, gender, race, and sexual orientation. All participants were then instructed, "You will be learning about the Fear of Future Anxiety (FOFA). You will then be asked questions to test your understanding of this disorder. Please read carefully." Participants then encountered information about FOFA that was designed to look and feel like online health resources, with concise descriptions, bulleted lists, and prevalence statistics. This approach was meant to reflect how online health information is usually encountered. Information varied in two ways: presentation format (scientific summary- see Appendix A or self-diagnostic test-see Appendix B) and prevalence rates (high-65% or low-15%) of the disorder in the college population. Participants were randomly assigned to one of four conditions, resulting in 27 participants in the scientific presentation/low prevalence condition, 27 participants in the diagnostic test/low prevalence condition, 26 participants in the scientific presentation/high prevalence condition, and 24 participants in the diagnostic test/high prevalence condition.

Participants were then presented with six statements describing symptoms to which they could rate their level of agreement on a 5-point scale where 1 represented "strongly disagree" and 5 represented "strongly agree." They were instructed, "The more you agree with these items, the more likely it is that you have symptoms of this disorder." Immediately following the FOFA information, participants encountered a manipulation check item that provided the correct definition of FOFA. Participants could respond "true" or "false." Participants had to respond correctly to be included in the study; the one participant who did not was excluded. Next, participants responded to four researcher-created items to assess 1) interest level in learning more about FOFA, 2) belief that they personally

had the disorder, 3) belief that someone they knew had the disorder, and 4) level of sympathy for people who have the disorder. Responses were made on a 5-point scale where 1 represented "strongly disagree" and 5 represented "strongly agree."

Participants also indicated how often they relied on online resources to learn about mental health issues and how often they believe others are to do this. The response options for both of these items ranged from 1 representing "never" to 5 representing "all the time." We then asked participants the percentage of college-aged students who experience symptoms of this disorder. The experimental conditions had indicated that prevalence rates were either 15% or 65%. The response options were 15%, 25%, 50%, 65%, and 80%.

Next participants responded to four previously published scales. On all four of these scales, there were five response options with 1 representing "strongly disagree" and 5 representing "strongly agree." The first scale included four items comprising the Behavioral Engagement sub-scale and two items from the Affective Engagement subscale of the *Social Media Engagement Scale* (Ni et al., 2020). The published reliabilities for this scale are .71 to .81 (Ni et al., 2020), and we achieved a reliability of .84. Participants then responded to the 6-item *Brief Resilience Scale* (Smith et al., 2009), The published reliabilities for this scale are .80 to .91 (Smith et al., 2009), and we achieved a reliability of .85. Participants responded to the 9-item F*ear of Missing Out Scale* (FOMO; Przybylski et al., 2013). The published reliabilities for this scale are .87 to .90 (Przybylski et al., 2013), and we achieved a reliability of .82. Participants responded to the 7-item *Leary Imposterism Scale* (Leary et al., 2000). The published reliability for this scale is .95 (Freeman et al., 2022), and we achieved a reliability of .93. Last, participants responded to was the 3-item *Sensitivity to Criticism Scale* (Rosenberg, 1965). Responses were made on a 5-point scale where 1 represented "almost never" and 5 represented "almost always" We achieved a reliability of .86.

All participants were debriefed at the end of the study that FOFA was a fictitious disorder and to the purpose of the study.

Results

Hypothesis 1: Comparison of Experimental Condition

A 2 (information presentation) X 2 (prevalence) MANOVA revealed no interactions or main effects among the four conditions on our dependent variables: interest in learning about FOFA, belief participants knew someone with the disorder, belief they personally had the disorder, and sympathy level for people with the disorder. See Table 1 for means and standard deviations.

Although these were not hypothesized, we found two significant main effects, both relating to participants' beliefs about the prevalence rates of the disorder. Participants who were provided high prevalence rates reported a higher prevalence rate (M = 2.19, SD = 1.51) than participants who were provided low prevalence rates (M = 3.92, SD = .71), [F(1, 99) = 57.65, p < .001]. We also found that participants who took the diagnostic quiz (M = 3.33, SD = 1.44) reported higher prevalence rates than those who read the scientific summary (M = 2.68, SD = 1.45), [F(1, 99) = 8.12, p = .005].

Hypothesis 2: Trait Predictors of Susceptibility

We examined relations among variables using Pearson's correlations and the data can be seen in Table 2. Participants who agreed more that they personally had FOFA were more sensitive to criticism, had lower resilience, had higher imposter syndrome, and had higher FOMO. These traits did not predict participants' agreement that they knew someone with FOFA. Participants who felt sympathy for those with FOFA had lower resilience. These represent weak to moderate correlations.

Table 1
Means (and Standard Deviations) for Dependent Variables Across Experimental Conditions

	Scientific Summary	Self-Diagnostic Test	Prevalence Low	Prevalence High
I believe that I might have this disorder.	3.38(1.07)	3.06(1.35)	3.30(1.14)	3.12(1.32)
I believe that I know someone with this disorder.	3.80(.99)	3.57(1.45)	3.85(1.22)	3.50(1.26)
I am interested in learning more about FOFA.	3.86(.93)	3.55(1.15)	3.81(1.04)	3.58(1.07)
I have great sympathy for those with this disorder.	4.14(.86)	4.10(.99)	4.23(.95)	4.00(.88)

Table 2
Correlation Matrix for Traits, Social Media Reliance, and FOFA Perceptions

	1	2	3	4	5	6	7	8	9	10
1 Resilience	-									
2 Sensitivity to criticism	-.31**	-								
3 FOMO	-.21*	.54**	-							
4 Imposter syndrome	-.23*	.43**	.49**	-						
5 Reliance on social media	-.09	.17	.25*	.22*	-					
6 Use social media for mental health info	-.20*	.22*	.30**	.08	.15	-				
7 Have FOFA	-.36**	.28**	.39**	.24*	.15	.24*	-			
8 Know someone with FOFA	-.07	-.01	.14	-.04	.08	.08	.56**	-		
9 Have sympathy for FOFA	-.25*	.05	.05	.08	.02	.26**	.48**	.44**	-	
10 Want to learn about FOFA	-.26**	.08	.04	-.06	.19	.25*	.41**	.21*	.28**	-

Note. * $p < .05$. ** $p < .01$.

Hypothesis 3: Reliance on Social Media

Participants who relied more on online mental health resources also agreed more that they personally had FOFA, were interested in learning about FOFA, and felt sympathy for people with FOFA (see Table 2). These are weak correlations. General reliance on social media did not significantly predict any of these variables related to FOFA.

Participants who agreed more that they rely on online resources to learn about mental health issues, had: lower resilience, higher sensitivity to criticism, and higher FOMO. General reliance on social media was predicted by imposter syndrome and FOMO. These are weak to moderate correlations.

Exploratory Analysis: Demographic Predictors

Next, we examined demographic categories as predictors of beliefs about FOFA and online mental health information using independent *t*-tests. Means and standard deviations can be seen in Table 3 and independent *t*-test data can be seen in Table 4. We

Table 3
Means (and Standard Deviations) for Dependent Variables Across Demographic Categories

	Hetero-sexual	Non-hetero-sexual	Black Adults	White Adults	Men	Women
I believe that I might have this disorder.	3.07 (1.26)	3.53* (1.13)	2.86 (1.22)	3.33* (1.28)	2.61 (1.16)	3.38** (1.22)
I believe that I know someone with this disorder.	3.64 (1.23)	3.82 (1.23)	3.59 (1.20)	3.80 (1.22)	3.13 (1.33)	3.86** (1.19)
I am interested in learning more about FOFA.	3.68 (.99)	3.74 (1.18)	3.66 (1.23)	3.65 (.97)	3.26 (1.25)	3.83* (.98)
I have great sympathy for those with this disorder.	4.05 (.92)	4.24 (.94)	4.21 (.98)	4.05 (.91)	3.83 (.94)	4.19* (.91)
I often use online resources to learn more about mental health issues.	3.51 (1.17)	3.97* (.75)	3.83 (.89)	3.62 (1.18)	3.26 (1.25)	3.75* (1.00)
Reliance on social media	3.69 (.74)	3.54 (.99)	3.39 (.95)	3.72 (.84)	3.14 (.88)	3.73** (.82)
People are too quick to believe mental health information they encounter online.	4.20* (.74)	3.87 (.07)	4.21 (.77)	4.02 (.97)	3.57 (1.04)	4.31** (.71)

Note. * $p \leq .05$. ** $p < .01$.

Table 4
Independent t-test Statistics for Comparisons Across Demographic Categories

				95% CI		
	df	***t*-value**	***p***	**Lower**	**Upper**	**Cohen's *d***
SEXUAL ORIENTATON						
Have FOFA	95	-1.82	.036	-0.96	0.04	1.21
Use social media for mental health info	95	-2.18	.016	-0.89	-0.04	1.03
People are too quick to believe mental health information they encounter online.	95	1.83	.036	0.03	0.70	0.88
RACE						
FOFA	82	-1.61	.053	0.34	0.77	1.26
GENDER						
Have FOFA	93	-2.66	.005	-1.34	-0.20	1.20
Know someone with FOFA	93	-2.49	.007	-1.31	-0.15	1.22
Want to learn about FOFA	93	-2.28	.013	-1.07	-.07	1.05
Have sympathy for FOFA	93	-1.67	.049	-0.81	0.07	0.92
Reliance on social media	93	-2.91	.002	-0.98	-0.18	0.84
Use social media for mental health info	93	-1.91	.029	-1.00	0.02	1.07
People are too quick to believe mental health information they encounter online.	93	-3.88	<.001	-1.12	-0.36	.80

compared Black adults to non-Hispanic White adults, our only two racial groups with sufficiently sized samples, and found that compared to Black adults, White adults agreed more that they had FOFA.

Compared to heterosexual adults, non-heterosexual adults were more likely to agree that they had FOFA, more likely to rely on online mental health resources, and less likely to agree that people are "too quick to believe mental health information they encounter online."

Compared to men, women were more likely to agree that they had FOFA, knew someone with FOFA, were interested in learning about FOFA, and felt sympathy for people with FOFA. Women also had higher social media reliance scores than did men, relied more on online mental health resources, and agreed more that people are "too quick to believe mental health information they encounter online."

Discussion

Hypothesis 1: Comparison of Experimental Conditions

We hypothesized that participants would be more susceptible to believing the fictitious disorder of FOFA if they took a self-diagnostic test after having been told the disorder had high prevalence. We failed to find support for this prediction. Regardless of the way the information was presented, participants seemed to believe that FOFA existed. For example, they indicated great sympathy for those with the disorder.

Participants believed there were higher prevalence rates when they were told the disorder had a high prevalence rate. This outcome is logical and expected. Additionally, we found that participants who took the self-diagnostic test picked higher prevalence rates than participants who encountered the scientific

summary about FOFA. In other words, when the information was presented in a personally relevant way, people assumed that the disorder was more common, even if they were told otherwise. This outcome was unexpected but might be explained by the false-consensus effect, where people make the cognitive mistake of assuming others share their beliefs and behaviors (Ross et al., 1977). Engaging with the self-diagnostic test might have led participants to identify with the disorder's characteristics, making them more likely to believe that others experience it too.

Hypothesis 2: Trait Predictors of Susceptibility

Providing some support for our hypothesis, the susceptibility to false online mental health information co-occurred, although weakly, with unhealthy psychological traits. Young adults who perceived themselves to have FOFA were more sensitive to criticism, were less resilient, and had a higher FOMO. Similarly, young adults who relied more on online resources for mental health information were more sensitive to criticism, were less resilient, and had higher FOMO; these findings support the previously established link between social media use and higher rates of mental health concerns (Nazari et al., 2023; Plackett et al., 2023). These relationships were weak, which suggests there is much to learn. There are likely other variables that better predict susceptibility to false online information, such as the demographic predictors found in our study. In addition, young adults' receptiveness to online mental health information might be so ubiquitous that it is not necessarily limited to those who are psychologically unhealthy.

Hypothesis 3: Reliance on Social Media

Participants who relied more on online resources to learn about mental health were also more likely to believe they had FOFO, feel sympathy for people with FOFA, and want to learn more about FOFA. These outcomes were specifically tied to using social media for mental health information, because general reliance on social media did not predict participants' belief about FOFA. This outcome provides mixed support for our hypothesis that adults who reported higher levels of reliance on social media would also be more likely to believe in the existence of FOFA. Instead, it seems that individuals who specifically trust the credibility of online mental health information are more susceptible to believing false or exaggerated claims about mental health conditions.

Imposter syndrome was the one psychological trait that predicted susceptibility to false online information but not a reliance on online resources for mental health information. One likely explanation is that imposter syndrome is characterized by fear, particularly fear of one's success being uncovered as fraudulent at some point in the future (Clance & Imes, 1978). This fear of an anxiety-provoking future event overlaps with the symptomology of FOFA, all of which focus on the fear of feeling anxiety in the future. However, imposter syndrome may not influence reliance on online mental health resources because its core concern is self-doubt and perceived fraudulence rather than a general tendency to seek or trust online information.

Exploratory Analysis: Demographic Predictors

Although we saw susceptibility to the false online information in the sample as a whole, young adults who were non-heterosexual, White, and women were more likely to believe that they personally had FOFA. Women and non-heterosexual adults also turned to online resources more for their mental health information than men and heterosexual adults. This pattern of receptiveness matches previous research showing that women are more likely than men to utilize, and have positive attitudes toward, online mental health programs (McCall et al., 2023; Smail-Crevier et al., 2019), while Black adults have historically harbored more negative perceptions of mental illness and treatment than White adults (Harris et al., 2020). Less information is known about non-heterosexual adults' engagement with online mental health information; however, Lucassen et al. (2018) reported that LGBTQ+ individuals were receptive to online information, especially in contemporary and engaging formats.

Compared to men, women also were more likely to agree that they knew someone with FOFA, felt sympathy for people with the disorder, and were interested in learning more. These responses may reflect women's socially-driven tendency to report higher levels of empathetic emotions (Pang et al., 2023). Another possibility is that women were more responsive because the information was presented online. The women in our study reported a greater, general reliance on social media than did men, and it is possible that the reliance on social media led them to trust the online content more.

Limitations

This study has limitations worthy of consideration. First, our sample was predominantly composed of college-aged, White women, and the use of a convenience sample further limited the demographic diversity of our participants. Another potential limitation of our study concerns the manner in which participants encountered information about FOFA. While we tried to use a presentation format similar to what individuals might encounter while searching online for health information, it is worth noting that encountering

information within the context of a scientific study may differ qualitatively with respect to credibility and trustworthiness, compared to online searches. It is possible that participants generally tend to trust information embedded in research studies, which may have dampened any potential differences that alternate formatting may have elicited. Additionally, we did not directly verify whether our materials truly mirrored what participants would typically see in their online searches. Finally, while we used symptoms that were designed to align with the characteristics of FOFA, they may have been broad enough that participants found it easy to agree with them, which could have influenced their responses. Future research could consider exploring the effects of presenting more unique or unconventional symptoms to assess whether they would yield different results.

Future Directions

There are several additional areas that could benefit from further investigation. Future research could investigate whether beliefs based on misinformation persist over time and whether repeated engagement with false health information strengthens its impact. Future studies could also examine how media literacy and the ability to critically evaluate content affect how people respond to online misinformation or test the impact of information formatting on the behaviors and attitudes of individuals engaging in real-world online searches of mental health information. Finally, future research could explore ways to help people become more skeptical of faulty online mental health information. Educational initiatives could help to mitigate the spread of misinformation and improve ability to identify reliable sources that support a more informed approach to mental health.

Conclusion

In sum, we found that young adults were susceptible to the false mental health information regardless of the format or purported prevalence rates, especially young adults who relied more heavily on social media for health information or were White, women, or non-heterosexual. This susceptibility to online mental health information is concerning because not all information online is credible (Starvaggi et al., 2024; Wilhelm et al., 2022). The problem may become self-perpetuating, because young adults who rely heavily on online sources of mental health information were also the ones who expressed interest in learning more and felt the most sympathy for people with FOFA. In other words, they may utilize low quality resources to gain information and then have an increased desire to engage further, both cognitively and emotionally, with the faulty information.

Our findings highlight the urgent need to address this issue through initiatives aimed at ensuring that trustworthy mental health information is both readily available and easily accessible. Additionally, efforts should focus on helping young adults develop stronger skills to critically evaluate online content. This will empower them to recognize credible sources and make informed decisions about their mental health.

References

Bastiaens, L., Cristofano-Casella, S. & Bastiaens, J. (2024). Prodromal or factitious psychosis: A cautionary tale in the era of increasing social media use by vulnerable individuals. *Journal of Psychiatric Practice*, *30*(2), 162-166. https://doi.org/10.1097/PRA.0000000000000771

Clance, P. R., & Imes, S. (1978). The imposter phenomenon in high achieving women: Dynamics and therapeutic interventions. *Psychotherapy: Theory, Research & Practice, 15*(3), 1-8. https://doi.org/10.1037/h0086006

Cohen, K., Dobias, M., Morris, R. & Schleider, J. (2023). Improving uptake of mental health crisis resources: Randomized test of a single-session intervention embedded in social media. *Journal of Behavioral and Cognitive Therapy, 33*(1), 24-34. https://doi.org/10.1016/j.jbct.2022.12.001

Freeman, K. J., Houghton, S., Carr, S. E., & Nestel, D. (2022). Measuring impostor phenomenon in healthcare simulation educators: A validation of the Clance Impostor Phenomenon Scale and Leary Impostorism Scale. *BMC Medical Education, 22*(1), 139. https://doi.org/10.1186/s12909-022-03190-4

Goodwin, R. D., Weinberger, A. H., Kim, J. H., Wu, M. & Galea, S. (2020). Trends in anxiety among adults in the United States, 2008-2018: Rapid increases among young adults. *Journal of Psychiatric Research, 130*, 441-446. https://doi.org/10.1016/j.jpsychires.2020.08.014

Harris, J. R. A., Crumb, L., Crowe, A. & McKinney, J. G. (2020). African Americans' perceptions of mental illness and preferences for treatment. *Journal of Counselor Practice, 11*, 1-33. https://doi.org/10.22229/afa1112020

Heiss, R., Bode, L., Adisuryo, Z. M., Brito, L., Cuardra, A., Gao, P., Han, Y., Hearst, M., Huang, K., Kinyua, A., Lin, R., Ma, Y., Manion, T. O., Roh, Y., Salazar, A., Yue, S. & Zhang, P., (2024). Debunking mental health misperceptions in short-form social media videos: An experimental test of scientific credibility cues. *Health Communication, 39*, 3059-3071. https://doi.org/10.1080/10410236.2023.2301201

Herrera-Peco, I., Fernández-Quijano, I. & Ruiz-Núñez, C. (2023). The role of social media as a resource for mental health care. *European Journal of Investigation in Health, Psychology and Education, 13*(6), 1026-1028. https://doi.org/10.3390/ejihpe13060078

Hoops, K., Nestadt, P. S. & Dredze, M. (2023). The case for social media standards on suicide. *The Lancet Psychiatry, 10*(9), 662-664. https://doi.org/10.1016/S2215-0366(23)00222-5

Jankowski, S. E., Yanos, P., Dixon, L. B., & Amsalem, D. (2023). Reducing public stigma towards psychosis: A conceptual framework for understanding the effects of social contact based brief video interventions. *Schizophrenia Bulletin, 49*(1), 99-107. https://doi.org/10.1093/schbul/sbac143

Leary, M. R., Patton, K., Orlando, A., & Funk, W. W. (2000). The imposter phenomenon: Self-perceptions, reflected appraisals, and interpersonal strategies. *Journal of Personality, 68*, 725- 756. https://doi.org/10.1111/1467-6494.00114

Li, J., Liu, D., Wan, C., Liang, Z. & Zhu, T. (2023). Empirical study of factors that influence the perceived usefulness of online mental health community members. *PsyCh Journal, 12*(2), 307-318. https://doi.org/10.1002/pchj.629

Lucassen, M., Samra, R., Iacovides, I., Fleming, T., Shepherd, M., Stasiak, K. & Wallace, L. (2018). How LGBT+ young people use the internet in relation to their mental health and envisage the use of e-therapy: Exploratory study. *Journal of Medical Internet Research, 6*(4), e11249. https://doi.org/10.2196/11249

Madden, E., Prior, K., Guckel, T., Garlick Bock, S., Bryant, Z., O'Dean, S., Nepal, S., Ward, C. & Thornton, L. (2024). 'What do I say? How do I say it?' Twitter as a knowledge dissemination tool for mental health research. *Journal or Health Communication, 29*, 20-33. https://doi.org/10.1080/10810730.2023.2278617

McCall, T., Foster, M. & Schwartz, T. A. (2023). Attitudes toward seeking mental health services and mobile technology to support the management of depression among Black American women: Cross-sectional survey study. *Journal of Medical Internet Research, 25*, e45766. https://doi.org/10.2196/45766

Naslund, J. A., Bondre, A., Torous, J. & Aschbrenner, K. A. (2020). Social media and mental health: Benefits, risks, and opportunities for research and practice. *Journal of Technology Behavioral Science, 5*, 245–257. https://doi.org/10.1007/s41347-020-00134-x

Nazari, A., Hosseinnia, M., Torkian, S. & Garmaroudi, G. (2023). Social media and mental health in students: A cross-sectional study during the COVID-19 pandemic. *BMC Psychiatry, 23*(1), ArtID 458. https://doi.org/10.1186/s12888-023-04859-w

Ni, X., Shao, X., Geng, Y., Qu, R., Niu, G., & Wang, Y. (2020). Development of the Social Media Engagement Scale for Adolescents. *Frontiers in Psychology, 11*, 701. https://doi.org/10.3389/fpsyg.2020.00701

Olvera, C., Stebbins, G. T., Goetz, C. G. & Kompoliti, K. (2021). TikTok tics: A pandemic within a pandemic. *Movement Disorders: Clinical Practice, 8*(8), 1200-1205. https://doi.org/10.1002/mdc3.13316

Pang, C., Li, W., Zhou, Y., Gao, T. & Han, S. (2023). Are women more empathetic than men? Questionnaire and EEG estimations of sex/gender differences in empathic ability. *Social Cognitive and Affective Neuroscience, 18*, nsad008. https://doi.org/10.1093/scan/nsad008

Plackett, R., Blyth, A. & Schartau, P. (2023). The impact of social media use interventions on mental well-being: Systematic review. *Journal of Medical Internet Research, 25*, e44922. https://doi.org/10.2196/44922

Pretorius, C., Chambers, D., Cowan, B. & Coyle, D. (2019). Young people seeking help online for mental health: Cross-sectional survey study. *Journal of Medical Internet Research Mental Health, 6*(8), e13524. https://doi.org/10.2196/13524

Przybylski, A. K., Murayama, K., DeHaan, C. R., & Gladwell, V. (2013). Motivational, emotional, and behavioral correlates of fear of missing out. *Computers in Human Behavior, 29*, 1814-1848. https://doi.org/10.1016/j.chb.2013.02.014

Ravi, D., Chandran, S. & Priya, S. (2022). Social media use for the dissemination of mental health information by medical students. Journal of *Psychiatry Spectrum, 1*(2), 117-120. https://doi.org/10.4103/jopsys.jopsys_20_22

Reed, O. S. (2023). Responses to mental health care posts by social media influencers: The moderating effects of previous experience and follower involvement. *Mass Communication and Society, 27*(3), 475-501. https://doi.org/10.1080/15205436.2023.2269542

Rosenberg, M. (1965). *Society and the Adolescent Self-Image*. Princeton, NJ: Princeton University Press.

Ross, L., Greene, D. & House, P. (1977). The "false consensus effect": An egocentric bias in social perception and attribution process. *Journal of*

Experimental Social Psychology, 13(3), 279-301. https://doi.org/10.1016/0022-1031(77)90049-X

Situmorang, D. D. B. (2020). Online/cyber counseling services in the COVID-10 outbreak: Are they really new? *Journal of Pastoral Care Counseling, 74*(3), 166-174. https://doi.org/10.1177/1542305020948170

Smail-Crevier, R., Powers, G., Noel, C. & Wang, J. (2019). Health-related internet usage and design feature preference for e-mental health programs among men and women. *Journal of Medical Internet Research, 21*(3), e11224. https://doi.org/10.2196/11224

Smith, B. W., Dalen, J., Wiggins, K., Tooley, E., Christopher, P. & Bernard, J. (2008). The Brief Resilience Scale: Assessing the ability to bounce back. *International Journal of Behavioral Medicine, 15*, 194-200. https://doi.org/10.1080/10705500802222972

Starvaggi, I., Dierckman, C. & Lorenzo-Luaces, L. (2024). Mental health misinformation on social media: Review and future directions. *Current Opinion in Psychology, 56*, 101738. https://doi.org/10.1016/j.copsyc.2023.101738

White, E. & Hanley, T. (2023). Therapist + social media = mental health influencer? Considering the research focusing upon key ethical issues around the use of social media by therapists. *Counseling and Psychotherapy Research, 23*(1), 1-5. https://doi.org/10.1002/capr.12577

Wilhelm, K., Handley, T., McHugh, C., Lowenstein, D. & Arrold, K. (2022). The quality of internet websites for people experiencing psychosis: Pilot expert assessment. *Journal of Medical Internet Research, 6*(4), e28135. https://doi.org/10.2196/28135

Zhang, C. C., Zaleski, G., Kailley, J. N., Teng, K. A., English, M., Riminchan, A. & Robillard, J. M. (2024). Debate: Social media content moderation may do more harm than good for youth mental health. *Child and Adolescent Mental Health, 29*(1), 104-106. https://doi.org/10.1111/camh.12689

Author Note

Correspondence concerning this manuscript should be addressed to Donna W. Nelson, Department of Psychology, Winthrop University, Rock Hill, SC 29732.

Email: nelsond@winthrop.edu.

Appendix A

Experimental Condition: Scientific Summary of Fear of Future Anxiety

Researchers have recently identified a new disorder called Fear of Future Anxiety (FOFA). The main symptom is worry that overwhelming anxiety will occur sometime in the near future.

This disorder is most common in young adults. Estimates are that 65% (or 15%) of college-aged people experience symptoms.

In order to be diagnosed, the individual must experience:

- Fear of future anxiety that causes life disruption for at least four weeks
- Worry and fear about having anxiety in the future that the individual can't control
- Avoidance of thinking about the future out of a fear that those thoughts will create anxiety

This disorder has been shown to drastically interfere with daily life activities for those who have it.

Appendix B

Experimental Condition: Diagnostic Test of Fear of Future Anxiety

Below is a tool to assess symptoms of Fear of Future Anxiety (FOFA). This disorder is most common in young adults. Estimates are that 65% (or 15%) of college-aged people experience symptoms.

The more you agree with these items, the more likely it is that you have symptoms of this disorder.

I often spend time worrying that something is about to trigger my anxiety.

I am not anxious at the moment but I fear that I could become anxious at any time.

I avoid thinking about the future because I worry that it will make me feel anxious.

I am fearful of how the future might make me feel anxious.

I avoid things that I think might trigger my anxiety about the future.

I worry that the future is going to be filled with anxiety.

Journal of Psychological Inquiry
2025, Vol. 29, No. 2, pp. 15-20

PRIMING STRENGTHS ENHANCES COLLEGE STUDENTS' COPING EFFICACY AND EMOTION REGULATION

PORSHA S. NESBITT AND DONNA W. NELSON

WINTHROP UNIVERSITY

Abstract – A growing number of college students face mental health challenges and identifying effective coping strategies is crucial. This study explored the role of a strengths-based intervention in enhancing emotion regulation and coping self-efficacy among college students, particularly those with a history of trauma. Participants (*N*=71) wrote about a prior traumatic event and were then randomly assigned to a control condition or a strengths intervention condition. The intervention emphasized awareness of personal strengths and highlighted their relevance to coping. Next, participants completed emotion ratings and measures of coping self-efficacy. We hypothesized that the strengths intervention would enhance self-efficacy in both coping and emotion regulation and reduce negative affect following trauma recall. Results revealed that participants in the strengths intervention condition, compared to the control group, experienced less negative emotion and more positive emotion. They also reported greater self-efficacy in managing emotions and coping. Additionally, significant differences were found across racial groups, with Black participants reporting higher emotion regulation and coping self-efficacy, compared to their White counterparts. These findings suggest that strengths-based interventions can effectively promote resilience and emotional adaptability in students with prior trauma.

Keywords: trauma, coping, strengths

A growing number of college students are experiencing mental health challenges. One significant contributing factor is that many students enter college with a history of prior trauma (Cusack et al., 2019). The transition to college life, with its academic and social pressures as well as newfound independence, can exacerbate these pre-existing issues (Garett et al., 2017). This study explored strategies to promote effective coping mechanisms for students confronted with trauma reminders. By understanding and addressing the specific needs of these students, colleges can implement supportive measures to foster a healthier and more resilient student population.

Prevalence and Impact of Trauma

Trauma can be defined as any event that is physically, emotionally, mentally, or psychologically threatening or stressful to a person's wellbeing. Trauma can occur once, on multiple occasions, or persistently for an extended period of time (Carr et al., 2018). In a study conducted by Benjet et al. (2016) that tested adults in 26 countries, 70.4% of participants had been exposed to at least one traumatic event in their lifetimes and 30.5% of their participants had been exposed to four or more traumatic events within their lifetimes.

Exposure to trauma can lead to negative long-term effects that detract from wellbeing. Trauma can lead to increased risks of mental health disorders including post-traumatic stress disorder, depression, hallucinations, and anxiety disorders (Kachadourian et al., 2021; Van Overloop et al., 2023). Trauma exposure can also lead people to engage in risky behaviors such as suicidal ideation, alcohol and substance abuse, and high-risk sexual behaviors (Van Overloop et al., 2023). Individuals who have experienced trauma also have higher annual health care costs, more medical visits, and lower health-related outcomes than those with no trauma exposure (Sowder et al., 2018). Indeed, trauma is linked with a variety of health conditions including chronic pain, gastrointestinal disorders, cardiovascular disorders, obesity, diabetes, cancer, and premature death due to toxic stress and lower immune functioning (D'arcy-Bewick et al., 2022; Sowder et al., 2018).

Importance of Coping Strategies

Utilizing effective coping strategies is an important component in helping people mitigate their

own adverse effects of trauma (Viana Machado et al., 2020). Coping strategies are cognitive and behavioral changes made by an individual to manage internal or external demands caused by situations that are taxing or exceed the individual's resources (Gustems-Carnicer & Calderón, 2016). Coping strategies can help individuals manage emotions and reduce stress associated with trauma across their lifespan (Perry & Cuellar, 2021). Moreover, identifying adaptive coping strategies can help mental and behavioral health professionals develop interventions that are more tailored to the needs of their clients to help support them in the most effective way possible (Kachadourian et al., 2021).

Emotion Regulation as Coping

Managing emotions is a fundamental aspect of coping (Troy et al., 2018). Emotion regulation is a type of emotion-focused coping that can be characterized by a person's ability to influence their emotions, when and how they experience emotions, and how they express them (Compas et al., 2014). Adaptive emotion regulation strategies can enhance psychological wellbeing while maladaptive emotion regulation can increase negative affect (Volkaert et al., 2019). Two key strategies are distraction and cognitive reappraisal. Distraction involves deliberately shifting attention away from emotional triggers to focus on something else, to avoid negative thoughts and emotions (Volkaert et al., 2019). Cognitive reappraisal involves reframing a negative event to perceive it in a more positive light (Troy et al., 2018; Volkaert et al., 2019). Developing strong emotion regulation skills is an important component to individuals' mental health and their coping journey (Compas et al., 2014).

Personal Strengths and Coping

Alongside emotion regulation, utilization of personal strengths can significantly enhance coping. Character strengths are positive, innate abilities rooted in a person's thoughts and behaviors, enabling them to operate at their highest potential (Gustems-Carnicer & Calderón, 2016; Proctor et al., 2011). Among a person's character strengths are "signature strengths" which are most closely linked to an individual's identity, sense of self, and well-being (Proctor et al., 2010). Research has shown that using character strengths is associated with higher life satisfaction and greater success in school and work (Park & Peterson, 2009). Moreover, utilizing character strengths has been linked to less risk of depression and anxiety, as well as improved psychological and physiological response to stress (Duan, 2016; Li et al., 2017). However, individuals often fail to fully recognize or appreciate their strengths (Hodges and Clifton, 2004) and this can result in underutilization of these strengths (Peterson & Seligman, 2004).

Research Question and Hypotheses

Although research has supported beneficial mental health effects of both emotion regulation and utilizing strengths, no research has yet examined the direct link between character strengths and emotion regulation. Given that utilization of strengths is associated with better adaptation to stress, we reasoned that personal strengths could serve as a resource to improve self-efficacy in coping and help individuals effectively regulate emotions linked to past trauma. Specifically, we hypothesized that raising the salience of individuals' strengths and highlighting their significance in coping with adverse life events would

- enhance self-efficacy in coping
- enhance self-efficacy in emotion regulation
- promote effective emotion regulation and therefore reduce negative emotion following the recall of trauma

Method

Participants

Participants were 71 students over the age of 18 in introductory and advanced psychology courses at a mid-sized university in the Southeast United States. These students were enrolled in courses that required them to complete a strengths assessment based on the 24 item *Brief Strengths Test* (Peterson & Seligman, 2004) during the first few weeks of their semester. The majority were women (69%) and heterosexual (67%) with an average age of 20.60 (SD = 6.20). The sample included an equal number of participants who identified as White (45%) and Black (45%), while 10% classified themselves as other races. Additionally, 43% self-identified as first-generation college students. Participants completed an anonymous survey either in paper format or online through Qualtrics. All participants received extra course credit in exchange for their participation. Informed consent was obtained to ensure that all participation was voluntary and alternate means of earning extra course credit was offered. Participants were informed about the nature of the survey, including questions related to past traumatic events, and were reminded that they could skip any item they preferred not to answer. Additionally, information about free counseling services was provided to address any distress that might arise from participating in the study.

Procedure and Materials

Trauma Ratings and Recall

After providing informed consent, participants responded to the 19-item *Trauma and Life Events (TALE) Checklist* (Carr et. al, 2018), designed to measure the extent to which they previously encountered various traumatic events such as loss of a parent or sexual assault. Responses were made on a Likert scale ranging from 0 (*never*) to 4 (*happened to me very often*). Next, participants were asked to recall and briefly write about their most traumatic or stressful life event. In accordance with the procedure followed by Troy et al. (2018), they then indicated the extent to which the event was traumatic and the extent to which the event exerts a negative impact on their current life. These ratings were made on a Likert scale ranging from 1 (*not at all*) to 10 (*extremely*).

Experimental Manipulation

Participants were then randomly assigned to a set of instructions outlining either the strengths-intervention condition or a control condition. In the strengths-intervention condition, participants were asked to list their top three signature strengths, as identified by the previous course assignment that required them to complete the 24 item *Brief Strengths Test* (Peterson & Seligman, 2004). This assessment occurred between one and two weeks prior to the study and involved scenario-based questions related to 24 different strengths such as perseverance, leadership, kindness and citizenship. After listing their top strengths, participants were asked to think about and describe how their strengths have helped them cope with adverse experiences in their life. In the control condition, participants were not asked about their strengths. Instead, they were asked to think about and describe the feelings evoked by the traumatic event they wrote about.

Emotion Ratings

Following the recall and writing tasks, participants completed emotion ratings based on Troy et al. (2018). Specifically, they indicated the extent to which they were currently feeling a series of positive and negative emotions such as anger, anxiety, happiness and optimism. Ratings were made on a five-point Likert scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree).*

Self-Efficacy Ratings

Next, participants responded to the 26-item *Coping Self-Efficacy Scale* (Chesney et al., 2006) to measure perceptions of their ability to perform coping behaviors after an adverse event. The prompt specified "When things aren't going well for you, or when you're having problems, how confident or certain are you that you can do the following…" Examples of coping behaviors included "make a plan of action and follow through" and "look for something good in a negative situation". Responses were made on an 11-point Likert scale ranging from 0 (*cannot do at all*) to 10 (*certainly can do*). Participants also completed the 10-item *Regulatory Emotional Self-Efficacy Scale* (Capara et al., 2008) to measure their self-efficacy in managing positive and negative affect through emotion regulation behaviors. The prompt specified "How well can you…". Examples of emotion regulation behaviors included "rejoice over your successes" and "avoid flying off the handle when you get angry." Responses were made on an 11-point Likert scale ranging from 0 (*cannot do at all*) to 10 (*certainly can do*).

Finally, participants responded to a series of demographic items to measure variables such as age, race/ethnicity, gender and socioeconomic status.

Results

Affect Ratings

We conducted independent groups t-tests to examine differences in affect across experimental conditions. We found that participants in the strengths condition experienced fewer negative emotions [$t(67) = 2.70$, $p = .009$, $d = .65$] and more positive emotions [$t(67) = -2.20$, $p = .03$, $d = -.53$] compared to the control group. Additionally, Cohen's d tests indicated that these differences reflected medium effect sizes.

Self-Efficacy Ratings

We conducted independent groups t-tests to examine differences in self-efficacy ratings across experimental conditions. We found that participants in the strengths condition had higher emotion regulation self-efficacy [$t(66) = 2.33$, $p = .02$, $d = -.56$] and higher coping self-efficacy [$t(63) = -2.08$, $p = .04$, $d = -.52$] compared to the control condition. Additionally, Cohen's d tests indicated these differences reflected medium effect sizes. See Table 1 for descriptive statistics across conditions.

Table 1

Means and Standard Deviations for Dependent Variables Across Condition

Dependent Variables	Control Condition		Strengths Condition	
	M	SD	M	SD
Emotion Regulation Self-Efficacy	5.32	1.87	6.19	1.88
Coping Self-Efficacy	4.75	1.8	5.69	1.96
Positive Affect	3.24	.67	3.52	.69
Negative Affect	3.12	.93	2.56	.58

Race Differences

Exploratory analyses revealed significant race differences within our study. We conducted independent groups t-tests comparing affect and self-efficacy for White and Black participants. Compared to their White counterparts, Black participants experienced fewer negative emotions [$t(60) = 3.13$, $p = .003$, $d = .79$] and rated themselves higher in emotion regulation self-efficacy [$t(59) = -2.66$, $p = .01$, $d = -.68$] and coping self-efficacy [$t(58) = -1.92$, $p = .03$, $d = -.49$]. Cohen's d tests indicated these differences reflected medium effect sizes. There were no significant differences in ratings of positive emotions for White and Black participants [$t(60) = -.49$, $p = .63$, $d = -.12$].

SES Differences in Trauma Exposure

Additional exploratory analyses examined differences in trauma across SES levels. A one-way ANOVA yielded a significant main effect indicating that participants who grew up somewhat poor experienced more trauma than other participants [$F(3,60) = 2.69$, $p = .05$, $\eta^2 = .12$]. A Cohen's d test indicated these differences reflected small effect sizes.

Discussion

In this study, we explored the relationship between character strengths, emotion regulation, and self-efficacy in coping. We hypothesized that a strengths-based intervention would enhance coping self-efficacy, improve emotion regulation self-efficacy, and foster effective emotion management after recall of a traumatic experience. In support of our hypotheses, our results showed that compared to the control condition, participants who completed the strengths intervention reported fewer negative emotions and more positive emotions after recalling trauma. These findings suggest that activating and utilizing character strengths can foster effective emotion regulation. Additionally, beyond experiencing a shift toward more positive emotional states, participants in the strengths condition showed increased confidence in their ability to manage emotions and cope with stress. This finding highlights the potential of strengths-based interventions as a useful tool for supporting emotional adaptability, particularly among college students with past trauma exposure.

Our results are consistent with previous research linking personal strengths to improved coping skills, emotional well-being and self-efficacy (Gustems-Carnicer & Calderón, 2016; Proctor et al., 2011; Wallace et al., 2023). Furthermore, our findings align with prior research that demonstrated that reappraisal is an effective strategy for increasing positive emotions and decreasing negative affect following trauma recall (Troy et al., 2018). Similarly, our findings support Bandura's (1997) coping self-efficacy theory, which emphasizes the importance of belief in one's ability to manage stress as a means of improving adjustment. By showing that priming personal strengths enhances both coping efficacy and emotion regulation efficacy, we extend prior work, offering a novel contribution to the body of literature focused on promoting emotional resilience in individuals with trauma.

Additionally, our exploratory analyses revealed significant demographic differences that merit closer examination. Specifically, Black participants reported higher emotion regulation and coping self-efficacy after recalling trauma, compared to their White counterparts. This may indicate that culturally specific mechanisms or different socialization processes exist that produce differences in emotional resilience across cultural groups. Moreover, our participants from lower SES backgrounds experienced more trauma than those from higher SES groups. This reinforces prior research showing that economically disadvantaged populations are more vulnerable to adverse life events (Yuan et al. 2022). In sum, the demographic differences we found in trauma exposure and coping efficacy across groups highlight the need for tailored interventions to address the unique challenges and coping resources of different populations.

Our findings have practical implications. Character strengths are an underexplored resource that can assist in trauma coping and emotional self-efficacy. Strengths-based interventions have the potential to be integrated into college mental health programs. For students with low SES and first-generation college students who may have more trauma exposure compared to other college students, a practical intervention that uses personal strengths can not only help students manage their trauma exposure but also help them navigate typical struggles that college students may face. College and university mental health services can begin to offer personal strengths assessments in their offices and use those results to tailor their interventions to the innate abilities and preferences of their clients. In the broader population, mental health professionals can use strengths-based interventions to promote resilience in people that have experienced trauma in their past.

Our study has both strengths and limitations worthy of consideration. One of the strengths of our study is that we utilized an experimental design to test our strengths intervention against a control condition. We also had a good balance of Black and White participants in our study, which allowed us to consider potential cultural differences that were reflected in our results. However, there were some limitations in our study that may have impacted our findings. For instance, our

sample lacked demographic diversity, as our participants were mostly Black and White young adults. This limits the generalizability of our findings. Moreover, our study utilized self-report measures, which introduces potential bias. Finally, the short-term study design prevented us from assessing long-term effects of our intervention. If we had conducted follow-up assessments of our participants' emotion regulation and coping over time, this may have provided valuable insights.

Future research should examine the long-term effects of strengths interventions on coping self-efficacy and emotion regulation, particularly among individuals with prior trauma. A longitudinal design could provide additional insight into the manner in which such interventions impact coping and adjustment. It could also be beneficial to explore the application of personal strengths to daily hassles to determine if strengths can help people navigate everyday struggles. Future research should also examine strengths-based interventions in more diverse populations including people in different age groups and with different cultural backgrounds and different gender identities. Investigating the underlying causes for the racial and SES differences in emotional resilience and coping observed in our sample would also be valuable.

Conclusion

Many college students have a history of prior trauma and are at risk for mental health challenges. Teaching students how to leverage their personal strengths can be a path to more effective coping and better adjustment. Raising awareness of how personal strengths can facilitate coping may boost self-efficacy in emotion regulation and coping. Furthermore, understanding how to use strengths can help students manage emotional responses when they are reminded of their trauma. By fostering better emotion regulation and coping abilities, strengths-based interventions can support resilience and reduce the long-term effects of trauma. Further research on strength-based interventions may provide valuable insights that can inform student mental health services and promote emotional resilience in individuals with past trauma.

References

Bandura, A. (1997). Self-efficacy: The exercise of control. *Choice Reviews Online*, *35*(3). https://doi.org/10.5860/choice.35-1826

Benjet, C., Bromet, E., Karam, E. G., Kessler, R. C., McLaughlin, K. A., Ruscio, A. M., Shahly, V., Stein, D. J., Petukhova, M., Hill, E., Alonso, J., Atwoli, L., Bunting, B., Bruffaerts, R., & Caldas-de-Almeida, J. M. (2016). The epidemiology of traumatic even exposure worldwide: Results from the World Mental Health Survey Consortium. *Psychological Medicine*, *46*(2), 327–343. https://doi.org/10.1017/S0033291715001981

Carr, S., Hardy, A., & Fornells-Ambrojo, M. (2018). The Trauma and Life Events (TALE) checklist: Development of a tool for improving routine screening in people with psychosis. *European Journal of Psychotraumatology*, *9*(1), 1512265. https://doi.org/10.1080/20008198.2018.1512265

Compas, B. E., Jaser, S. S., Dunbar, J. P., Watson, K. H., Bettis, A. H., Gruhn, M. A., & Williams, E. K. (2013). Coping and emotion regulation from childhood to early adulthood: Points of convergence and divergence. *Australian Journal of Psychology*, *66*(2), 71–81. https://doi.org/10.1111/ajpy.12043

Cusack, S. E., Hicks, T. A., Bourdon, J., Sheerin, C. M., Overstreet, C. M., Kendler, K. S., Dick, D. M., & Amstadter, A. B. (2019). Prevalence and predictors of PTSD among a college sample. *Journal of American College Health*, *67*(2), 123–131. https://doi.org/10.1080/07448481.2018.1462824

D'arcy-Bewick, S., Terracciano, A., Turiano, N., Sutin, A. R., Long, R., & O'Súilleabháin, P. S. (2022). Childhood abuse and neglect, and mortality risk in adulthood: A systematic review and meta-analysis. *Child Abuse & Neglect*, *134*, 105922. https://doi.org/10.1016/j.chiabu.2022.105922

Duan, W. (2016). The benefits of personal strengths in mental health of stressed students: A longitudinal investigation. *Quality of Life Research*, *25*(11), 2879–2888. https://doi.org/10.1007/s11136-016-1320-8

Garett, R., Liu, S., & Young, S. D. (2017). A longitudinal analysis of stress among incoming college freshmen. *Journal of American College Health*, *65*(5), 331–338. https://doi.org/10.1080/07448481.2017.1312413

Gustems-Carnicer, J., & Calderón, C. (2015). Virtues and character strengths related to approach coping strategies of college students. *Social Psychology of Education*, *19*(1), 77–95. https://doi.org/10.1007/s11218-015-9305-y

Hodges, T. D., & Clifton, D. O. (2004). Strengths-Based Development in Practice. In P. A. Linley & S. Joseph (Eds.), *Positive psychology in practice* (pp. 256–268). John Wiley & Sons, Inc.. https://doi.org/10.1002/9780470939338.ch16

Kachadourian, L. K., Harpaz-Rotem, I., Tsai, J., Southwick, S., & Pietrzak, R. H. (2021). Mindfulness as a mediator between trauma exposure and mental health outcomes: Results from the national health and resilience in veterans study. *Psychological Trauma: Theory,*

Research, Practice, and Policy, *13*(2), 223–230. https://doi.org/10.1037/tra0000995

Li, T., Duan, W., & Guo, P. (2017). Character strengths, social anxiety, and physiological stress reactivity. *PeerJ*, *5*, 5626–5636. https://doi.org/10.7717/peerj.3396

Park, N., & Peterson, C. (2009). Character strengths: Research and practice. *Journal of College and Character*, *10*(4). https://doi.org/10.2202/1940-1639.1042

Perry, Y., & Cuellar, M. J. (2021). Coping methods used by college undergraduate and graduate students while experiencing childhood adversities and traumas. *Journal of Child & Adolescent Trauma*, *15*, 451–459. https://doi.org/10.1007/s40653-021-00371-z

Peterson, C., & Seligman, M. E. P. (2004). *Character Strengths and Virtues: A Handbook and Classification*. Oxford University Press.

Proctor, C., Maltby, J., & Linley, P. A. (2010). Strengths use as a predictor of well-being and health-related quality of life. *Journal of Happiness Studies*, *12*(1), 153–169. https://doi.org/10.1007/s10902-009-9181-2

Sowder, K. L., Knight, L. A., & Fishalow, J. (2018). Trauma exposure and health: A review of outcomes and pathways. *Journal of Aggression, Maltreatment & Trauma*, *27*(10), 1041–1059. https://doi.org/10.1080/10926771.2017.1422841

Troy, A. S., Shallcross, A. J., Brunner, A., Friedman, R., & Jones, M. C. (2018). Cognitive reappraisal and acceptance: Effects on emotion, physiology, and perceived cognitive costs. *Emotion*, *18*(1), 58–74. https://doi.org/10.1037/emo0000371

Van Overloop, E., Arms-Chavez, C., Carol, R. N., & LoBello, S. G. (2023). Effects of adverse childhood experiences and chronic health conditions on current depression. *Community Mental Health Journal*, *59*, 1208–1216. https://doi.org/10.1007/s10597-023-01103-3

Viana Machado, A., Volchan, E., Figueira, I., Aguiar, C., Xavier, M., Souza, G. G. L., Sobral, A. P., de Oliveira, L., & Mocaiber, I. (2020). Association between habitual use of coping strategies and posttraumatic stress symptoms in a non-clinical sample of college students: A Bayesian approach. *PLOS ONE*, *15*(2), 1–17. https://doi.org/10.1371/journal.pone.0228661

Volkaert, B., Wante, L., Van Beveren, M.-L., Vervoort, L., & Braet, C. (2019). Training adaptive emotion regulation skills in early adolescents: The effects of distraction, acceptance, cognitive reappraisal, and problem solving. *Cognitive Therapy and Research*, *44*, 678–696. https://doi.org/10.1007/s10608-019-10073-4

Wallace, N., Parente, A., & McGrath, R. E. (2023). Character Strengths as Moderators of the Relationship Between Adverse Childhood Experiences and Negative Health Outcomes. *International Journal of Applied Positive Psychology*, *8*, 429–452. https://doi.org/10.1007/s41042-023-00097-3

Yuan, N. P., Kram-Brooks, N. A. Z., Ellsworth-Kopkowski, A., & Hamby, S. (2022). Adverse experiences and positive and negative responses among Appalachian young people. *Journal of Rural Mental Health*, *46*(4), 237–248. American Psychological Association. https://doi.org/10.1037/rmh0000216

Author Note

This research was supported by the Ronald E. McNair Post-Baccalaureate Achievement Program.

Correspondence concerning this article should be addressed to Dr. Donna Nelson, Department of Psychology, Winthrop University, Rock Hill, SC 29732. Email: nelsond@winthrop.edu.

Journal of Psychological Inquiry
2025, Vol. 29, No. 2, pp. 21-28

EMOTION REGULATION AND MENTAL HEALTH OUTCOMES BETWEEN SOCIOECONOMIC CLASSES AMONG UNIVERSITY STUDENTS

REBECCA REARDON, FEATHER CHOKBENGBOUN, SAMMY GRAY, AND YESEUL NAM

UNIVERSITY OF CENTRAL ARKANSAS

Abstract – This study investigated emotion regulation and its impact on mental health. This modified replication of Troy et al. (2017), showed that socioeconomic status (SES) moderated the relationship between cognitive reappraisal ability (CRA) and depression while ruling out life stress. Our first hypothesis replicated the original study, the second predicted the interaction effect of CRA and SES on satisfaction with life (SWL), and the third examined whether SES moderates emotional suppression ability (ESA) and depression. Undergraduate students (N = 118) completed an online survey. Our findings were inconsistent with those of Troy et al. (2017). SES did not moderate between CRA and depression, CRA and SWL, or ESA and depression. The results indicated the benefits of emotion regulation on mental health may apply regardless of their SES.

Keywords: emotional regulation, cognitive reappraisal ability, socioeconomic status, mental health, emotion suppression, depression

Every human being experiences a vast multitude of differing emotions throughout each day, but how one chooses to express or suppress these emotions may have a larger impact on our mental well-being than some may think. Some individuals are inherently more effective in regulating their emotions toward specific situations. Specifically, how an individual chooses to regulate their emotions and interpret their thoughts can affect negative mental health outcomes, such as depression and personal dissatisfaction (Troy et al., 2013). Our research represents a tailored replication of an investigation conducted by Troy et al. (2017), which delved into the intricate relationship between cognitive reappraisal ability (CRA), socioeconomic status (SES), life stress, and symptoms of depression. Building upon the foundations of the Troy et al. (2017) study, our research seeks to contribute to the understanding of how SES might moderate the associations between three sets of variables: CRA and depression, CRA and satisfaction with life (SWL), and emotion suppression ability (ESA) and depression.

The effect of CRA and SES on depression (Hypothesis 1: direct replication)

Troy et al. (2017) examined the interplay between CRA, SES, and depression. SES can be defined as a person's relative social standing determined by their household income, educational achievements, and overall assets (Kraus et al., 2012). A person's ownership over these assets can determine their ability to control their external environment and overall satisfaction. Individuals with lower SES experience heightened stress but possess fewer coping resources, leading to decreased psychological well-being (Finkelstein et al., 2007; Navarro-Carrillo et al., 2020). Consequently, they often feel less control over their lives with their limited resources (Gallo et al., 2005; Kraus et al., 2009). Those who lack these benefits are subject to harsher financial circumstances, having to rely on mental fortitude and emotional revaluation to bring happiness into their lives (Kraus et al., 2012).

As well as this, CRA, defined as the capacity to reframe emotional situations to alter their emotional impact (Goldin et al., 2012), was examined in relation to SES and its impact on a person's negative thoughts and feelings toward themselves or their environment otherwise known as depression (Troy et al., 2017). The findings revealed a moderation effect which confirmed their hypothesis, suggesting that higher CRA was significantly associated with having lower depressive symptoms in individuals with lower SES, although the same relationship was notably weaker in individuals with higher SES after controlling for life stress. Furthermore, cognitive reappraisal is a dependable emotion regulation strategy, particularly helpful in uncontrollable situations (Roos & Bennett, 2022). According to Brandao et al.

(2023), individuals who struggle to regulate their emotions are at greater risk of developing depressive symptoms. This indicates that emotion regulation can be critical for mental health, especially when individuals lack sufficient material resources to control their situations.

Especially in cases of low SES where individuals lack control over their life trajectory (Gallo et al., 2005; Kraus et al., 2009), CRA could play a critical role in determining mental health, whereas it may have less of an effect on high SES individuals. Research conducted by Cohen et al. (1983) contributes to our understanding of life stress as a potential confounder in studies exploring the relationship between emotion regulation and psychological health issues like depression. Based on these findings, we aim to expand the research and increase generalizability through direct and modified replication. Our first hypothesis is a direct replication of Troy et al. (2017), investigating if there is truly an interaction between an individual's CRA and SES on depressive symptoms after controlling for life stress. By controlling for life stress, we aim to build on insights from Cohen et al. (1983) and isolate the specific impacts of SES and emotion regulation techniques on mental health.

The effect of CRA and SES on Life Satisfaction (Hypothesis 2: expansion)

In our second hypothesis, our focus shifted to whether SES moderates the link between CRA and SWL while ruling out the potentially confounding effects of life stress. In this context, SWL is an individual's cognitive evaluation of their own life, measuring whether they feel that their life is meaningful and holds significant value (Mehrotra et al., 2018). Taking this into account, CRA in specific contexts is also linked to positive psychological effects, such as higher SWL, a greater sense of self-esteem, and improved overall well-being (Gross, 1998; Troy et al., 2013). Also, SES can be related to SWL: A meta-analysis showed that individuals with high SES were more likely to be satisfied with life (Moreno-Agostino et al., 2021). High SES individuals are less concerned about their finances, which leads to an increase in life satisfaction, while those with lower SES are faced with financial difficulties and are more dissatisfied with their circumstances and lives (Ansah et al., 2022). Since high SES individuals have more control over their lives (Gallo et al., 2005; Kraus et al., 2009), they do not necessarily need to have the same level of control over their emotions, so the ability to control their emotions may not be related to their SWL in the same way it would be for those of low SES (Heidemeier, 2017). Therefore, we expand the findings of Troy et al. (2017), and predict the interaction between CRA and SES on SWL: upper-class individuals will not have a relationship between CRA and SWL, while lower-class individuals will have a positive relationship between the two.

The effect of ESA and SES on Depression (Hypothesis 3: expansion)

In our third hypothesis, we seek to modify and extend Troy et al. (2017) by focusing on ESA instead of CRA, while ruling out the confounding effects of life stress. We aim to shed light on a different facet of emotion regulation (i.e., ESA), as well as contribute valuable insights into how various emotion regulation strategies might interact with SES to influence mental health outcomes. In this regard, ESA refers to an individual's capacity to consciously inhibit or restrain the expression of their emotions (Gross, 1998). In other words, ESA is the capability of an individual to suppress their emotions and feelings when necessary, not the frequency with which they are suppressed. Gross and John (2003) provide valuable insights into the mechanisms of emotion regulation. Their Emotion Regulation Questionnaire (ERQ), which includes subscales for both cognitive reappraisal and suppression, provides a foundation for investigating ESA. Gross and John (2003) demonstrated that individuals who frequently employ emotion suppression strategies may experience negative psychological consequences, such as increased stress and depressive symptoms. While their findings demonstrate the role that suppression frequency plays in health outcomes, our interest lies in the effects of suppression ability. Snibbe and Markus (2005) explored how educational attainment, closely related to SES, affects agency and choice in individuals. They found that higher SES individuals tend to perceive themselves as having more control over their life choices and situations, whereas lower SES individuals often feel more constrained.

Based on these findings, we expect the strength of ESA's effect on depressive severity to be moderated by each individual's SES. While the act of suppressing emotions can either help or hurt depending on the context, we posit that the ability to suppress and control one's emotions is a universally beneficial skill (Bonanno et al., 2004; Burton et al., 2016). However, that skill may be more beneficial for those who have relatively less agency and control over their standing in the socioeconomic landscape, as they often cannot afford to let their emotions get the best of them. For high SES individuals, we hypothesize that ESA and depression will have a weak negative relationship, but for low SES individuals, we hypothesize a strong negative relationship.

Overview of Current Research

Our study seeks to understand how individuals from different socioeconomic backgrounds navigate the challenges of emotion regulation strategies and their implications. The sample for our study consists of 118 university students at University of Central Arkansas (UCA). Based on the literature, we expect that SES may exert varying influences on the relationships between our variables, with the strength of these relationships contingent upon the individual's SES. This study is part of the Collaborative Replication and Education Project (CREP) and was conducted for a graded class assignment at UCA. All research data and materials are shared in Open Science Framework (https://osf.io/cu6p4/). The replication plan and hypothesis were registered before data collection (https://osf.io/t6hfb).

Method

Participants

A total of 118 participants were recruited for this study. As part of CREP, the sample size had been determined to be more than 100 university participants for each university in advance. Those who were under the minimum age of 18, who finished the study in less than 3 minutes, or who did not fully complete the survey were excluded from the final data analysis as pre-registered. Participants were adults with ages ranging from 18-47 (M = 21.24, SD = 4.50). Participants were predominantly women, with 83.9% identifying as female, 12.7% identifying as male, 2.5% identifying as non-binary, and 0.8% preferred not to say. The participants' racial identity consisted of 76.3% being Caucasian, 12.7% being African American, 5.1% being multiple races, 3.4% identifying as other, 1.7% being Asian, and 0.8% preferred not to say. Their ethnic background included 84.7% identifying as not Hispanic/Latino, 6.8% identifying that they were Hispanic/Latino, and 4.2% preferring not to say.

Procedure

This current study was approved by the Institutional Review Board through the authors' institution, as we have replicated the study conducted by Troy et al., (2017). Participants were recruited through the university's SONA system, through convenience sampling, and were provided online consent to participate in this study. Participants completed an online survey that took approximately 15 minutes to complete, which started with demographic questions asking for gender, age, racial/ethnic identity, ethnicity, student status, and current annual family income (SES). After answering the demographic questions, participants answered questions regarding CRA, habitual cognitive reappraisal use, depressive symptoms, and life stress using the original scales from Troy et al. (2017), along with the additional scales of SWL, ESA, and habitual emotion suppression use (HESU) from the current study. Upon completion of the online survey, participants were thanked, debriefed, and compensated for their time with SONA credits.

Measures

Socioeconomic Status (SES)

We measured SES at the beginning of the survey via the demographic questions. We used the exact question and scale to replicate Troy et al. (2017). We asked participants what their family income was with this question "How high is your current annual family income?", rated on a 1 ("$10,000 or less") to 12 ("$200,000 or above") scale (M = 21.24, SD = 3.50), as a measurement for SES (Table 1). Many individuals find it hard to determine and report their exact annual family income level, therefore we asked them to score it using a series of estimated ranges, as was done by Troy et al. (2017).

Table 1

Distribution of Annual Family Income

Annual Family Income	% of Total
$10,000 or less	9.3 %
$10,001-$20,000	6.8 %
$20,001-$30,000	7.6 %
$30,001-$40,000	10.2 %
$40,001-$50,000	7.6 %
$50,001-$60,000	10.2 %
$60,001-$70,000	7.6 %
$70,001-$80,000	4.2 %
$80,001-$90,000	5.1 %
$90,001-$100,000	7.6 %
$100,001-$200,000	17.8 %
$200,000 or more	5.9 %

Cognitive Reappraisal Ability (CRA)

We measured CRA with an 8-item measure that was adapted from the Emotion Regulation Questionnaire created by Gross and John (2003), as used in Troy et al. (2017). The questions were modified to measure an

individual's ability to use cognitive reappraisal with strong reliability (α = .90). Participants were asked to rate the questions on a scale of 1 (*strongly disagree*) to 7 (*strongly agree*), with some examples including: "When I really want to, I am very capable of changing what I'm thinking about when I want to feel less negative emotion (such as sadness or anger)," and "When I really want to, I am very capable of reconsidering what relevance the situation really has for me if a situation is likely to make me upset." Higher scores indicated a greater ability to reappraise one's cognitions about emotions, and the total CRA score was measured by calculating the mean of all 8 questions (M = 4.78, SD = 1.14).

Habitual Cognitive Reappraisal Use (HCRU)

Habitual cognitive reappraisal use was measured using a subscale of the ERQ (Gross & John, 2003), and consists of six questions which are answered on a scale of 1 (*strongly disagree*) to 7 (*strongly agree*), as used in Troy et al. (2017). Some questions asked were: 'When I want to feel less negative emotion (such as sadness or anger), I change what I'm thinking about," and "I control my emotions by changing the way I think about the situation I'm in." Participants with higher scores in this category indicated they were more likely to use cognitive reappraisal as an emotion regulation strategy. The total score was measured for each person by calculating the mean of all six questions (M = 4.84, SD = 1.09, α = .89).

Depressive Symptoms

Depressive symptoms were measured using a 5-item questionnaire adapted from the Center for Epidemiologic Studies Depression Scale (Radloff, 1977). Participants were asked to report symptoms of depression that they noticed within the past week, on a scale of 0 (*rarely or none of the time*) to 3 (*most of the time*). Examples of symptoms inquired in this section of the survey include: "I felt depressed," and "I had crying spells." Higher scores indicated a greater presence of depressive symptoms, and the total score was calculated by taking the sum of all 5 questions (M = 10.86, SD = 3.63, α = .75).

Life Stress

Life stress was measured using a 4-item questionnaire that was adapted from the Perceived Stress Scale (Cohen et al., 1983). Participants were asked to report their perceptions of stress over the past two years, with items being rated on a scale of 1 (*never*) to 5 (*very often*). Some examples of questions included are: "In the past two years, how often have you felt that things were going your way? (R)" and "In the past two years, how often have you felt that you were unable to control the important things in your life?" Higher scores indicated a greater presence of life stress in individuals, and the total score was assessed by taking the mean of all 4 questions (M = 3.04, SD = 0.71, α = .74).

Satisfaction with Life (SWL)

We measured SWL using the 5-item SWL scale questionnaire created by Diener et al. (1985) with participants being asked how content they are with life, on a scale of 1 (*strongly disagree*) and 7 (*strongly agree*). Some examples of questions included are: "In most ways, my life is close to ideal," and "If I could live my life over, I would change almost nothing." Higher scores of SWL indicated that the individual was overall more satisfied with the circumstances of their life, and the total score was then assessed by taking the mean of all 5 questions (M = 4.09, SD = 1.23, α = .85).

Emotion Suppression Ability (ESA)

We measured ESA by using a 4-item scale, adapted from Gross and John's (2003) ERQ, with participants being asked to report their ability to suppress emotions on a scale of 1 to 7 (1 being *strongly disagree* and 7 being *strongly agree*). We adjusted the items similarly to CRA to measure the emotion suppression ability, rather than the frequency of their use. Some examples of questions included are: "When I really want to, I am capable of keeping emotions to myself," and "When I really want to, I am capable of controlling my emotions by not expressing them." Higher scores indicated a greater ability to suppress one's emotions when necessary, and the total score was assessed by taking the mean of all 4 questions (M = 4.96, SD = 1.26, α = .81).

Habitual Emotion Suppression Use (HESU)

We measured HESU using a questionnaire taken directly from Gross and John's (2003) ERQ and consisted of 4 questions. Participants were asked to report their HESU on a scale of 1 (*strongly disagree*) to 7 (*strongly agree*). Example questions include: "I control my emotions by not expressing them," and "I keep my emotions to myself." Higher scores indicated more frequent or habitual use of emotion suppression in the surveyed individuals, and the total score was assessed by taking the mean of all 4 questions (M = 3.89, SD = 1.33, α = .78).

Table 2

Correlations for Study Variables

Variable	1	2	3	4	5	6	7
1. CRA	—						
2. ESA	.13	—					
3. SES	.15	.10	—				
4. Depressive symptoms	-.34***	-.07	-.10	—			
5. SWLS	.46***	-.01	.09	-.49***	—		
6. HCRU	.74***	.09	-.03	-.33***	.44***	—	
7. HESU	-.13	.52***	-.11	.17	-.20*	-.07	—
8. Life stress	-.39***	-.00	-.12	.53***	-.56***	-.36***	.20*

*$p < .05$. **$p < .01$. ***$p < .001$

Results

As a first analysis, we examined the associations between all of the measures in the study. The results of these Pearson's correlations are reported in Table 2.

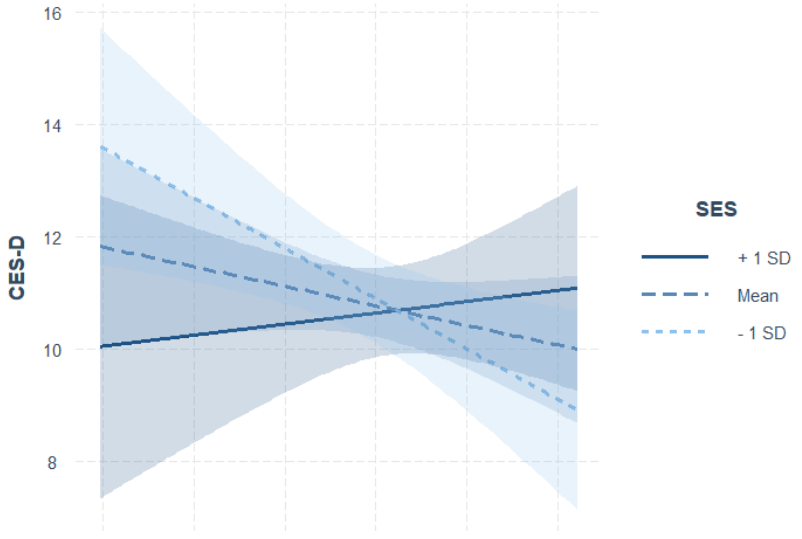
For the interaction analyses, the continuous variables in the regression models were mean-centered. To control for other variables, we included them as predictors in the regression models.

To examine hypothesis 1, we conducted a multiple regression to test the interaction between SES and CRA on depression, which was not significant after controlling for life stress, β = 0.12, t = 1.64, p = .104, 95% CI = [-0.02, 0.26]. These findings were inconsistent with Troy et al. (2017), as SES did not moderate the relationship between CRA and depression. That said, when we examined the simple slopes by SES, the patterns were consistent with the hypothesis. For those with higher SES, simple slope analyses indicated that CRA was not associated with depressive symptoms, β = -0.02, t = -0.04, p = .96. However, for those lower in SES, higher CRA was negatively associated with depression, β = -0.84, t = -2.45, p = .02. For lower SES participants, when they perceived that they are good at cognitive reappraisal, they reported less depression.

We added demographic variables (age, sex, and race), and habitual reappraisal use individually as control variables to the model described above by Troy et al. (2017). Gender was dummy-coded (1: women, 0: men and other options). When we controlled for life stress and gender, the interaction became significant (Figure 1), β = 0.15, t = 2.12, p = .036, 95% CI = [0.01, 0.30]. As predicted by the hypothesis, for lower SES, higher CRA was significantly related to lower depression, β = -1.05, t = -2.77, p = .007, 95% CI = [-1.81, -0.30]. And it was not significant for higher SES, β = 0.31, t = 0.72, p = .471, 95% CI = [-0.55, 1.18]. Also, when life stress and HCRU were controlled for, the interaction was marginally significant, β = 0.12, t = 1.76, p = .081, 95% CI = [-0.02, 0.27]. For lower SES, CRA was not related to depression, β = -0.68, t = -1.48, p = .143, 95% CI = [-1.60, 0.23]. And it was not significant for higher SES, β = 0.43, t = 0.80, p = .425, 95% CI = [-

Figure 1

The Interaction between CRA and SES on Depressive Symptoms after Controlling for Life Stress and Gender



0.64, 1.51]. In other models, the interaction between CRA and SES was not significant.

When assessing hypothesis 2, we conducted a multiple regression to evaluate the interaction between CRA and SES on SWL, which was not significant after controlling for life stress, $\beta = -0.01$, $t = -0.24$, $p = .812$, CI = [-0.05, 0.04]. However, the main effect of CRA was significant, which indicates that higher CRA was related to higher satisfaction with life, $\beta = 0.31$, $t = 3.50$, $p < .001$, CI = [0.13, 0.48]. That being said, SES was not found to be a moderator between CRA and SWL, and there was a positive association between the two variables regardless of one's SES. When we examined the effect of demographics or HCRU individually, the interactions between CRA and SES were still not significant.

To examine hypothesis 3, we conducted a multiple regression to test the interaction between SES and ESA on depression and found that it was not significant after controlling for life stress, $\beta = 0.11$, $t = 1.55$, $p = .124$, CI = [-0.03, 0.25]. For individuals with higher SES, simple slope analyses indicated that ESA was not associated with depressive symptoms, $\beta = 0.23$, $t = 0.65$, $p = .52$. For individuals with lower SES, higher ESA was marginally associated with having fewer depressive symptoms, $\beta = -0.54$, $t = -1.70$, $p = .09$. Therefore, SES did not moderate the relationship between ESA and depressive symptoms. When we tested the effect of demographics or HESU individually, the interactions between ESA and SES were still not significant.

Discussion

Our goal for this present study is to measure the moderating effects of SES on the link between CRA and depression, CRA and SWL, and ESA and depression while ruling out the confounding effects of life stress. For each relationship, we predicted that SES would be a moderator between the variables. However, the results of this study did not support our hypotheses as we found that SES was not a moderating factor between any of the variables used.

Implications of our Study

In general, our study result did not replicate the result of Troy et al. (2017). That said, we discovered some evidence supporting the direct replication hypothesis. There was the significant interaction between CRA and SES on depression as expected when controlled for life stress and gender. Specifically, when controlling for life stress and gender, we found that individuals who perceived themselves to be effective at changing their emotions reported lower levels of depression, but this effect was primarily observed in lower SES participants. This suggests that the ability to regulate emotions through CRA may play a crucial role in mitigating depressive symptoms, particularly for those facing socioeconomic difficulties, as consistent with previous research (Chen & Miller, 2012). Also, given the research about the gender differences in the emotion regulation ability (e.g., Min et al., 2023), controlling for gender might have been able to show the predicted relationship more clearly. Additionally, we found the marginally significant interaction between CRA and SES on depression when controlled for life stress and HCRU.

Furthermore, we found that higher CRA was associated with fewer depressive symptoms, indicating that those who feel more capable of employing cognitive reappraisal strategies experience fewer symptoms of depression. These findings are consistent with previous studies, showing that better emotion regulation strategies can help overall mental well-being (Roos & Bennet, 2022). We also observed a correlation between higher CRA and greater SWL across all SES groups, suggesting that effective emotional regulation may enhance overall life satisfaction, regardless of socioeconomic circumstances. That being said, Troy et al. (2013) and Gross (1998), also found that CRA is linked to positive psychological effects, one being SWL. Interestingly, we found no relationship between ESA and depression, indicating that the mechanisms underlying these variables may differ.

Strengths of our Study

Our study provides valuable insights into the role of CRA in the context of SES and depression. Our findings suggest that higher CRA is associated with reduced depressive symptoms across different socioeconomic backgrounds. Future studies should aim to replicate these results in more diverse and representative samples to better understand the relationship between emotion regulation strategies, SES, and mental health. Through our systematic replication and adaptation of the Troy et al. (2017) research, we hope to enrich the existing literature on emotion regulation by offering a comprehensive view of how SES and emotion regulation strategies influence depressive symptoms and satisfaction with life. Our findings carry the potential to illuminate the varying experiences of individuals from diverse socioeconomic backgrounds and their strategies for managing emotional distress. Ultimately, we hope that the knowledge generated from this study can inform interventions and support systems tailored to the unique needs of different socioeconomic groups, promoting mental health and well-being for all.

Limitations and Future Research Directions

The discrepancies between our findings and those of Troy et al. (2017) highlight the influence of participant demographics. The inconsistent findings

could be due to the focus on female participants in Study 2 in Troy et al. (2017), which can alter the significant interaction between CRA and SES when not taking the male population into account. As well as this, Troy et al. (2017) study did not focus on college students which can lead to contrasting results as college students have more stressors in life while juggling school work and jobs.

One of the limitations of our research was the sample size of the population tested, leading to a lack of variation, which consisted entirely of students from a university in the West South Central region of the U.S., with the vast majority identifying as female (83.9%), severely limiting the sample's scope and diversity. This lack of diversity in our population limits the generalizability of our findings and raises concerns about whether the results can be applied to other populations or contexts. Future research would benefit from incorporating a more diverse population sample inclusive of gender and ethnicity. Also, the cross-sectional design of our study, meaning that all of the data was collected at one point in time rather than longitudinally collecting data repeatedly over a longer period, may have restricted our ability to draw conclusions regarding the relationships between variables.

In addition, the SES measurement we used from Troy et al. (2017), had individuals report their average annual family income which was limited to a series of broad categories, all of which would have to be mentally estimated and selected on the spot with little time to prepare a thoughtful response. Notably, out of the 12 SES categories available for selection, almost 1/5 of our sample (17.8%) selected the "$100,001 - $200,000" response. This is an incredibly broad estimate, with a range of nearly one hundred thousand dollars, which could hinder our analysis concerning SES moderation. It may benefit future researchers to include a more thorough measure of SES among participants, potentially incorporating past and present opportunities for education, housing, employment, and social relationships or organizations (Antonoplis, 2023). Moreover, examining other potential moderators, such as cultural or religious factors (Chesser et al., 2018; Nam et al., 2018), can contribute to a more comprehensive understanding of these relationships.

References

Ansah, K., Dey, N. E. Y., Adade, A. E., & Agbadi, P. (2022). Determinants of life satisfaction among Ghanaians aged 15 to 49 years: A further analysis of the 2017/2018 Multiple Cluster Indicator Survey. *PLoS ONE, 17*(1). 1-18. https://doi.org/10.1371/journal.pone.0261164

Antonoplis, S. (2023). Studying socioeconomic status: Conceptual problems and an alternative path forward. *Association for Psychological Science, 18*(2). 275-292. https://doi.org/10.1177/17456916221093615

Bonanno, G. A., Papa, A., Lalande, K., Westphal, M., & Coifman, K. (2004). The importance of being flexible: The ability to both enhance and suppress emotional expression predicts long-term adjustment. *Psychological Science*, *15*(7), 482–487. https://doi.org/10.1111/j.0956-7976.2004.00705.x

Brandão, T., Brites, R., Hipólito, J., & Nunes, O. (2023). Perceived emotional invalidation, emotion regulation, depression, and attachment in adults: A moderated-mediation analysis. *Current Psychology, 42*(18), 15773-15781. https://doi.org/10.1007/s12144-022-02809-5

Burton, C. L., & Bonanno, G. A. (2016). Measuring ability to enhance and suppress emotional expression: The flexible regulation of emotional expression (FREE) scale. *Psychological Assessment*, *28*(8), 929–941. https://doi.org/10.1037/pas0000231

Chen, E. & Miller, G. E. (2012). "Shift-and-persist" strategies: Why low socioeconomic status isn't always bad for health. *Perspectives of Psychological Science*, *7*(2), 135-158. https://doi.org/10.1177/1745691612436694

Chesser, S., Swanson, S., Garey, E., & Hood, R. W. (2018). Religious and non-religious predictors of life satisfaction in a sample of American college students. *Mental Health, Religion & Culture*, *21*(4), 418–428. https://doi.org/10.1080/13674676.2018.1504905

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior*, *24*(4), 385–396. https://doi.org/10.2307/2136404

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, *49*(1), 71–75. https://doi.org/10.1207/s15327752jpa4901_13

Finkelstein, D. M., Kubzansky, L. D., Capitman, J., & Goodman, E. (2007). Socioeconomic differences in adolescent stress: The role of psychological resources. *Journal of Adolescent Health, 40*(2), 127-134. https://doi.org/10.1016/j.jadohealth.2006.10.006

Gallo, L. C., Matthews, K. A., Bogart, L. M., & Vranceanu, A. M. (2005). Socioeconomic status, resources, psychological experiences, and emotional responses: A test of the reserve capacity model. *Journal of Personality and Social Psychology, 88*(2), 386–399. https://doi.org/10.1037/0022-3514.88.2.386

Gross, J. J. (1998). Antecedent- and response-focused emotion regulation: Divergent consequences for

experience, expression, and physiology. *Journal of Personality and Social Psychology*, *74*, 224–237. https://doi.org/10.1037/0022-3514.74.1.224

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*(2), 348–362. https://doi.org/10.1037/0022-3514.85.2.348

Goldin, P. R., Ziv, M., Jazaieri, H., Werner, K., Kraemer, H., Heimberg, R. G., & Gross, J. J. (2012). Cognitive reappraisal self-efficacy mediates the effects of individual cognitive-behavioral therapy for social anxiety disorder. *Journal of Consulting and Clinical Psychology*, *80*(6), 1034–1040. https://doi.org/10.1037/a0028555

Heidemeier, H. (2017). Are economic conditions related to nonnormative life satisfaction development? Evaluating the relative impact of economic conditions, personality, and subjective health. *European Journal of Social Psychology*, *47*(4), 383–398. https://doi.org/10.1002/ejsp.2227

Kraus, M. W., Piff, P. K., & Keltner, D. (2009). Social class, sense of control, and social explanation. *Journal of Personality and Social Psychology, 97*(6), 992–1004. https://doi.org/10.1037/a0016357

Kraus, M. W., Piff, P. K., Mendoza-Denton, R., Rheinschmidt, M. L., & Keltner, D. (2012). Social class, solipsism, and contextualism: How the rich are different from the poor. *Psychological Review*, *119*(3), 546–572. https://doi.org/10.1037/a0028756

Mehrotra, B., Pandey, N. M., Kumar, P., Sinha, S., & Tiwari, S. C. (2018). Life satisfaction and its determinants in rural aging population of Lucknow, India. *Journal of Psychosocial Research, 13*(1), 33–41. https://doi.org/10.32381/JPR.2018.13.01.4

Min, J., Koenig, J., Nashiro, K., Yoo, H. J., Cho, C., Thayer, J. F., & Mather, M. (2023). Sex differences in neural correlates of emotion regulation in relation to resting heart rate variability. *Brain Topography, 36*(5), 698-709. https://doi.org/10.1007/s10548-023-00974-9

Moreno-Agostino, D., de la Fuente, J., Leonardi, M., Koskinen, S., Tobiasz-Adamczyk, B., Sánchez-Niubò, A., Chatterji, S., Haro, J. M., Ayuso-Mateos, J. L., & Miret, M. (2021). Mediators of the socioeconomic status and life satisfaction relationship in older adults: A multi-country structural equation modeling approach. *Aging & Mental Health, 25*(3), 585–592. https://doi.org/10.1080/13607863.2019.1698513

Nam, Y., Kim, Y.-H., & Tam, K. K.-P. (2018). Effects of emotion suppression on life satisfaction in Americans and Chinese. *Journal of Cross-Cultural Psychology*, *49*(1), 149-160. https://doi.org/10.1177/0022022117736525

Navarro-Carrillo, G., Alonso-Ferres, M., Moya, M., & Valor-Segura, I. (2020). Socioeconomic status and psychological well-being: Revisiting the role of subjective socioeconomic status. *Frontiers in Psychology, 11*, 1-15. https://doi.org/10.3389/fpsyg.2020.01303

Radloff, L. S. (1977). The ces-d scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, *1*(3), 385-401. https://doi.org/10.1177/014662167700100306

Roos, L. G., & Bennett, J. M. (2023). Reappraisal and health: How habitual reappraisal and reappraisal ability interact to protect against life stress in young adults. *Emotion, 23*(5), 1360–1372. https://doi.org/10.1037/emo0001154

Snibbe, A. C., & Markus, H. R. (2005). You can't always get what you want: Educational attainment, agency, and choice. *Journal of Personality and Social Psychology*, *88*(4), 703–720. https://doi.org/10.1037/0022-3514.88.4.703

Troy, A. S., Ford, B. Q., McRae, K., Zarolia, P., & Mauss, I. B. (2017). Change the things you can: Emotion regulation is more beneficial for people from lower than from higher socioeconomic status. *Emotion, 17*(1), 141–154. https://doi.org/10.1037/emo0000210

Troy, A. S., Shallcross, A. J., & Mauss, I. B. (2013). A person-by-situation approach to emotion regulation: Cognitive reappraisal can either help or hurt, depending on the context. *Psychological Science*, *24*(12), 2505–2514. https://doi.org/10.1177/0956797613496434

Author Note

This research was conducted as part of a course research project with guidance from a faculty mentor, Dr. Yeseul Nam.

Correspondence concerning this article should be addressed to Dr. Yeseul Nam, University of Central Arkansas, Conway, AR, 72035; ynam@uca.edu

Journal of Psychological Inquiry
2025, Vol. 29, No. 2, pp. 29-37

VIDEO VERSUS TEXT SPOILERS AND ENJOYMENT OF NARRATIVES

CONNER K. RYAN, KEITH WYLIE, AND SAMUEL R. BROWN

EMPORIA STATE UNIVERSITY

Abstract – Common knowledge suggests that spoilers ruin stories for people—but is this really the case? In the current study, we investigated the effect of spoilers and narrative formats (video vs. text) on college students' enjoyment of narratives. In this 2 (format: text, video) × 2 (spoiler: present, not) factorial study, Qualtrics randomly assigned 63 undergraduate students to one of four groups: short film without spoiler, short film with spoiler, text without spoiler, or text with spoiler. After either reading the short story or watching the short film, participants filled out a survey to measure their perceived enjoyment of the narrative. We expected to find that participants who watched a video with a spoiler would enjoy the video less than those that watched the video with no spoiler and those that read the text with a spoiler or without a spoiler. We found that spoilers did not reduce participants' enjoyment of narratives and that there was no interaction between format and spoilers. However, we found a significant interaction between spoilers and format on participants' reported processing fluency/plot comprehension, which could potentially explain our results. Finally, participants reported that they actively dislike and try to avoid spoilers, furthering the idea that people often over-forecast the spoiler effect. Spoilers are increasingly relevant and hard to avoid in today's highly connected and media-focused society. Although research on the psychological effect of spoilers is limited and inconsistent, our findings suggest that people may overestimate the effect of spoilers on narrative enjoyment.

Keywords: spoilers, stories, amusement, media format, narratives

Imagine waiting for hours in line to watch Star Wars: The Empire Strikes Back (Kershner, 1980) in theaters, only to have two men in a pickup truck drive by, loudly exclaiming, "Darth Vader is Luke's father". My father (CR) experienced this exact event on the opening night of the movie in 1980. He ultimately enjoyed the movie, despite the spoiler. However, he often recounts the exasperation he felt that day. Leavitt and Christenfeld (2011, 2013) define spoilers as revealed plot developments that affect overall enjoyment or feelings of anticipation. As with the case of my father, spoilers often cause frustration. As the name suggests, many people feel that spoilers can ruin their enjoyment of a narrative. However, this may not be the case.

Since 2011, there has been increased interest in research on the effect of narrative spoilers on participant's enjoyment of text and media, known as the spoiler effect. Much of the recent research on the spoiler effect stems from a pair of articles published by Leavitt and Christenfeld (2011, 2013). Originally, Leavitt and Christenfeld (2011) concluded that spoilers do not ruin stories and that they increase enjoyment. They suggested that improved understanding of the narrative and increased anticipation of upcoming plot devices enhanced viewers' experiences. In a follow-up study, they replicated the finding and found that spoilers increase enjoyment by improving processing fluency (Leavitt & Christenfeld, 2013).

Previous spoiler studies frequently point to the idea that improved processing fluency leads to greater narrative enjoyment (e.g., Johnson & Rosenbaum, 2018; Leavitt & Christenfeld, 2013; Levine et al., 2016). Processing fluency is a subjective measure of an observer's ability to parse information (Alter & Oppenheimer, 2009). Typically, higher levels processing fluency correlate with increased positive evaluations of a subject (Rober et al., 2004). In the context of spoiler research, processing fluency is synonymous with plot comprehension (e.g., how easy it is to follow the narrative). Put simply, the belief is that spoilers make it easier for participants to understand the plot. An important consideration to make when examining processing fluency in conjunction with spoilers is that (in theory) the effect of processing fluency on enjoyment scales directly with the complexity of the narrative. The effect on simple stories may be small or nonexistent,

while the effect on more complex stories is more prominent (Leavitt & Christenfeld, 2013). Whereas seeing a spoiler for simple children's stories such as Goodnight Moon and The Very Hungry Caterpillar may not affect one's ability to process the narrative, seeing a spoiler for more complex, intricate, or nonlinear stories like those of the mystery-thriller book/film Fight Club or the drama-suspense film Inception could help to contextualize some of the more convoluted narrative elements.

More recent studies corroborate the idea that spoilers may not affect narrative enjoyment (Johnson & Rosenbaum, 2018; Johnson et al., 2020). For example, Johnson and Rosenbaum (2018) state that spoilers have little to no effect on participants' perceived amusement. Later, Johnson et al. (2020) found that those who enjoy the horror genre may find increased enjoyment when experiencing scary or suspenseful plot information before watching. The authors hypothesized that those viewers enjoyed the increased anticipation the spoiler created (Johnson et al., 2020). Past research shows that making predictions about an unknown outcome significantly decreases participants' perceived enjoyment of an event, and removing the unknown factor eliminates this negative effect (Mandel & Nowlis, 2008).

Not all contemporary studies replicate the neutral or positive effect of spoilers. Multiple recent studies suggest that spoilers decrease participants' enjoyment (Daniel & Katz, 2019; Johnson & Rosenbaum 2015; Levine et al., 2016; Yan & Tsang, 2016). Research points towards several factors that could explain these findings, including the location of the spoiler within the narrative (e.g., beginning, middle, or end), the narrative delivery medium, the method of measuring participants' enjoyment, and participants' preexisting beliefs about the spoiler effect. For instance, researchers testing the spoiler effect on different measures of engagement, such as perceived fun, suspense, and overall enjoyment found that spoilers lowered participants' ratings of each of these categories (Johnson & Rosenbaum, 2015). In another study, researchers found that the method of presentation of a narrative influences the spoiler effect. Whereas spoilers did not affect participants' enjoyment of text, they significantly reduced participants' enjoyment of television shows (Daniel & Katz, 2019). Additionally, research shows that participants will go to great lengths to avoid spoilers for narratives they enjoy, regardless of the strength of the spoiler effect (Maxwell, 2022).

We are interested in the effects of spoilers on college students' enjoyment of narratives, and how these effects will interact with different forms of media. Previous research suggests that spoilers decrease enjoyment for video stimuli but not necessarily text stimuli (Daniel & Katz, 2019; Johnson & Rosenbaum, 2015). We predicted that if a group of participants watched a video with a spoiler, then they would enjoy the video less than a group that watched the video with no spoiler and groups that read the text with a spoiler or without a spoiler. This is relevant in today's social media-focused society, where spoilers are increasingly difficult to avoid.

Method

Participants

Participants in this study were 64 college students (18 years & older) enrolled in undergraduate psychology courses in the fall of 2022 at a small midwestern university. Participants reported their age (M = 20.64, SD = 5.01), gender (52 women, 10 men, 2 non-binary) classification (37 freshman, 10 sophomores, 9 juniors, 8 seniors) and race/ethnicity (54 Caucasian, 7 Hispanic, 3 Latino/a/e/x, 3 Black, 3 Asian/Polynesian/Indian). Recruitment instructions specified that participants must bring their cell phone and a pair of earphones. Participants possibly earned course credit for their participation, but their instructors made other alternatives available. We obtained our participants via Psychology Department Research Pool sign-up in SONA (participant pool management software).

Materials

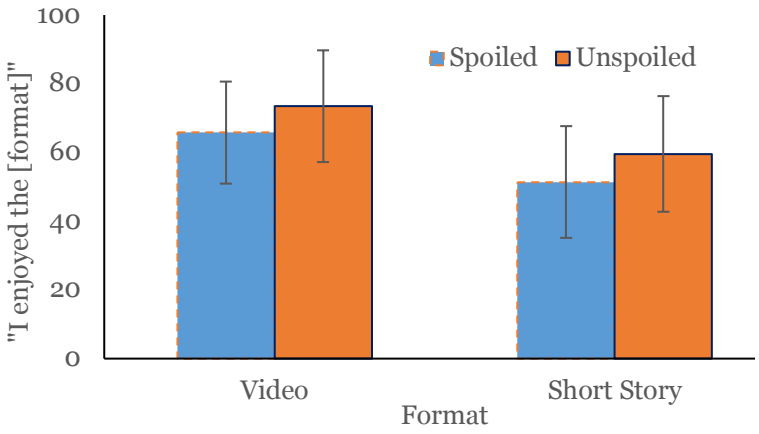
"To Build a Fire" by Jack London is a short story about a man exploring the Yukon in freezing temperatures. After falling into a pool of water, the man fails to build a fire and freezes to death. A shortened version of this story is 5176 words (London, 2008). "To Build a Fire" by Nexus Studios is a faithful animated film adaptation of London's short story. The video is 12 minutes and 59 seconds in length (London, 2008; Nexus Studios, 2016).

The video and text conditions each had two separate prompts to introduce the media. The non-spoiler prompt read, "Below, you will see [description of media]. Please [watch the video/read the story] in its entirety before continuing to the next slide." The spoiler prompt read, "Below, you will see [description of media]. This story follows a man and his dog in a frozen environment. After falling into a frigid pool of water, the man fails to build a fire. The man attempts to kill his canine companion for warmth but fails to do so. Ultimately, because of his failures, the man freezes to death. Please [watch the video/read the story] in its entirety before continuing to the next slide."

Depending on the assigned spoiler condition the Story Survey had either eight (no spoiler) or nine

Figure 1

Mean Enjoyment Ratings Across Conditions



Note. Mean perceived enjoyment scores by participants who either read a short story without a spoiler (n = 14), read a short story with a spoiler (n = 15), watched a short video without a spoiler (n = 16), or watched a video with a spoiler (n = 18). Error bars show 95% confidence intervals for the mean estimates. Higher scores indicate more enjoyment.

(spoiler) items. All groups completed eight slider-scale items designed to evaluate participants' perceptions of the story (e.g., "I felt transported into the story" where 0 = not at all and 100 = absolutely). The spoiler condition surveys contained one additional slider scale item, "I felt irritated after being exposed to the ending of the story," where 0 = not at all and 100 = absolutely. Within the Story Survey, we used the slider scale item, "I enjoyed the [format: short video/short story]" (0 = not at all, 100 = absolutely) to measure participants' narrative enjoyment and, "The plot was easy to follow" (0 = not at all, 100 = absolutely) to measure participants' perceived processing fluency (see Appendix for survey items). The Demographic Survey contained four items (age, gender, student classification, race/ethnicity).

Design and Procedure

We obtained IRB approval prior to conducting any procedures of this 2 (format: text, video) × 2 (spoiler: present, not) factorial study. We read the informed consent form aloud as participants followed along in Qualtrics (a secure web-based survey-building program). Qualtrics randomly assigned participants to one of four groups: short film without spoiler, short film with spoiler, text without spoiler, or text with spoiler. Participants then read their assigned prompt, before either reading the short story or watching the short film. Then, participants completed the Story Survey items, followed by the Demographic Survey. Lastly, participants read the debriefing statement.

Results

To test our hypothesis, we used two independent variables: the format (video/short story) of the story we presented to the participants and the presence of a spoiler (present/not present). The dependent variable was the participants' enjoyment of the video/short story. We calculated mean and standard deviations of the enjoyment slider scale, "I enjoyed the [format]" for each of the groups (Figure 1). We compared the enjoyment scores for each of the four groups by using a 2 (spoiler condition: spoiler, non-spoiler) × 2 (format: video, short story) completely randomized factorial ANOVA. For analyses with significant main effects or interactions ($p < .05$), we calculated simple effects using one-way ANOVAs and determined the effect size with partial η^2. We expected to find that participants who watched a video with a spoiler would enjoy the video less than a group that watched the video with no spoiler and groups that read the text with a spoiler or without a spoiler.

The main effect of format was not significant, $F(1, 59) = 3.33$, $p = .073$, partial $\eta^2 = .05$. The levels of enjoyment for the video group ($M = 69.89$, $SD = 30.86$, $n = 34$) were not significantly different from the levels of enjoyment for the short story group ($M = 55.49$, $SD = 31.11$, $n = 29$; see Figure 1). The main effect of spoilers was not significant, $F(1, 59) = 1.08$, $p = .30$, partial $\eta^2 = .02$. The levels for the spoiler group ($M = 58.59$, $SD = 33.4$, $n = 33$) were not significantly different from the non-spoiled group ($M = 66.79$, $SD = 29.34$, $n = 30$). The interaction of format and spoiler presence was also not significant, $F(1, 59) < 0.01$, $p = .99$, partial $\eta^2 < .01$. In other words, story format and spoiler presence did not affect college student's enjoyment of the video/text.

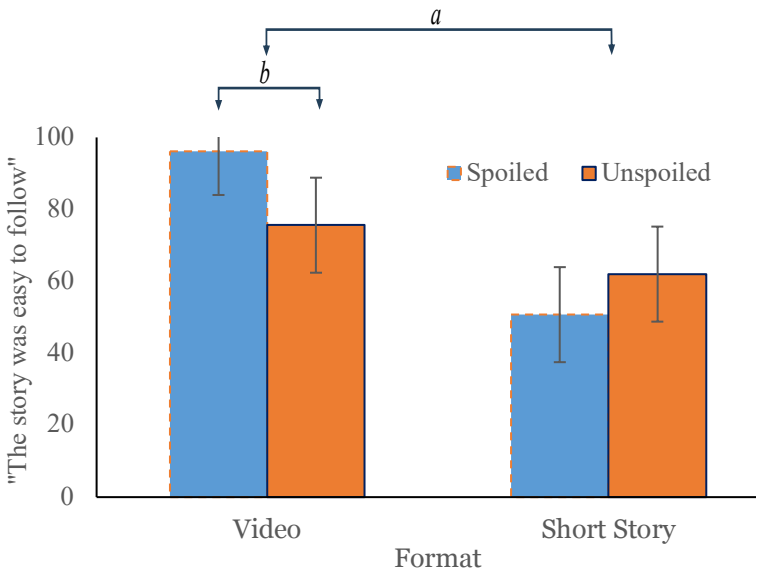
We also conducted exploratory analyses on the effect of spoilers and format on processing fluency. Although we did not make *a priori* hypotheses about processing fluency, previous research suggests that spoilers might increase processing fluency, thereby affecting enjoyment. The main effect of format was significant, $F(1, 60) = 20.78$, $p < .001$, partial $\eta^2 = .26$. The levels of plot comprehension for the video conditions ($M = 86.76$, $SD = 31.21$, $n = 33$) were significantly higher than the levels of plot comprehension for the short story conditions ($M = 56.37$, $SD = 31.21$, $n = 30$). In other words, participants in the video condition found the plot easier to follow than those in the text condition. There

was no main effect of spoilers of spoilers on plot comprehension, $F(1, 60) = 0.51$, $p < .26$, partial $\eta^2 = .01$. Further analyses revealed a significant interaction between format and spoiler presence $F(1, 60) = 5.65$, $p = .02$, partial $\eta^2 = .09$. Probing the interaction revealed a simple effect of spoilers on participants in the video condition, $F(1, 59) = 5.24$, $p = .03$. Specifically, those in the spoiler condition rated their plot comprehension as higher ($M = 75.60$, $SD = 7.94$) than those in the non-spoiled condition ($M = 68.80$, $SD = 30.74$). There was no simple effect of spoilers on plot comprehension of participants in the text condition, $F(1, 59) = 1.46$, $p = .23$. See Figure 2. Put simply, participants who saw a spoiler in the video condition reported higher plot comprehension than those in the video condition who did not see a spoiler. In the text condition, however, spoilers did not affect the reported plot comprehension of participants.

We also measured participants' attitudes towards spoilers—including irritation after experiencing a plot spoiler, belief that spoilers reduce enjoyment, and spoiler avoidance. We collected these data by utilizing a 0-100 slider scale and analyzed them via descriptive statistics. Participants reported minor irritation after exposure to plot spoilers in the study ($M = 38.67$, $SD = 25.70$, $n = 30$). They also largely believed that plot spoilers reduce their enjoyment of narratives ($M = 59.95$, $SD = 31.79$, $n = 60$) and reported that they actively attempt to avoid plot spoilers ($M = 67.92$, $SD = 30.21$, $n = 63$). Essentially, participants carry a negative attitude towards spoilers and actively try to avoid them regardless of their actual effect on enjoyment.

Figure 2

Mean Comprehension Ratings Across Conditions



Note. Mean comprehension ratings by participants who either read a short story without a spoiler ($n = 14$), read a short story with a spoiler ($n = 15$), watched a short video without a spoiler ($n = 15$), or watched a video with a spoiler ($n = 18$). Error bars show 95% confidence intervals for the mean estimates. Higher scores indicate higher levels of self-reported plot comprehension. Italicized lowercase letters indicate a statistically significant relationship ($p < .05$).

Discussion

The purpose of this study was to see if spoilers negatively affected college students' enjoyment of stories and whether the format of the story influenced the spoiler effect. We hypothesized that participants who watched a video with a spoiler would enjoy the video less than a group that watched the video with no spoiler and groups that read the text with a spoiler or without a spoiler. Upon collecting and analyzing data, we found that spoilers did not reduce students' enjoyment of the short story or video. These findings do not support our original hypothesis.

We developed our design with the goal of combining elements from several previous studies. For example, the first study that sparked academic interest in spoilers had college students read a short story (with or without a spoiler), then rate their enjoyment (Leavitt & Christenfeld, 2011). Future replication attempts stuck to a similar format but changed elements such as the genre and format of the stories (Daniel & Katz, 2019; Johnson et al., 2020; Yan & Tsang, 2016). Daniel and Katz (2019) compared the effect of spoilers on enjoyment of short stories versus television shows. However, they did not use the same stories across conditions (Daniel & Katz, 2019). We wanted our study to be a direct comparison of the effect of spoilers on video versus text, which is why we elected to use the same narrative for all conditions. Additionally, whereas many of the existing studies on spoilers utilized Likert-type scales for data collection (e.g., Johnson & Rosenbaum, 2018; Leavitt & Christenfeld, 2011; Leavitt & Christenfeld, 2013), we elected to use slider scales. We did so to allow for greater freedom of choice in participants' answers.

Ultimately, we did not find a statistically significant difference between groups that read a short story with or without

a spoiler and groups that watched a short film with or without a spoiler. These results reflect the studies that found no negative effect of spoilers (Johnson & Rosenbaum, 2018; Johnson et al., 2020; Leavitt & Christenfeld, 2011; Leavitt & Christenfeld, 2013). However, based on our results, we were also unable to recreate the increase in enjoyment shown in some studies (Leavitt & Christenfeld, 2011; Leavitt & Christenfeld, 2013). Our results were unable to replicate any study that found a significant negative effect of spoilers (Daniel & Katz, 2019; Johnson & Rosenbaum, 2015; Levine et al., 2016; Yan & Tsang, 2016). We were also unable to replicate findings which found that format (video, television, text, etc.) moderates the effect of spoilers (Daniel & Katz, 2019).

However, we found that plot comprehension produced significant results. While spoilers alone had no effect, narrative format and the interaction between format and spoilers significantly affected plot comprehension. Overall, participants in our study found the plot of the video easier to comprehend than the plot of the short story. Spoilers themselves did not have a significant effect on overall plot comprehension. However, the presence of spoilers produced different outcomes for each of the format conditions. Participants who received a spoiler in the video condition rated their plot comprehension as significantly higher than those who saw the video without a spoiler—whereas spoilers had no effect on plot comprehension in the text condition. This outcome is different from previous studies that frequently demonstrated an increase in plot comprehension among participants who saw a spoiler (Johnson & Rosenbaum, 2018; Johnson et al., 2018; Leavitt & Christenfeld, 2013). Potential explanations for this finding include the possibility that the spoiler prompt was not powerful enough, or potentially that the short story was too complex for participants in this study.

We believe that it is difficult to generalize the findings of this study. Two of our results reflect consistent findings from other studies. For example, participants over-forecast the spoiler effect (Yan & Tsang, 2016). In the case of our study, participants' beliefs about spoilers did not reflect the findings. Our results are also consistent with studies that found that, regardless of the effect of spoilers on enjoyment, participants report that they try to avoid plot spoilers as much as possible (Maxwell, 2022). However, we believe that it is difficult to generalize the main effect of spoilers and format on enjoyment of narratives. Like some existing studies (e.g., Johnson & Rosenbaum, 2018; Johnson et al., 2020; Leavitt & Christenfeld, 2011; Leavitt & Christenfeld, 2013) we found that spoilers did not significantly negatively affect participants' enjoyment of narratives. However, our data are dissimilar to other recent findings (e.g., Johnson & Rosenbaum 2015; Levine et al., 2016; Yan & Tsang, 2016)—including the study that ours was most closely related to (Daniel & Katz, 2019). Compared to these existing studies—regardless of whether they corroborate our findings—our data lack significantly in one key area. Past studies had significantly larger samples, which helps to improve internal validity and account for variability. Given this significant difference in sample size, it is difficult to draw conclusions or generalizations about how spoilers affect participants' enjoyment of narratives or the interactions between spoilers and narrative format.

There are several limitations that limit the generalizability of our study. Firstly, we did not include a question asking participants if they had read or watched "To Build a Fire" before. This potentially introduces a confound if any participants had previous knowledge of the story. The methodology of having participants in a standard classroom complete the study on their phones may have allowed for distractions, like incoming calls or notifications. The classroom setting also meant that participants were likely able to see the screens of other people's devices, potentially allowing participants to guess the experimental manipulation. We also had to remove several participants from each session because they did not read the requirement to bring their headphones with them. Finally, we believe that using slider scale items for our survey was a mistake. Slider scales generally provide more freedom of choice and help to avoid participants feeling forced into an answer with which they did not agree. However, due to the small sample size, the slider scale items introduced a significant amount of variability, which made both data analysis and generalizing results difficult. According to the central limit theorem, studies require a sample size of roughly 30 participants per group to achieve normal distribution. With an average group size of 15.75, it is difficult to assume normal distribution within our study (Kwak & Kim, 2017). That said, Levene's test did not indicate a violation of normality in our current study ($p = .72$). Regardless, it is difficult to parse whether the high standard deviation for each group is representative of the population or if it is due to individual differences, outliers, selecting answers at random without reading the questions, etc. Collecting data from a larger sample would help to significantly alleviate such issues. Additionally, to mitigate the risks posed by participants randomly selecting answers, we recommend that future studies include attention check items (e.g., "For this item, please select answer 'D'") to identify such cases. However, if we

were to continue collecting data with the same sample, we would elect to use Likert scale items to help reduce measurement error, decrease within-group variability, and simplify data analysis. Nonetheless, although our current data may be underpowered, they provide interesting preliminary insights on spoilers and suggest directions for future research.

Despite the limitations of this study, it still contributes to the relatively small body of literature on the topic of spoilers. We were able to reaffirm some consistent findings of other studies. Participants dislike and attempt to actively avoid spoilers. Other studies have found that this avoidance may be due to fear of missing out (Maxwell, 2022). We also found that the interaction between spoilers and format significantly influenced participants' plot comprehension. Previous authors have mentioned that increased processing fluency from spoilers may allow participants to better understand the plot, which could mitigate the other negative effects of spoilers (Leavitt & Christenfeld, 2013). Another possible explanation is that plot engagement and transportation (feeling immersed within a narrative) play a significant role in moderating the spoiler effect, as perhaps participants' engagement with and excitement about a narrative predict the effect of spoilers on their enjoyment. Previous research found that spoilers did not decrease enjoyment of short stories but did decrease enjoyment of television shows (Daniel & Katz, 2019). If participants were not engaged with our selected narratives, it could potentially explain the outcome. Ultimately, this study serves as a blueprint for further replications and improvements to increase internal validity and generalizability, while also providing some additional insight and reaffirmations to previous findings.

We believe the next logical step for this study is to conduct similar replications while reducing confounds and other issues. Increasing sample size while maintaining internal validity would likely be the biggest factor in improving upon this study. It is important to find a video and short story that are closer to each other in terms of time to complete. We would also recommend choosing a shorter story overall to hopefully improve engagement. Future studies should also utilize existing multi-question measurements for enjoyment, rather than measuring it through a single question. For broader studies on spoilers, we recommend looking closely at the interactions between fear of missing out and the spoiler effect. We also believe that it is important to examine the interaction between engagement and the spoiler effect—which could potentially help to explain some of the variances in findings between studies on this topic. Ultimately, the best way forward is to continue researching the spoiler effect to add to the small research pool and improve our psychological understanding of this increasingly prevalent topic.

References

Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the tribes of fluency to form a metacognitive nation. *Personality and Social Psychology Review*, 13(3), 219–235. https://doi.org/10.1177/1088868309341564

Daniel, T. A., & Katz, J. S. (2019). Spoilers affect the enjoyment of television episodes but not short stories. *Psychological Reports*, *122*(5), 1794–1807. https://doi.org/10.1177/0033294118793971

Johnson, B. K., & Rosenbaum, J. E. (2015). Spoiler alert: Consequences of narrative spoilers for dimensions of enjoyment, appreciation, and transportation. *Communication Research*, *42*(8), 1068–1088. https://doi.org/10.1177/0093650214564051

Johnson, B. K., & Rosenbaum, J. E. (2018). (Don't) tell me how it ends: Spoilers, enjoyment, and involvement in television and film. *Media Psychology*, *21*(4), 582–612. https://doi.org/10.1080/15213269.2017.1338964

Johnson, B. K., Udvardi, A., Eden, A., & Rosenbaum, J. E. (2020). Spoilers go bump in the night: Impacts of minor and major reveals on horror film enjoyment. *Journal of Media Psychology: Theories, Methods, and Applications*, *32*(1), 14–25. https://doi.org/10.1027/1864-1105/a000252

Kershner, I. (Director). (1980). *The Empire Strikes Back* [Film]. Lucasfilm Ltd.

Kwak, S. G., & Kim, J. H. (2017). Central limit theorem: The cornerstone of modern statistics. *Korean journal of anesthesiology*, 70(2), 144–156. https://doi.org/10.4097/kjae.2017.70.2.144

Leavitt, J. D., & Christenfeld, N. J. S. (2011). Story spoilers don't spoil stories. *Psychological Science*, *22*(9), 1152–1154. https://doi.org/10.1177/0956797611417007

Leavitt, J. D., & Christenfeld, N. J. S. (2013). The fluency of spoilers: Why giving away endings improves stories. *Scientific Study of Literature*, *3*(1), 93–104. https://doi.org/10.1075/ssol.3.1.09lea

Levine, W. H., Betzner, M., & Autry, K. S. (2016). The effect of spoilers on the enjoyment of short stories. *Discourse Processes*, *53*(7), 513–531. https://doi.org/10.1080/0163853X.2016.1141350

London, J. (2008). *To build a fire and other favorite stories*. Dover Publications.

Mandel, N., & Nowlis, S. M. (2008). The effect of making a prediction about the outcome of a consumption experience on the enjoyment of that experience. *Journal of Consumer Research*, *35*(1), 9–20. https://doi.org/10.1086/527339

Maxwell, L. C. (2022). Spoilers ahead, proceed with caution: How engagement, enjoyment, and FoMO predict avoidance of spoilers. *Psychology of Popular Media*, *11*(2), 163–172. https://doi.org/10.1037/ppm0000362

Nexus Studios. (2018). *To Build A fire. To Build a Fire*. Retrieved September 7, 2022, from https://www.youtube.com/watch?v=LzRrPQJlUwM

Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review*, 8(4), 364–382. https://doi.org/10.1207/s15327957pspr0804_3

Yan, D., & Tsang, A. S. L. (2016). The misforecasted spoiler effect: Underlying mechanism and boundary conditions. *Journal of Consumer Psychology*, *26*(1), 81–90. https://doi.org/10.1016/j.jcps.2015.05.003

Author Note

Conner K. Ryan is now at the Department of Psychology, University of Wyoming.

Keith Wylie is now at the Department of Behavioral Sciences, Flagler College.

Correspondence concerning this article should be addressed to Conner K. Ryan, 1000 E. University Ave, Dept. 3415, Laramie WY 82071. Email: cryan19@uwyo.edu

Appendix

Story Survey Items

Q6.1 On a scale of 0 to 100, where 0 means "not at all" and 100 means "absolutely":

	0 10 20 30 40 50 60 70 80 90 100
I felt transported into the story. ()	
The story was suspenseful. ()	
I found the story exciting. ()	
The plot was easy to follow. ()	
I liked the story. ()	

Page Break

Q6.2 On a scale of 0 to 100, where 0 means "not at all" and 100 means "absolutely":

	0 10 20 30 40 50 60 70 80 90 100
I enjoyed the ${e://Field/format}. ()	

Page Break

Display This Question:

If spoiler = yes

Q6.3 On a scale of 0 to 100, where 0 means "not at all" and 100 means "absolutely":

	0 10 20 30 40 50 60 70 80 90 100
I felt irritated after being exposed to the ending of the story. ()	

Page Break

Q6.4 On a scale of 0 to 100, where 0 means "not at all" and 100 means "absolutely":

	0 10 20 30 40 50 60 70 80 90 100
I think that plot spoilers reduce my enjoyment of narratives. ()	
I do my best to avoid plot spoilers. ()	

Journal of Psychological Inquiry
2025, Vol. 29, No. 2, pp. 38-48

PERCEIVER CHARACTERISTICS AND TARGET RACE INFLUENCE PERCEPTIONS OF THE IMPOSTER PHENOMENON

EBONIE WHITE AND MERRY J. SLEIGH

WINTHROP UNIVERSITY

Abstract – We sought to better understand adults' perceptions of successful women who varied in race and their experience with the imposter phenomenon. Participants responded to scales to assess their own imposter phenomenon and empathy. Next, we randomly assigned participants to one of four conditions where they read about a highly educated woman working in a male-dominated context. The conditions described the woman as either Black or White. The two "Imposter" conditions described the woman as feeling like an imposter. The other two "Control" conditions described the woman as feeling confident although finding it difficult to work in a male dominated industry. After a manipulation check, we asked participants their perceptions of the woman. Compared to Black participants, White participants reported higher levels of imposter phenomenon. Similarly contradicting previous research, our participants did not assume that Black women in our scenarios were more likely to be experiencing imposter feelings. Our participants were more impressed with the Black women who was succeeding in a male-dominated workplace than with the White women in the same circumstance. Adults who experienced imposter phenomenon themselves were more understanding of Black women, while adults higher in cognitive or affective empathy reported more positive perceptions of the White women.

The imposter phenomenon is the feeling that one's success is fraudulent and one's abilities are being overestimated by others (Bhama et al., 2021; Clark et al., 2021). An intense sense of incompetency and doubt often accompanies these feelings (Narayanamoorthy et al., 2023). People who experience imposter feelings also report fear of failure, fear of being deceptive, concerns about their competency, internal conflict, and external attributions of success (Chakraverty, 2022; Chakraverty & Rishi, 2022; Huecker et al., 2023). Imposter feelings are common, with over 90% of college students reporting moderate or higher levels (Fassi et al., 2020) and prevalence rates at over 50% and higher in graduate students, nurses, and medical students (Bravata et al., 2020a).

Ironically, people with imposter feelings have often reached an elevated level of success in highly competitive environments, such as academia or the medical field (Bhama et al. 2021, Narayanamoorthy et al., 2023). Adults who have achieved formal leadership roles, characterized by high visibility and high levels of responsibility, are more likely to report experiencing imposter feelings (Kark & Peters, 2022). Tefwik (2022) explored potential advantages of imposter feelings and discovered that individuals, who experience these feelings in the workplace, are more likely to be other-focused, resulting in higher evaluative ratings for interpersonal effectiveness than their co-workers.

Although people who experience imposter feelings are often successful, researchers have consistently linked the experience to a range of negative outcomes. Imposter feelings are associated with higher levels of depression, anxiety, compromised educational attainment, and lower self-esteem (Bravata et al., 2020a; Canning et al., 2019; Wyat et al., 2019). The experience of imposter feelings often leads individuals to engage in a cycle of over-preparation or procrastination, driven by a desire for perfectionism (Huecker et al., 2023; Pannhausen et al., 2022). Imposter feelings co-occur with maladaptive perfectionistic behaviors such as dwelling on mistakes, experiencing self-doubt, unrealistic expectations, and fear of failure (Cokley et al., 2018; Sheveleva et al., 2023). Perhaps resulting from these perfectionism-driven behaviors, higher rates of imposter feelings also predict higher rates of burn-out symptomology (Jaremka et al., 2020; Vaa et al., 2023). For example, adults experiencing these feelings report feeling depersonalization and emotional exhaustion, along with experiencing lower levels of personal

accomplishment (Alrayyes et al., 2020; Liu et al., 2022; Wang et al., 2019).

Although both men and women can experience imposter feelings, it occurs at higher rates in women. This gender difference was documented over two decades ago (Clance & Imes, 1978) and is still seen, particularly in industries dominated by men (Collins et al., 2020; Clark & Payne, 2024). Compared to their male counterparts, higher levels of imposter feelings are consistently documented in female adults (Fleischhauer et al., 2021), college students (Pákozdy et al., 2023), graduate students (Cowle et al., 2018), and surgical residents (Narayanamoorthy et al., 2023). When quota systems are used, women may be especially prone to doubt their hiring qualifications and thus, even more susceptible to feeling like an imposter (Wall & Troisi, 2020). College students who associated themselves with negative aspects of femininity were more likely to engage in social comparison and experience greater levels of imposter feelings (Fassi et al., 2020).

Race and ethnicity also impact prevalence rates of imposter feelings, with underrepresented groups more vulnerable to experiencing them (Bravata et al., 2020a; Heslop et al., 2023). One reason is that Black, Indigenous, and other people of color are often underrepresented in highly competitive work and academic settings, leading to a perception of not belonging and, thus, feeling like an imposter (Afran et al., 2020). Being one of few people of color often leads to feeling like a token group member rather than a qualified worker (Wall & Troisi, 2020; Wyatt et al., 2019). For example, ethnically underrepresented medical students were more likely to report imposter feelings when enrolled at a primarily White institution compared to a historically Black college or university (Rice et al., 2023). Cokley et al. (2022) found that stress levels are high among Black attorneys, who represent a small minority (about 5%) of all attorneys in the United States; these stress levels served as a mediator between Black attorneys' imposter feelings and negative health outcomes of stress and depression. Another contributing factor is that minority individuals are more likely to experience microaggressions which creates a sense of feeling like an outsider and predict higher rates of imposter feelings (Chakraverty, 2022). Encounters of discrimination and implicit bias elicit feelings of exclusion and thus, amplify the experience of imposter feelings (Cawcutt, 2021; Jenkins et al., 2022; Meadhbh et al., 2023).

Minority women seem to be especially vulnerable to imposter feelings because of the intersection of two risk factors, gender and race. An investigation of first-generation, minoritized status women attending college revealed that imposter feelings were a common experience (Jackson & Colson-Fearon, 2022). Similarly, female medical students from underrepresented groups reported more imposter feelings than their male counterparts regardless of whether they were at a primarily White or primarily Black educational institution (Rice et al., 2023). In the workplace, women of color's feeling like an imposter are likely exacerbated because of work cultures that implicitly or explicitly promote patriarchal views (Collins et al., 2020; Jenkins et al., 2022).

Although imposter feelings are common in women from marginalized groups, researchers have argued that highly educated Black women are especially prone to this self-perception (Chrousos & Mentis, 2020; Fields & Cunningham-Williams, 2021). In a comparison of Asian American, Latinx, and Black college students, Stone-Sabali et al. (2024) found that the Black students were more vulnerable to imposter feelings because of their greater proneness for shame and anxiety. In addition, Black women are significantly underrepresented in leadership roles, which can facilitate an individual Black woman's perception that she does not belong and is an imposter (Manongsong & Ghosh, 2021).

Much of the previous research has focused on characteristics and outcomes of people experiencing imposter feelings. More recently, researchers have started to investigate how people's perceptions impact the process of labeling. People with implicit bias may unfairly label minority women as being imposters whether it is true or not (Feenstra et al., 2020; McGee et al., 2021). These labeled women are typically high achieving; thus, the mislabeling is not the result of their poor performance (Bahama et al., 2021). Instead, the mislabeling is likely the result of biased perceivers, emphasizing demographic attributes of the target woman, such as race and gender (Cawcutt et al., 2021; Simons, 2021). Such labeling may create a self-fulfilling prophecy, because women who perceive themselves to be the subject of negative stigmas and stereotypes in the work force are more likely to report having imposter feelings (Collins et al., 2020). For example, among female software engineers, an increased awareness of gender stigma was associated with increased experiences of imposter feelings (Maji & Dixit, 2023). Ironically, Chakraverty and Rishi (2022) found that while women in certain STEM disciplines were more likely to experience imposter feelings than men, the perpetuators who elicited those feelings included both women and men.

Because imposter feelings are so prevalent, researchers have worked to identify factors that may

decrease the experience, particularly for Black women. Manongsong and Ghosh (2021) argued that a network of supportive others can enable minoritized women to develop a positive leader identity. Another intervention strategy focused on helping women identify the presence and consequences of imposter feelings, with awareness of those feelings minimizing its impact (Parker et al., 2023). These interventions have an underpinning element of empathy and understanding, leading us to wonder whether empathy might reduce attributions of imposter feelings.

Our study focused on attributions of imposter feelings in the workplace, varying in whether the symptoms were explicitly framed as imposter feelings or not and whether the person experiencing it was a Black or White woman. We also investigated how the perceiver's race, empathy level, and personal imposter feelings impacted their judgments of the women in the imposter feelings situation. We hypothesized that:

1. Because marginalized groups report higher rates of imposter feelings, Black participants would report higher rates of imposter feelings than White participants.

2. Because of implicit biases, participants would be more likely to be impressed that Black women have achieved success compared to White women.

3. Participants with higher rates of imposter feelings would report more empathy and understanding for the women for whom imposter feelings were explicitly described (imposter condition) than for the women who was not described as feeling like an imposter (control condition).

4. Higher levels of affective or cognitive empathy would predict more positive reactions to the women in the imposter conditions than women in the control conditions.

We had an additional exploratory question. Because of documented tendencies for adults to over-attribute the imposter feeling to marginalized groups, participants might have been more likely to perceive the two Black women as experiencing imposter feelings in comparison to the two White women. However, a competing possibility is that in the condition specifically stating that the woman feels like an imposter, participants might have been more likely to perceive the two imposter condition women as experiencing imposter feelings compared to the two control condition women. In other words, we were interested in whether race and condition would interact, or if one of the two variables (race or imposter status) would prove more salient to young adults.

Method

The study was not preregistered. Three researcher-created items that assessed perceptions of the experimental condition were accidentally miscoded by the researchers and thus were not included in the analyses. These items were similar to other items that were included. All other measures in the study are reported.

Participants

We originally had 121 participants respond to the survey; however, participants who did not correctly respond to our three validity check items were eliminated (see Materials section below). All remaining participants responded to the majority of survey items. Items that were skipped were analyzed as missing data, and no outliers were removed. Thus, our final sample included 113 adults with a mean age of 23.88 (SD = 9.22). When we asked participants their identity, 93 were cisgender women, 14 were cisgender men, 1 was transgender, one identified as other, and 4 were non-binary. When we asked participants their sexual orientation, 76 participants identified as heterosexual, 5 as gay/lesbian, 22 as bisexual, 2 as asexual, 4 as pansexual, and 4 as other. When we asked participants to categorize their race, 58 were White, 43 were Black, 6 were Latinx/Hispanic, 3 were Asian, 2 were multi-racial, and 2 was "Other."

This study was approved by our institutional IRB. We recruited participants through personal contact, undergraduate classrooms, Facebook, e-mail, and social networking sites. Potential participants were issued an invitation that included the web address to link to the online survey. Once participants went to that address, they encountered the informed consent form. Choosing to go to the web address was completely voluntary on the part of the participants and required participants to click a link indicating that they agreed to participate.

Materials and Procedure

First, participants responded to the Clance Imposter Phenomenon Score (CIPS; Clance, 1985). Twenty items evaluated the extent to which the participant experienced feelings of fraudulence or unworthiness associated with their accomplishments. A sample item was "I have often succeeded on a test or task even though I was afraid that I would not do well before I undertook the task." Participants made responses on a 5-point Likert scale where 1 represented "not at all true" and 5 represented "very true." The published reliability for this scale ranged from 0.85 to 0.96 (Mak et al., 2019), and we had a Cronbach's alpha of .91.

Next, participants responded to the Basic Empathy Scale (Jolliffe & Farrington, 2006). These 20

items assessed cognitive and affective empathy. Affective empathy refers to the ability to feel what others feel. A sample item was "I tend to feel scared when I am with friends who are afraid." Cognitive empathy refers to the ability to mentally understand the feelings of others. A sample item was "When someone is feeling 'down' I can usually understand how they feel." The average Cronbach's alphas for these scales are .81 (Cabedo-Peris et al., 2021). We had a Cronbach's alpha of .76 for cognitive empathy, and .84 for affective empathy.

We then randomly assigned participants to one of four conditions. The conditions can be seen in the Appendix. In each condition, participants read a brief narrative of a woman who was working in a male-dominated environment where imposter feelings would be likely. For example, all four narratives stated that the character "is currently in a high-ranking position in her career at Company…She is in a male dominated career field, and she is the only woman working at her office." The conditions varied in two ways. First, the woman was described as either Black or White. Second, the conditions differed in how specifically imposter feelings were mentioned. Two conditions stated that the women "feels like an imposter but hides those feelings from everyone around her" and included symptoms of imposter phenomenon (imposter condition). The other two conditions did not contain the word imposter (control condition). Instead, the woman was described as feeling confident and enjoying the challenges of her job. However, the last sentence of these two conditions stated that "It's difficult to work in a male dominated industry but she is making it happen." The created conditions were labeled Black Control, Black Imposter, White Control and White Imposter.

Immediately after reading the story, participants encountered three items that served as a validity check to ensure that participants read the story. The first question asked the participant to recall the name of the company stated in the story. The second asked the participant to choose the answer that best describes the woman they read about. The third asked them to choose the answer that best described the job the woman mentioned.

Next, participants encountered five researcher-created items that measured participants' perceptions of the woman. One item directly assessed our manipulation by stating, "This woman is probably experiencing imposter syndrome." (Note that we did not provide participants with a definition of imposter syndrome.) Two items assessed whether participants perceived the woman positively or negatively: "I am impressed by this woman" and "This woman is probably annoying to work with." We also assessed participants' pity and empathy for the woman. The two items were "I empathize with this woman and her situation" and "I feel sorry for this woman." Participants then encountered an additional researcher-created item that stated, "If I were in this situation, I would experience self-doubt." Responses were made on a 5-point Likert scale where 1 represented "strongly disagree" and 5 represented "strongly agree."

Last, participants encountered five commonly used demographic questions that assessed age, race, gender, and sexual orientation.

Results

Hypothesis One and Participant Race Comparisons

We conducted independent t-tests to compare the responses of participants who self-identified as either Black or White. Compared to White participants, Black participants scored lower on the imposter phenomenon scale (CIPS), $t(99) = 3.71$, $p < .001$, $d = .75$, 95% CI [.34, 1.15]. The mean for White participants was 68.98 (SD = 13.51) and for Black participants was 59.0 (SD = 13.17). Means can be seen in Figure 1. This finding contradicts our hypothesis that Black participants would report higher rates of imposter feelings and reflect a moderate effect. Compared to White participants, Black participants were less likely to agree that they "would doubt themselves in the same situation" as the woman in the scenario, $t(69) = 1.86$, $p = .034$, $d = .46$, CI [-.03, .95]; this reflects a small effect. Compared to White participants, Black participants also scored lower on the affective empathy scale, $t(99) = 2.50$, $p = .007$, $d = .50$, CI [.10, .90], reflecting a moderate effect. We found no

Figure 1
Imposter Syndrome Score of Participants

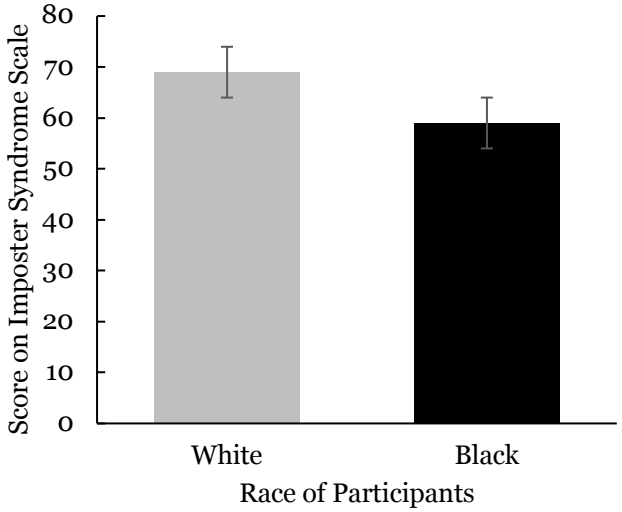


Table 1
Means and Standard Deviations for Dependent Variables Across Race and Imposter Conditions

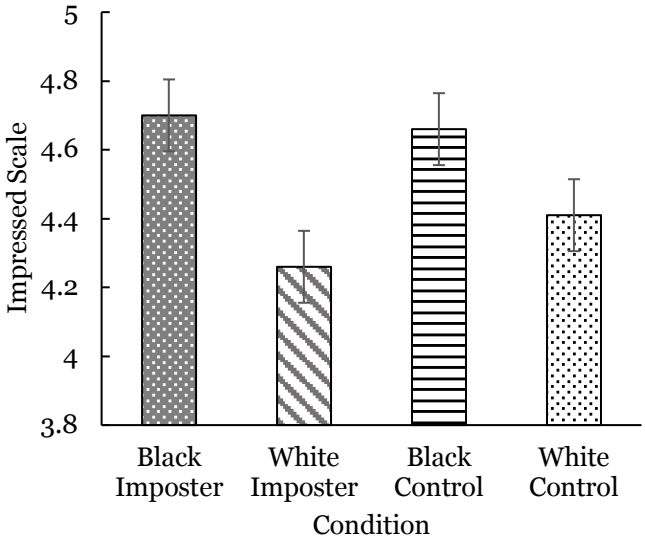
	Imposter Condition		Control Condition	
	Black Woman	White Woman	Black Woman	White Woman
Impressed By	4.70 (.54)	4.26 (.92)	4.66 (.55)	4.41 (.95)
Feel Sorry For	3.37 (1.18)	2.65 (1.34)	2.25 (1.24)	2.07 (1.22)
Empathize With	4.33 (.68)	4.04 (1.26)	3.78 (1.04)	3.90 (1.21)
Annoying to Work With	1.33 (.68)	1.65 (1.03)	1.41 (.76)	1.66 (1.01)
Experiencing Imposter	3.44 (1.09)	4.04 (1.04)	3.06 (1.41)	2.97 (1.40)

differences between Black and White participants' scores on the cognitive empathy scale [$t(99) = .81$, $p = .21$, *ns*].

Hypothesis Two and Condition Comparisons

We compared the responses to the four researcher-created items that assessed participants' perceptions of the described woman using a 2(race) X 2(imposter status) MANOVA. The means and standard deviations can be seen in Table 1. We found no significant interactions. In support of our hypothesis, there was a main effect for race, where participants agreed more that they were "impressed by this woman" for the Black, versus White conditions, $F(1, 111) = 5.70$, $p = .019$, $\eta 2 = .05$. This represents a small effect. See Figure 2 for the means. There was a main effect for the imposter condition, with participants agreeing more that they felt "sorry for this woman" for the imposter, versus the control, conditions, $F(1, 111) = 12.84$, $p < .001$ $\eta 2 = .03$. This represents a small effect. We also found a main effect for the imposter condition, where participants who read the imposter condition scenarios were more likely to agree that the woman was "probably feeling imposter syndrome," $F(1, 111) = 9.26$, $p = .003$ $\eta 2 = .09$. This represents a medium effect. We found no main effects for how much participants "empathize with this woman and her situation" [$F(1, 111) = .18$, $p = .669$, *ns*] or for how much participants agreed that the "woman is probably annoying to work with" [$F(1, 111) = 2.89$, $p = .092$, *ns*].

Figure 2
Means for Participants Impressed by Different Conditions



Hypothesis Three and Participants' Imposter Phenomenon Scores as Predictors

Because we found no interactions, we split the data file between the four conditions to examine relations among variables. Participants with higher CIPS agreed more that they would "doubt themselves in the same situation" when they read about the Black Control woman [$r(21) = .66$, $p = .001$] and Black Imposter woman [$r(23) = .55$, $p = .006$]. These reflect moderate to strong correlations. Participants with higher CIPS also agreed more that they "empathize with" [$r(27) = .41$, $p = .035$] and felt "sorry" [$r(27) = .41$, $p = .035$] for the Black Imposter woman. These reflect moderate correlations. Participants' CIPS did not predict their perceptions of the White women. This analysis did not support our hypothesis that participants with higher rates of imposter feelings would report more empathy and understanding for the Imposter, versus the Control, woman.

Hypothesis Four and Participants' Empathy Scores as Predictors

Our examination of empathy did not support our hypothesis that higher levels of empathy would predict more positive reactions to the Imposter, versus the Control, woman. Instead, the race of the target woman was more influential. Higher scores on the affective empathy scale correlated with participants agreeing more that were impressed with the White Control woman [$r(30) = .47$, $p = .010$] and White Imposter woman [$r(24) = .52$, $p = .010$]. These reflect moderate correlations. The higher participants' affective empathy the less

they agreed that the White Control woman [$r(29) = -.45$, $p = .014$] and White Imposter woman [$r(24) = -.52$, $p = .009$] were "probably annoying to work with." These reflect moderate correlations. Participants' affective empathy levels did not predict their perceptions of the Black women.

The higher participants' cognitive empathy, the more they felt empathy for the Black Control woman [$r(32) = .35$, $p = .047$], Black Imposter woman [$r(27) = .51$, $p = .006$], and White Control woman. [$r(30) = .44$, $p = .015$]. These reflect weak to moderate correlations. Cognitive empathy did not predict empathy for the White Imposter woman. Instead, participants with higher cognitive empathy responded to the White Imposter woman by agreeing more that they "were impressed" [$r(24) = .72$, $p < .001$] and agreeing less that she was "probably annoying to work with" [$r(24) = -.68$, $p < .001$]. These are strong correlations.

Discussion

Hypothesis One: Imposter Feelings by Participant Race

We hypothesized that, in line with previous research, our Black participants would report higher rates of imposter feelings than our White participants. This hypothesis was not supported. In fact, Black participants in our study had lower imposter scores and disagreed that they would doubt themselves if they were in the imposter situation described in our narratives. Our sample was drawn largely from young adults with some college education, and our scenario described someone who was advanced in her career. College students may not be able to predict how much doubt they would feel in a situation that they had not yet encountered. However, this argument does not explain why Black, versus White, participants reported less imposter feelings (a clear contradiction with previous research). One possibility is that, because of the age of our participants, they experienced some portion of their schooling during the isolation of COVID-19; the forced separation from peers may have decreased opportunities for experiences of social discrimination and comparison. Social comparison, particularly in a face-to-face context, is known to increase imposter feelings, particularly among women (Fassi et al., 2020; Fraenza, 2016).

Most of our Black participants were recruited from a racially diverse campus, which may also have reduced their experiences with social discrimination, social comparison, and thus, imposter feelings. In support of this argument, Bernard et al. (2024) recently found that Black college students report higher rates of imposter feelings when their campus racial climate was negative, a link mediated by social anxiety. In other words, perceived campus inequities related to higher anxiety in Black students, and those Black students reported more imposter feelings (Bernard et al., 2024). The racially diverse nature of the campus from which our Black participants were recruited also may have afforded them a certain amount of social support. In systematic literature review, Bravata et al. (2020b) found that increased social support is related to lower levels of imposter feelings. In conjunction, these findings suggest the need for continued investigation into contextual and historical factors that may work alongside other intersectional factors that impact an individual's vulnerability to imposter feelings.

An exploratory question in our research was whether race or imposter status would determine which women, or woman, would be perceived as experiencing imposter feelings. In our study, imposter status emerged as more influential. Participants who read either of the imposter conditions agreed more strongly that the woman was experiencing imposter feelings. They also reported feeling sorry for the two women who were explicitly described as experiencing imposter feelings compared to the two women who were not. This finding provides evidence that our manipulation worked because we described these women as feeling like imposters, and our participants agreed that the women felt like imposters. In other words, our use of the word "imposter" in the narrative description was salient enough to get participants to attribute "imposter syndrome" to the women in that specific condition. (The wording of the scenarios can be seen in the Appendix.) Previous research suggested that adults over-attribute imposter feelings to marginalized groups (Feenstra et al., 2020; McGee et al., 2021), which would have predicted that the Black women in our study would have been seen as "feeling like an imposter;" however, we did not see this in our study. Recall that our Black participants reported lower levels of imposter feelings than our White participants, and thus, perceptions on these items may reflect the personal experiences of our participants.

Hypothesis Two: Impressed Perceptions by Target Race

We hypothesized that participants would be more impressed by the successful Black women compared to the successful White women. This prediction was supported. The women in our scenario were described as succeeding in a male-dominated world and being highly educated. Being impressed by the Black woman may reveal an implicit bias that this level of achievement is less common for minority women and thus, more noteworthy; this implicit bias has been clearly established by previous researchers (e.g., Iheduru-Anderson, 2022;

Pogrebna et al., 2024). Black women may internalize the global perception that Black women's success is uncommon, leading to the higher rates of imposter feelings typically reported for Black women (Noskeau et al., 2021; Wall & Troisi, 2020). An alternative and more optimistic explanation might be that our participants assumed that the workplace was dominated specifically by White men, versus Black men, a perception that would make sense given typical workplace demographics (Afran et al., 2020). If this assumption was made, participants may have been particularly impressed by the Black women because their minority status included both race and gender, whereas White women only represented gender minority status. In other words, the Black women might have been perceived as impressive because they achieved workplace success despite membership in two historically disadvantaged groups in the workplace (Dill & Duffy, 2022; Pogrebna et al., 2024).

Hypothesis Three: Imposter Feelings Predict Empathy and Understanding

We hypothesized that increased personal feelings related to believing one was an imposter would predict greater empathy and understanding for the two women in the conditions where imposter feelings were clearly described. This hypothesis was not supported. Participants with more personal feelings related to believing they were an imposter were more likely to relate to the Black, versus the White women, regardless of her imposter condition. These participants, high in imposter feelings, also were more likely to agree that they empathized with and felt sorry specifically for the Black Imposter woman. In other words, participants' own imposter feelings predicted their responses to the women's race more than to the women's imposter experience. Perhaps adults who are aware of imposter feelings assumed that all Black women experience imposter feelings, and because they personally know that imposter feelings are unpleasant, they responded with empathy and pity. Such an assumption might be fair because Black women are more likely to report imposter feelings (Chrousos & Mentis, 2020; Fields & Cunningham-Williams, 2021). Another possibility is that the result reveals a microaggression. White participants in our study were more likely to experience imposter syndrome – perhaps they revealed an implicit bias by assuming if they, personally, were feeling like an imposter, a Black woman would be even more likely to feel like an imposter. This argument would be in line with previous researchers who documented that people over-attribute the imposter label to minority women (Cawcutt et al., 2021; Feenstra et al., 2020; McGee et al., 2021; Simons, 2021).

Hypothesis Four: Empathy Score and Perceptions of Conditions

We also hypothesized that higher levels of affective or cognitive empathy would predict more positive reactions to the women in the imposter conditions than women in the control conditions. This hypothesis was not supported, and like our earlier presented findings, the race of the woman was more predictive of perceptions than her imposter status. Participants who were higher in affective empathy (e.g., more capable of feeling what others feel) were more impressed with the White woman and believed her to be more less annoying to have as a co-worker. One possible explanation is that this outcome was driven by our White participants. White participants scored higher on affective empathy, and it might be these same White participants who preferred the White woman in the workplace. In other words, White adults may be showing a bias for their own race, a bias that exists across race and culture (Meissner & Brigham, 2001; Wong et al., 2020). A similar argument might be that White adults, who were adept at feeling what others feel, had an especially easy time identifying and connecting with the emotions of White people.

When we examined cognitive empathy (e.g., the ability to understand another's situation) we found a somewhat similar pattern. Participants higher in cognitive empathy reported feeling more empathy for all the women except the White Imposter. Instead of feeling empathy for her, they were impressed by her and believed her to be less annoying. This pattern is close to what we saw with affective empathy; however, affective empathy predicted positive feelings for both White women, while cognitive empathy predicted positive feelings specifically for the White Imposter woman. People high in cognitive empathy analyze situations with logic in addition to emotion. The data suggest that there was something about the combination of "White" and the inclusion of the word "imposter" in the scenario that led these participants to be particularly impressed with the woman. Maybe these two terms (White, imposter) seemed less compatible than Black and imposter, leading to deeper processing and reflecting some implicit bias.

Limitations and Conclusion

The main limitation of this study is the lack of diversity within the sample. There was diversity in race. However, our sample had a disproportionately high percentage of women (87%) compared to men. This gender disparity might be especially relevant given that our narratives focused on women characters. In addition to broadening participant demographics, future researchers may want to collect more data about adults'

own experiences at work or in school to start identifying specific experiential factors that may drive perceptions. Future researchers could also focus on how perceptions translate toward behaviors. For example, does having empathy for someone experiencing imposter feelings impact interactions with that individual?

In sum, our data did not clearly support our hypotheses. Compared to Black participants, White participants reported higher levels of imposter feelings. Similarly contradicting previous research, our participants did not assume that our Black women were more likely to be experiencing imposter feelings. Although we expected our imposter condition to influence perceptions, we found stronger patterns related to race. Our participants were more impressed with the Black women who was succeeding in a male-dominated workplace than with the White women in the same circumstance. Adults who experienced imposter feelings themselves were more understanding of Black women, while adults higher in empathy, who were disproportionately White women, reported more positive perceptions of the White women. These findings suggest that imposter feelings may not always be higher, or perceived to be higher, in Black, compared to White, women. In addition, our findings add to the limited research examining how perceiver characteristics impact attributions of imposter feelings by demonstrating that perceivers' empathy and personal experiences with imposter feelings influence their perceptions.

References

Alrayyes S., Dar U. F., Alrayes M., Alghutayghit A., & Alrayyes N. (2020). Burnout and imposter phenomenon among Saudi young adults. The strings in the puppet show of psychological morbidity. *Saudi Medical Journal, 41*(2), 189-194. https://doi.org/10.15537/smj.2020.2.24841

Bernard, D. L., Stone-Sabali, S., Volpe, V. V., Matthews, I. J., & Smith, L. V. (2024). Campus racial climate and imposter phenomenon among Black college students: The mediating roles of social anxiety and self-esteem. *Journal of Diversity in Higher Education*. Advance online publication. https://doi.org/10.1037/dhe0000590

Bhama, A. R., Ritz, E. M., Anand, R. J., Auyang, E. D., Lipman, J., Greenberg, J. A., & Kapadia, M. R. (2021). Imposter phenomenon in surgical trainees: Clance Imposter Phenomenon Scale assessment in general surgery residents. *Journal of the American College of Surgeons, 233*, 633-638. https://doi.org/10.1016/j.jamcollsurg.2021.07.681

Bravata, D. M., Madhusudhan, D. K., Boroff, M., & Cokley, K. O. (2020a). Commentary: Prevalence, predictors, and treatment of imposter phenomenon: A systematic review. *Journal of Mental Health and Clinical Psychology, 4*(3), 12-16. Retrieved from https://www.mentalhealthjournal.org/articles/commentary-prevalence-predictors-and-treatment-of-imposter-syndrome-a-systematic-review.pdf

Bravata, D. M., Watts, S. A. Keefer, A. L., Madhusudhan, D. K., Taylor, K. T., Clark, D. M., Nelson, R. S., Cokley, K. O., Hagg, H. K. (2020b). Prevalence, predictors, and treatment of imposter syndrome: A systematic review. *Journal of General Internal Medicine*, *35*, 1252-1275. https://doi.org/10.1007/s11606-019-05364-1

Cabedo-Peris, J., Martí-Vilar, M., Merino-Soto, C., & Ortiz-Morán, M. (2021). Basic Empathy Scale: A systematic review and reliability generalization meta-analysis. *Healthcare Basel, Switzerland*, *10*(1), 29. https://doi.org/10.3390/healthcare10010029

Canning, E. A., LaCosse, J., Kroeper, K. M., & Murphy, M. C. (2019). Feeling like an imposter: The effect of perceived classroom competition on the daily psychological experiences of first-generation college students. *Social Psychological and Personality Science, 11*(5), 647–657. https://doi.org/10.1177/1948550619882032

Chakraverty, D. (2022). Imposter phenomenon and identity-based microaggression among Hispanic/Latinx individuals in science, technology, engineering, and mathematics: A qualitative exploration. *Violence and Gender*, *9*(3), 135-141. https://doi.org/10.1089/vio.2021.0061

Chakraverty, D. & Rishi, M. (2021). Imposter phenomenon and discipline-specific experiences of violence in science, technology, engineering, and mathematics. *Violence and Gender*. 9. https://doi.org/10.1089/vio.2021.0025

Chrousos, G. P., & Mentis, A. F. (2020). Imposter phenomenon threatens diversity. *Science, 367*(6479), 749–750. https://doi.org/10.1126/science.aba8039

Clance, P. R. (1985). *The imposter phenomenon: When success makes you feel like a fake*. Bantam Books.

Clance, P. R., & Imes, S. A. (1978). The imposter phenomenon in high achieving women: Dynamics and therapeutic intervention. *Psychotherapy: Theory, Research & Practice, 15*(3), 241–247. https://doi.org/10.1037/h0086006

Clark, P., Holden, C., Russell, M., & Downs, H. (2021). The imposter phenomenon in mental health professionals: Relationships among compassion fatigue, burnout, and compassion satisfaction.

Contemporary Family Therapy, 185–197. https://doi.org/10.1007/s10591-021-09580-y

Clark, C. T. & Payne, J. L. (2024). Gender diversity in the psychiatric workforce: It's still a (White) man's word in psychiatry. *Child and Adolescent Psychiatric Clinics of North America*, *33*, 45-52. https://doi.org/10.1016/j.chc.2023.06.005

Cokley, K., Krueger, N., Garba, R., Bailey, M., Harris, K., Hall, S., Coleman, C. & Archer, J. (2022). Lawyering while Black: Perceived stress as a mediator of imposter feelings, race-related stress and mental health among Black attorneys. *Journal of Black Psychology*, *48*(2), 2060232. https://doi.org/10.1177/00957984211070216

Cokley, K., Stone, S., Krueger, N., Bailey, M., Garba, R., & Hurst, A. (2018). Self-esteem as a mediator of the link between perfectionism and the imposter phenomenon. *Personality and Individual Differences, 135*, 292–297. https://doi.org/10.1016/j.paid.2018.07.032

Collins, K. H., Price E. F., Hanson L., & Neaves D. (2020). Consequences of stereotype threat and imposter phenomenon: The personal journey from STEM-practitioner to STEM-educator for four women of color. *Taboo. The Journal of Culture and Education*, *19*(4), 161–180. Retrieved from https://digitalscholarship.unlv.edu/taboo/vol19/iss4/10

Cowle, M. E., Nealis, L. J., Sherry, S. B., Hewitt, P. L. & Flett, G. L. (2018). Perfectionism and academic difficulties in graduate students: Testing incremental prediction and gender moderation. *Personality and Individual Differences*, *123*, 223-228. https://doi.org/10.1016/j.paid.2017.11.027

Dill, J. & Duffy, M. (2022). Structural racism and Black women's employment in the US health care sector. *Health Aff (Millwood)*, *41*, 265-272. https://doi.org/10.1377/hlthaff.2021.01400

Fassi, F., Yanagida, T. & Kollmayer, M. (2020). Imposters dare to compare: Associations between the imposter phenomenon, gender typing, and social comparison orientation in university students. *Frontiers in Psychology*, *11*, ArtID: 1225. https://doi.org/10.3389/fpsyg.2020.01225

Feenstra, S., Begeny, C. T., Ryan, M. K., Rink, F. A., Stoker, J. I., & Jordan, J. (2020). Contextualizing the imposter "phenomenon." *Frontiers in Psychology, 11*. https://doi.org/10.3389/fpsyg.2020.575024

Fields, L. N., & Cunningham-Williams, R. M. (2021). Experiences with imposter phenomenon and authenticity at research-intensive schools of social work: A case study on Black female faculty. *Advances in Social Work, 21*(2/3), 354–373. https://doi.org/10.18060/24124

Fleischhaeuer, M., Wossidlo, J., Michael, L. & Enge, S. (2021). The imposter phenomenon: Toward a better understanding of the nomological networks and gender differences. *Frontiers in Psychology*, *12*, ArtID: 764030. https://doi.org/10.3389/fpsyg.2021.764030

Fraenza, C. B. (2016). The role of social influence in anxiety and the imposter phenomenon. *Online Learning*, *20*, 230-243. https://files.eric.ed.gov/fulltext/EJ1105971.pdf

Iheduru-Anderson, K., Okoro, F. O., & Moore, S. S. (2022). Diversity and inclusion or tokens? A qualitative study of Black women academic nurse leaders in the United States. *Global Qualitative Nursing Research*, *18*, epub. https://doi.org/10.1177/23333936211073116

Jackson, A. & Colson-Fearon, B. (2022). Managing intersectional invisibility and hypervisibility during the transition to college among first-generation women of color. *Psychology of Women Quarterly*, *46*(3), 354-371. https://doi.org/10.1177/03616843221106087

Jaremka, L. M., Ackerman, J. M., Gawronski, B., Rule, N. O., Sweeny, K., Tropp, L. R., Metz, M. A., Molina, L., Ryan, W. S., & Vick, S. B. (2020). Common academic experiences no one talks about: Repeated rejection, imposter phenomenon, and burnout. *Perspectives on Psychological Science, 15*(3), 519-543. https://doi.org/10.1177/1745691619898848

Jenkins, P. R., Roberts-Lewis, K., Hernandez, P. M., Segura, D., & Greensdale, V. (2022). A pilot study: Exploring the psychological demise among women of color in the academy. *Researcher: An Interdisciplinary Journal, 30*(2), 51-74.

Jolliffe, D., & Farrington, D. P., (2006). Development and validation of the Basic Empathy Scale. *Journal of Adolescence, 29*(4), 589-611. https://doi.org/10.1016/j.adolescence.2005.08.010

Kark, R., Meister, A. & Peters, K. (2022). Now you see me, now you don't: A conceptual model of the antecedents and consequences of leader imposterism. *Journal of Management*, *48*, 1948-1979. https://doi.org/10.1177/01492063211020358

Liu, R.Q., Davidson, J., Van Hooren, T. A., Van Koughnett, Julie A. M., Jones, S., & Ott, M. C. (2022). Imposterism and anxiety contribute to burnout among resident physicians. *Medical Teacher, 44*(7), 758-764. https://doi.org/10.1080/0142159X.2022.2028751

Maji, S., & Dixit, S. (2023). Gender stigma consciousness, imposter phenomenon, and self-silencing: A mediational relationship. *Psychological Studies, 54*. https://doi.org/10.1007/s12646-023-00724-8

Mak, K. K. L., Kleitman, S., & Abbott, M. J. (2019). Imposter phenomenon measurement scales: A

systematic review. *Frontiers in Psychology*, *10*(671). https://doi.org/10.3389/fpsyg.2019.00671

Manongsong, A. M. & Ghosh, R. (2021). Developing the positive identity of minoritized women leaders in higher education: How can multiple and diverse developers help with overcoming the imposter phenomenon? *Human Resource Development Review*, *20*(4), 436-485. https://doi.org/10.1177/15344843211040732

Meadhbh, M. Ó., Tiffany Chiu, Y. L., Wong, B., & Horsburgh, J. (2023). Deindividualising Imposter phenomenon: Imposter work among marginalised STEM undergraduates in the UK. *Sociology, 57*(4), 749-766. https://doi.org/10.1177/00380385221117380

McGee, E. O., Botchway, P. K., Naphan-Kingery, D. E., Brockman, A. J., Houston, S., & White, D. T. (2021). Racism camouflaged as imposterism and the impact on black STEM doctoral students. *Race Ethnicity and Education*, *25*(4), 487-507. https://doi.org/10.1080/13613324.2021.1924137

Meissner, C. A. & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology and Public Policy Law*, *7*, 3 - 35. https://doi.org/10.1037//1076-8971.7.1.3

Narayanamoorthy S., McLaren R., Pendam R., & Minkoff H. (2023). Are women residents of surgical specialties at a higher risk of developing imposter phenomenon? *American Journal of Surgery*, *227*, 48-51. https://doi.org/10.1016/j.amjsurg.2023.09.025

Noskeau, R., Santos, A., Wang, W. (2021). Connecting the dots between mindset and imposter phenomenon, via fear of failure and goal orientation, in working adults. *Frontiers in Psychology*, *12*, 588438. https://doi.org/10.3389/fpsyg.2021.588438

Pákozdy, C., Askew, J., Dyer, J., Gately, P., Martin, L., Mavor, K. I. & Brown, G. R. (2023). The imposter phenomenon and its relationship with self-efficacy, perfectionism and happiness in university students. *Current Psychology*, *43*, 5153-5162. https://doi.org/10.1007/s12144-023-04672-4

Pannhausen, S., Klug, K. & Rohrmann, S. (2020). Never good enough: The relation between the imposter phenomenon and multidimensional perfectionism. *Current Psychology*, *41*, 888-901. https://doi.org/10.1007/s12144-020-00613-7

Parker, J. S., Haskins, N., Coleman-Chatman, T., Duru, C., Ramirez, D., Chae, N., Fulmore, K. & Nortey, A., (2023). Dismantling the imposter syndrome: A case study of early career women of color in school mental health fields. *Journal for Specialists in Group Work*, *49*, 3-24. https://doi.org/10.1080/01933922.2023.2281670

Pogrebna, G., Angelopoulos, S., Motsi-Omoijiade, I., Kharlamov, A., & Tkachenko, N. (2024). The impact of intersectional racial and gender biases on minority female leadership over two centuries. *Scientific Reports, 14,* 111. https://doi.org/10.1038/s41598-023-50392-x

Rice, J., Rosario-Williams, B., West-Livingston, L., Savage, D., Wilensky, J. A. & Landry, A. (2023). Imposter syndrome among minority medical students who are underrepresented in medicine. *Journal of the National Medical Association*, *115*(2), 191-198. https://doi.org/10.1016/j.jnma.2023.01.012

Sheveleva, M. S., Permyakova, T. M., Kornienko, D. S. (2023). Perfectionism, the imposter phenomenon, self-esteem, and personality traits among Russian college students. *Psychology in Russia*, *16*(3), 132-148. https://doi.org/10.11621/pir.2023.0310

Stone-Sabali, S., Uanhoro, J. O., McClain, S., Bernard, D., Makari, S., & Chapman-Hilliard, C. (2024). Imposterism and psychological distress among college students of color: A moderation analysis of shame-proneness, race, gender, and race-gender interactions. *Current Psychology*, *43*(4), 1-17. https://doi.org/10.1007/s12144-023-04579-0

Vaa Stelling B. E., Andersen C. A., Suarez D. A., Nordhues H. C., Hafferty F. W., Beckman T. J., & Sawatsky A. P. (2023). Fitting in while standing out: Professional identity formation, imposter phenomenon, and burnout in early-career faculty physicians. *Academic Medicine: Journal of the Association of American Medical Colleges, 98*(4), 514-520. https://doi.org/10.1097/ACM.0000000000005049

Wall, S. L. & Troisi, K. (2020). Impact of gender bias on women surgeons: A South African perspective. *Journal of Medical Ethics*, *46*(11), 785-786. https://doi.org/10.1136/medethics-2020-106442

Wang, K. T., Sheveleva, M. S., & Permyakova, T. M. (2019). Imposter syndrome among Russian students: The link between perfectionism and psychological distress. *Personality and Individual Differences, 143*, 1–6. https://doi.org/10.1016/j.paid.2019.02.005

Wong, H. K., Stephen, I. D., Keeble, D. R. T. (2020). The own-race bias for face recognition in a multiracial society. *Frontiers in Psychology*, *11*, 208. https://doi.org/10.3389/fpsyg.2020.00208

Wyatt, G. E., Chin, D., Milburn, N., Hamilton, A., Lopez, S., Kim, A., Stone, J. D. & Belcher, H. M. E. (2019). Mentoring the mentors of students from diverse backgrounds for research. *American Journal of Orthopsychiatry*, *89*(3), 321-328. https://doi.org/10.1037/ort0000414

Appendix

Experimental Conditions

Imposter Condition	A Black/White woman is currently in a high-ranking position in her career at Company AI. She has spent years working hard to earn 2 degrees that have qualified her for the position. She is the first person in her family to graduate from college, and they are very proud of her. She is in a male dominated career field, and she is the only woman working at her office. She is having trouble accepting her accomplishments, wondering if she deserves the praise that she is getting. She worries whether she can keep doing as good of a job in the future as she has done in the past and whether she will ultimately disappoint her family. Often, she feels like an imposter but hides those feelings from everyone around her.
Control Condition	A Black/White woman is currently in a high-ranking position in her career at Company CD. She has spent years working hard to earn 2 degrees that have qualified her for the position. She is the first person in her family to graduate from college, and they are very proud of her. She is in a male dominated career field, and she is the only woman working at her office. She feels confident in her ability and enjoys the praise she gets from others. She has to work hard but she enjoys the challenge. She feels especially proud of herself when she faces a tough situation and work and figures out a solution to the problem. It's difficult to work in a male dominated industry but she is making it happen.

Journal of Psychological Inquiry
2025, Vol. 29, No. 2, pp. 49-61

INFANT LEARNING AND ARTIFICIAL INTELLIGENCE: EXPLORING PROBABILISTIC CONNECTIONS FOR COGNITIVE ADVANCEMENT

RONG FAN AND CHELSEA ANDREWS

UNIVERSITY OF WISCONSIN-MADISON

Abstract – There is well-established theory in both infant learning and artificial intelligence (AI) suggesting significant overlap in probabilistic learning mechanisms. Research on infant learning, such as Saffran's (1996) work on transitional probabilities, demonstrates that infants use statistical cues to segment speech and develop word understanding. Similarly, AI models, particularly generative models like neural networks, are trained through statistical methods to simulate human-like cognitive processes (Goodfellow et al., 2016). This paper investigates the intersection of these probabilistic mechanisms, focusing on key studies in infant probabilistic reasoning and the development of language AI models, such as GPT-4.0. The review highlights the shared cognitive functions between infants and AI algorithms through a comparative analysis, emphasizing how both systems predict and process complex inputs from their environments. The discussion moves beyond algorithmic processing to consider the broader implications for advancements in cognitive sciences and AI. By bridging these disciplines, this review positions AI not just as a technology but as a potentially significant model for cognitive development research.

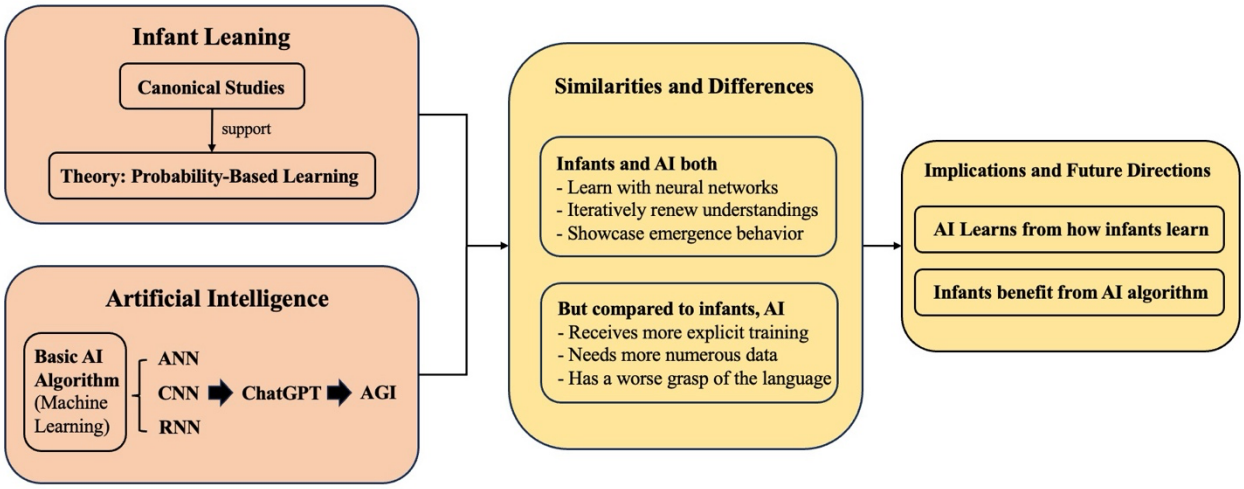
Keywords: infant learning, artificial intelligence, probabilistic learning, cognitive development, generative AI, language acquisition

The method by which infants learn to understand the world differs from that of adults in various ways. Within the domain of language, for example, infants do not immediately grasp the sophisticated vocabulary and intricate grammatical structures mastered by their caregivers. Instead, they develop an understanding of their immediate surroundings by extracting the meanings of individual words from the structures and semantics of the sentences (Wojcik & Saffran, 2015). Here, "structure" denotes the word order in sentences, while "semantics" pertains to the connections between surrounding words. As Saffran (2020) notes, infants discern patterns and regularities; they unconsciously employ a probabilistic method to comprehend language.

Acquiring a similar probabilistic learning technique is Artificial Intelligence (AI), which refers to the capacity of digital computers to perform tasks commonly associated with intelligent beings (Copeland, 2023). Recently, public interest in AI has surged substantially due to Generative AI, which can generate various types of content, including text, imagery, audio, and synthetic data (Goodfellow et al., 2014). The interest in AI continues to grow with the advent of generative AI through Large Language Models (LLM), such as Generative Pre-trained Transformer (GPT) 3.5, which can engage in human-like conversations with remarkable intelligence. GPT 4.0 has garnered increased attention in AI development, and the most recent Sora by OpenAI is a foundational step toward achieving Artificial General Intelligence (AGI) (OpenAI, 2024). Despite its extreme sophistication, generative AI also operates through statistical learning, closely resembling human language acquisition (Dupoux, 2018).

Given that both infants and artificial intelligence systems rely on probabilistic mechanisms to process information, an intriguing question arises: Are AI models imitating the learning processes of human beings? If so, to what extent can these models replicate the sophisticated cognitive functions of early human development? By exploring their similarities and differences, this review aims to assess whether AI models—such as ChatGPT—can replicate or enhance our understanding of human cognitive development (Figure 1). Addressing this question has broader implications, both for advancing AI and for shedding light on how human cognition, especially in early developmental stages, might be better understood and potentially simulated through AI frameworks.

Figure 1.
Conceptual Map: Parallels Between Infant Learning and Artificial Intelligence.



Note. This figure illustrates the logical relationships between various sections of the paper, highlighting the key themes and connections explored in the exploration of infant learning and artificial intelligence.

Infant Learning

Existing Studies

Several significant experiments have demonstrated infants' ability to process probability. Xu and Garcia (2008) meticulously explored how infants could make inferences from samples to populations and vice versa, revealing an innate capacity to process probability. Their experiments involved a controlled setup where infants were shown a population of ping-pong balls of varying colors. They were then shown a series of samples drawn from this population, with some samples accurately representing the diversity of the population and others deliberately skewed. Notably, when the total population was later revealed, infants showed significantly longer gaze durations for the intentionally skewed samples. These results suggest that even at a very young age, infants can predict information about a population (color distribution, in this case) based on their understanding of a sample. Further reinforcing these findings, when infants were briefly exposed to the population before seeing the samples, one representative and the other skewed, infants stared longer at the skewed samples. This indicates that infants used the base rate information obtained from the population to form predictions about which sample was more likely. These reactions were robust indicators of infants' probabilistic processing abilities.

The work of Shultz and Nobandegani (2022) delved deeper into infants' ability to discern probability and predict outcomes. Their experiments echoed the setup of Xu and Garcia (2008), where infants were initially shown the population of objects, followed by samples. The infants' surprise was measurable when presented with skewed samples, suggesting a discrepancy between their predictions and the observed sample. When the sequence was reversed—revealing a disproportionate sample first and the population afterward—the infants' surprise reactions were still prominent, further illustrating their capacity to process both sample-to-population and population-to-sample inferences.

Building on the understanding of infants' skill in handling probability, Saffran et al. (1996) made a significant step toward deciphering infant language acquisition by investigating how infants use probabilities to discern word boundaries. Saffran et al. presented infants with a stream of sounds from an invented language and measured their listening times to both familiar and novel words in that language. The data showed that infants listened longer to novel words, indicating recognition of their unfamiliarity. This suggested that, after merely two minutes of listening, infants could identify the boundaries of unfamiliar words and distinguish them from familiar vocabulary. In the same study, Saffran et al. conducted a complementary experiment where infants were exposed to words and part-words from the same artificial language. Part-words were created by combining the initial syllable of one word

with the final syllable of another. Since part-words included actual word boundaries, they were technically novel to the infants. The infants' longer listening times to these part-words further confirmed that they had learned to identify words as discrete units. This study demonstrates that infants could discern word boundaries solely based on sequential statistics.

Wojcik and Saffran's (2015) study broadened the understanding of infants' statistical learning capabilities by examining their ability to discern regularities within and between sentences. Their findings revealed that infants could recognize not only the "horizontal" relationships of novel nouns that co-occurred within the same sentence but also the "vertical" relationships of nouns based on their consistent positioning within different sentences. The infants could then make more sophisticated predictions based on the regularities they identified across sentences. This ability is foundational for grasping the complex nature of sentence construction and semantics.

Collectively, these experiments illuminate how adeptly infants process probability. From discerning relationships between samples and populations to encoding linguistic regularities, infants demonstrate a remarkable capacity for probabilistic learning. The convergence of these findings underscores the depth of infants' engagement with probabilistic mechanisms, providing compelling evidence of their sophisticated cognitive abilities. As we examine these experiments, it becomes increasingly evident that probability serves as a cornerstone in infants' multifaceted language acquisition journey.

Probability-Based Learning

The experiments conducted empirically consolidate theories regarding infant learning, particularly within the intricate landscape of language acquisition. These studies establish the paradigm of probability-based learning in infants, a concept continuously explored within psychology. This section delves into how infants leverage probability and statistical regularities to acquire language skills.

Acquiring an unfamiliar language poses a formidable challenge for infants, necessitating the initial task of segmenting individual words from the continuous stream of speech (Estes, 2009; Saffran et al., 1996). In essence, infants break down complex sentences into digestible fragments. Once this partitioning is accomplished, the subsequent phases focus on vocabulary (understanding individual words) and grammar (connecting words to make sense of the sentence).

Mastering vocabulary is a process shown to be both statistical and probabilistic, playing a pivotal role in infants' linguistic journey. Toddlers showcase a remarkable ability to infer the meaning of novel words encountered within sentences, demonstrating their adeptness in discerning repeating sounds and predicting subsequent words based on the regularities of fluent language spoken to them (Saffran, 2020; Wojcik & Saffran, 2015). For example, infants can use word order and argument structure to deduce the meaning of a novel verb (Naigles, 1990) and employ familiar verbs to anticipate the semantic properties of upcoming nouns (Wojcik & Saffran, 2015). Through discerning statistical regularity, infants extract unknown words from continuous speech and link them to referents (Estes, 2009; Saffran, 2020). Consequently, infants begin to understand words by their surrounding words, relative position in the sentence, phonotactic patterns, etc. With these probability cues, infants form a primitive understanding of each word. Then, learning becomes iterative: when the word is heard again, infants compare their initial understanding with the new contextual usage (probability cues within the new context). This process enables infants to subtly adjust and refine their comprehension of vocabulary over time (Arciuli & Monaghan, 2009). With each exposure and interaction, infants iteratively fine-tune their understanding until it aligns sufficiently with the context, gradually expanding and sharpening their lexicon (Arciuli & Monaghan, 2009). In essence, infants learn vocabulary through continuous refinement, gradually shaping their rudimentary lexicon to better suit the nuances of their linguistic environment.

The study of infant grammar learning is notably focused on Artificial Grammar Learning (AGL), a paradigm in cognitive psychology and linguistics that evaluates subjects' ability to learn constructed grammar in a laboratory setting (Miller, 1963; Pothos, 2007). In a recent study on AGL, Gervain et al. (2020) generalized infants' ability to learn abstract grammatical and structural properties of language. They argue that young infants can learn simple repetition-based structures, discriminate them from other structures, and discern inconsistencies in unfamiliar structures. Meanwhile, infants identify functors—words whose purpose is to signal grammatical relationships rather than lexical meaning—as anchor points for structural signals that mark parts of speech. As infants detect these grammatical regularities, their expectations of grammatical structure and what they hear allow them to continuously modify and refine their understanding of grammar (Gervain et al., 2020; Saffran, 2020).

In conclusion, infants employ statistical methods to acquire the structural patterns of their native language. They establish and refine associations between sounds and meanings through iterative exposure to linguistic input (Arciuli & Monaghan, 2009). This early stage of infant language acquisition bears striking similarities to the learning processes observed in artificial intelligence when acquiring human language.

Artificial Intelligence

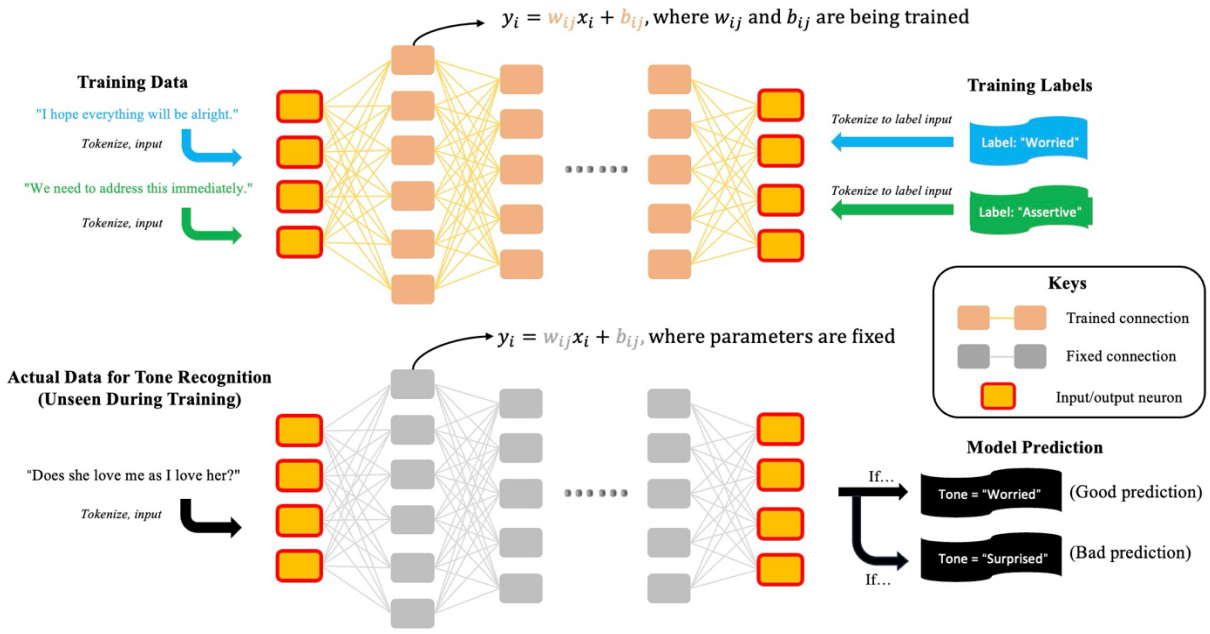
AI Algorithm and Existing AI Models

Intelligence has been defined as the ability to adapt to one's surrounding environment and effectively handle new situations (Abbate, 2023; Anastasi, 1986). Recent advancements in generative AI have progressively impressed the public with the intelligence showcased by AI algorithms. Achieving AI capabilities comparable to GPT has been a gradual process characterized by a learning trajectory. This journey, known as machine learning, involves extracting meaningful representations from data across various levels of abstraction and subsequently optimizing them (Wang et al., 2023). This learning process is fundamentally statistical, akin to how humans acquire language, albeit with a significantly larger volume of data input (Dupoux, 2018).

Machine learning models come in various types, applications, and levels of complexity, but most employ iterative epochs to refine estimations. In language learning, machines often leverage Artificial Neural Networks (ANNs) to enhance performance. Figure 2

Figure 2.
Artificial Neural Network Architecture for Tone Recognition

(a) Top Figure, the model during the training phase.
(b) Bottom figure, the model in practical use.



Note. The model receives tokenized sentences as input, with each box representing a neuron conducting linear regression. Neurons within hidden layers receive input from the preceding layer, perform linear regression, and transmit predictions to the subsequent layer. This process continues until the results reach the output layer, where a final prediction vector is formed. Tokenized labels are provided to the output layer, which compares them with the final prediction. Through training epochs, the model adjusts weights and biases for each neuron to optimize the final prediction's alignment with the given labels.

Once the model is trained, the neurons retain their configurations. When presented with new data, the model endeavors to predict the tone of the sentences. High-quality predictions and elevated accuracy signify a well-suited model, whereas poorer predictions indicate areas for improvement.

depicts the general workings of an ANN, using an example of tone recognition. Sentences in human language are tokenized, establishing a one-to-one correspondence between words and arrays of binary numbers, making abstract sentences understandable for machines (Mikolov et al., 2013). During the learning phase, tokenized labels are assigned to each input sentence. These labels can encompass aspects such as one (joyful, sarcastic, formal, etc.), topic (psychology, cooking, gossip, etc.), or the context (conversation, speech, scientific paper) of the sentence. During training, the ANN seeks to establish (typically linear) relationships in a multidimensional space to connect the input sentence with its corresponding label (Goodfellow et al., 2016; Figure 2a). A network of neurons is deployed to construct a model capable of capturing the intricate relationship between words and their assigned labels. With each training epoch, the weights and biases of neurons undergo adjustments to minimize errors, striving for a better fit (Goodfellow et al., 2016). Once training is complete, the model can be evaluated with unseen sentences to assess its performance in predicting the assigned label (Figure 2b).

This ANN acts as a black box: the model seeks only to establish relationships and does not truly understand the language. Even with high prediction accuracy, it remains difficult for users—and even experts—to explain why a specific set of weights and biases models the data correctly (Zhang et al., 2016). To reduce this uncertainty, more advanced machine learning models for language processing include algorithms such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs). CNNs are particularly effective at identifying and processing spatial features, making them ideal for handling visual data, such as images. Each image can be thought of as a grid of pixels, each containing color values for red, green, and blue. CNNs work by applying filters that highlight different aspects of the image. For instance, when distinguishing between images of humans and horses, a CNN might use filters that emphasize the number of legs. A filter that identifies four vertical structures might indicate a horse, while one that detects two might suggest a human. This ability to recognize distinct spatial features is key to how CNNs classify images.

RNNs, in contrast, excel at analyzing data where sequence or order is important, such as time series data or text. They work by processing sequences one element at a time, maintaining a memory of what has been processed so far to use in making decisions about the current element. This is particularly useful in language processing. For example, in understanding a sentence, an RNN processes each word in the context of the preceding words, allowing it to interpret the sentence with an awareness of the evolving meaning from start to finish.

While advanced AI models, exemplified by the transformer model used in GPT, incorporate many complex techniques, the fundamental building blocks often include CNN and RNN architectures and logic (Kim, 2014; Lipton et al., 2015). At the heart of the transformer model is a sophisticated mechanism known as "multi-head attention" (Vaswani et al., 2017). This system enables a dynamic and context-aware analysis of sequences, functioning like a sophisticated interrogation process where each word in a sentence simultaneously queries every other word to understand its function and relationship within the text. Each query gathers distinct "answers" or features, similar to the way filters in CNNs identify relevant spatial features in visual data. These accumulated features are layered, much like filters in a CNN, refining and enhancing the contextual understanding of each word. By processing all words concurrently and retaining their interrelations over time, transformers integrate a temporal dimension reminiscent of RNNs' handling of sequential data. Consequently, despite their simplicity, CNNs and RNNs are foundational components of many sophisticated AI models today.

While advanced model architecture is crucial, the breadth and depth of the training data are also essential in shaping AI's intelligence. The Large Language Model (LLM) plays a significant role in developing adaptable, general language systems (Brown et al., 2020). LLMs analyze extensive language pattern data to identify statistical regularities (features) and word order (time dependence) to enhance models' understanding of language vocabulary and grammar. Although most models do not fully understand the language, they recognize correlations. The LLM-trained model then attempts to comprehend input and generate output in the learned language (Brown et al., 2020). GPT is a transformer model that is pre-trained on vast amounts of text data. It has demonstrated exceptional performance across various natural language processing tasks, including question answering, text summarization, language translation, and text generation (Abbate, 2023).

Despite its complex nature, GPT uses a probabilistic concept mechanism that aligns with basic machine learning paradigms, although it is considerably larger and more complex, with numerous add-ons that form the transformer structure (Abbate, 2023). With the expansive data provided by LLMs, the complex model learns to respond in a human-like manner when users provide input to ChatGPT, showcasing not only

intelligence but also adaptability. ChatGPT can even answer open-ended questions for which there is no predetermined answer, such as, "Write me a paragraph about AI as a real friend to a user." In this example, GPT spontaneously generates a creative response, demonstrating its ability to adapt to novel scenarios (OpenAI, 2023).

It is essential to note that, despite its remarkable capabilities, GPT operates based on vast collections of regressions. It does not possess a fundamental understanding of users' input as language itself. Instead, GPT relies on statistical regularities derived from LLMs and algorithms to generate the most probable responses (Brown et al., 2020). This probabilistic nature grants GPT versatility in responding to a wide array of problems, although it may occasionally produce incorrect answers while confidently asserting their validity (Abbate, 2023).

Artificial General Intelligence (AGI)

In the realm of technology, few recent developments have sparked as much debate as Artificial General Intelligence (AGI) (Faraboschi et al., 2023). Although a universally accepted definition of AGI is lacking, its central attribute is intelligence capable of solving almost any task, often with superior performance to humans (Shevlin et al., 2019). AGI aims not only to mimic human thought processes but also to autonomously complete tasks that humans typically face.

The increasing parallels between AI and human learning processes have become a source of both fascination and apprehension in the public sphere (Faraboschi et al., 2023). Concerns arise as AI systems become capable of performing tasks on par with, or even better than, humans. For instance, AGI can exhibit self-reliance by making decisions based on its learned patterns and data with minimal human intervention. In the realm of language acquisition, AGI could conceivably attain a genuine understanding of human language, equaling or even surpassing the comprehension abilities of infants. Current research in AI language acquisition predominantly revolves around "AI learning from humans," an approach in which human logic is encoded into AI algorithms to enhance language processing performance (Johnson et al., 2022). The trajectory of AI development hints at a potential future where AGI might contribute to enhancing language acquisition processes for infants.

Despite the natural proficiency that infants display in learning language, the rapid advancement of AI capabilities suggests that AGI may eventually offer novel methods for language learning. Reflecting on the evolution of AI, the transition from classical programming to machine learning significantly boosted the language learning capabilities of machines. Given the pace of development in current AI models and the breakthroughs they have achieved, another paradigm shift in future models' language learning capabilities seems plausible. This shift could introduce new patterns and methodologies that not only mirror but also potentially surpass human language acquisition capabilities. While no existing model currently approaches this level of intelligence, the prospect is an exciting possibility that may unfold in the not-too-distant future.

Efforts toward AGI are promising, with the release of Sora by OpenAI in February 2024 marking a significant milestone. Representing a fundamental step toward AGI, Sora demonstrates the capability to generate photorealistic videos and comprehend human instructions with unprecedented accuracy (OpenAI, 2024). This level of intelligence and autonomy is both a significant advancement and a source of alarm, primarily because of Sora's potential to disrupt fields such as movie production and advertising. Public apprehension regarding Sora's intelligence surpassing human capabilities is palpable, with concerns about job displacement and the restructuring of modern industries looming large in discussions about the implications of advanced AI technologies.

A canonical threshold of machine intelligence is the Turing test, which determines whether a machine can demonstrate intelligence equivalent to or indistinguishable from that of a human (“Turing Test,” 2024). While no AI has yet passed the Turing test (Powell, 2019), the trend suggests that AI is inevitably becoming more human-like. These concerns are valid, and ethical considerations are paramount in the development of AI toward AGI, drawing parallels with Asimov's (2004) laws of robotics.

Overlapping Principles

Similarities Between Infant Learning and AI

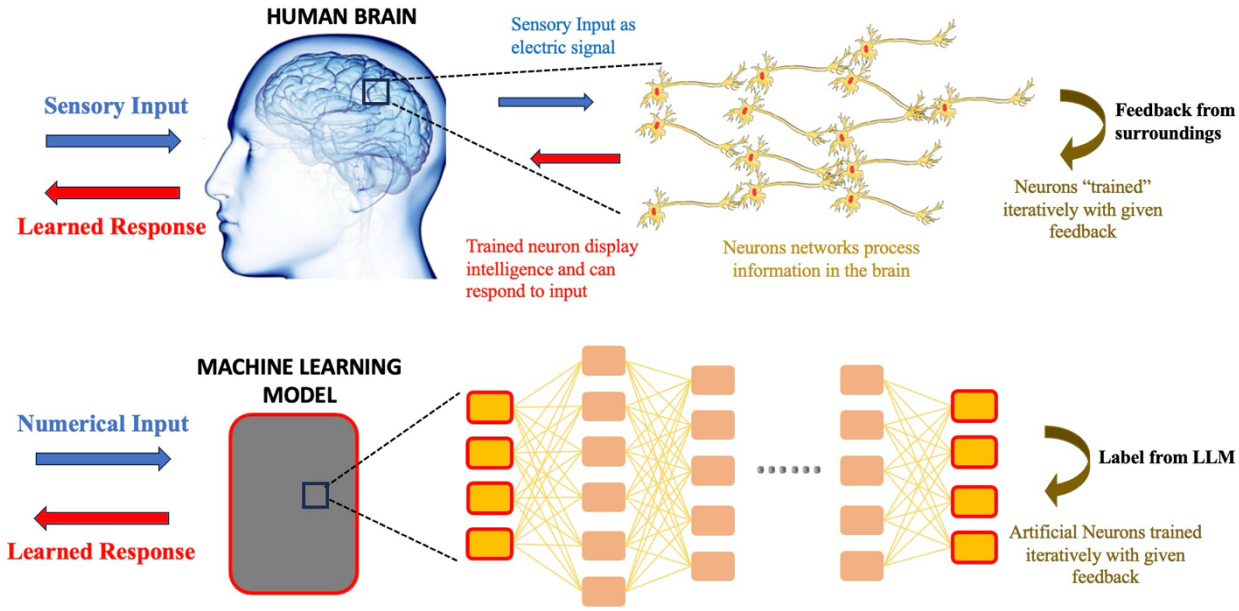
As we explore the processes of infant learning and artificial intelligence, it becomes apparent that their statistical learning mechanisms are closely intertwined (Abbate, 2023; Dupoux, 2018; Wang et al., 2023). Dupoux (2018) characterizes infants and AI as "learners" when confronted with new information and suggests a learning strategy that applies to both human infants and AI: start with simplicity and rely on actual raw data. Since the learning processes of infants and AI align so closely, we attempt to answer the question raised in the introduction: Does machine learning genuinely model infant learning? Before addressing this question, it is essential to note that this paper focuses on the early stages of infant learning, during which infants have not

yet comprehended language. Once infants develop enough understanding to generally comprehend sentences spoken by caregivers and begin explicitly learning the language (akin to an elementary-level English course), they have progressed beyond the scope of this discussion.

When acquiring spoken language, the auditory system transmits sound signals to the brain cortex for processing, where individual neurons function within a sophisticated system that facilitates comprehension (Hickok & Poeppel, 2015). While the intricate workings of neurons and the brain are beyond the scope of this paper, it is crucial to acknowledge the neurons' pivotal role in information processing, enabling humans to derive meaning from linguistic signals (Figure 3a). At this early stage, infants encounter words devoid of understanding, similar to the initial stages of AI grappling with human language. In language acquisition, infants rely on statistical learning mechanisms (Saffran, 2003) to discern structural and semantic regularities within sentences, gradually developing linguistic comprehension by iteratively refining their lexicon (Arciuli & Monaghan, 2009). Similarly, AI demonstrates adeptness at deciphering the semantics embedded within sequential data, including natural language and biological sequences (Wang et al., 2023). Using statistical methodologies, AI acquires proficiency in human language by discerning patterns and regularities through extensive data-driven processes facilitated by neurons in ANN and labels provided by LLM, as shown in Figure 3b (Dupoux, 2018; Goodfellow et al., 2016; Wang et al., 2023). This approach allows AI to grasp aspects of lexicon and

Figure 3.
Comparative Analysis of Human and Artificial Neural Networks

(a) Top figure, human neural network.
(b) Bottom figure, Artificial Neural Network (ANN). Utilizing numerical signals as input, individual neurons within the model collaborate to conduct regressions and generate predictions. Through iterative training by comparing these predictions with labels, the neurons change coefficients in their regression to enhance their responses over time, exhibiting intelligence.



Note. When exposed to sensory input signals, the individual neurons within the human neural network respond accordingly. Through iterative adjustments with contextual cues, the neurons improve their responses over time, represented by an improved understanding of vocabulary, showcasing intelligence. In parallel, individual neurons within the Artificial Neural Network collaborate to conduct regressions and generate predictions for given numerical signals. Through iterative training by comparing these predictions with labels, the neurons change coefficients in their regression to enhance their responses over time, exhibiting intelligence.

grammatical structures inherent in human language, resembling the learning trajectory observed in infants.

The predictive capabilities observed in AI models, in a sense, also mirror the learning process of infants. Both involve processing input information through a neural network, where groups of individual neurons collaboratively perform language acquisition tasks (Goodfellow et al., 2016; Hickok & Poeppel, 2015), albeit with significant differences between the functions of infant neurons and artificial neurons. In infant learning, as outlined previously, infants reassess and refine their understanding of words through iterative encounters. Upon hearing a word previously encountered, they retrieve their initial interpretation and juxtapose it with the word's current usage, considering various probabilistic cues. These cues encompass both vertical and horizontal relationships within language structures, transitional probabilities, and phonotactic patterns, among others. A fine-tuning of comprehension then occurs based on the new context, as described by Arciuli and Monaghan (2009). Grammar acquisition in infants operates under a similar mechanism. Infants compare their pre-existing grammatical framework to new linguistic contexts, adjusting as necessary (Gervain et al., 2020). The process described above resembles one epoch in the machine learning paradigm. As shown in Figure 3a, the neural network of the infant brain, which receives linguistic signals as inputs, processes these signals and generates an output reflecting preliminary comprehension. This output, when confronted with new contextual data (akin to a label in ANN), is subject to revaluation and adjustment, mirroring the process by which an ANN adapts the weights and biases of its neurons to better fit provided labels (see Figure 3b). Occasionally, infants find that their initial understanding is incongruent with the new context—similar to an ANN's initial set of poorly performing coefficients that result in significant loss. However, just as an ANN iterates and optimizes its coefficients in successive epochs, infants also refine their linguistic understandings upon realizing discrepancies in new encounters. Indeed, a parallel can be drawn between the functioning of human neural networks and artificial neural networks, with both involving inputs, predictions, and iterative adjustments based on feedback.

Another notable similarity between infants and AI lies in the phenomenon of emergence. As Wei et al. (2022) argue, emergence occurs when quantitative changes in a system lead to qualitative changes in behavior. In essence, both infants and AI demonstrate intelligence that transcends the information provided in their training data, representing a qualitative shift indicative of genuine intelligence. As infants progress in language acquisition and master vocabulary and grammar, their understanding surpasses probabilities and correlations, marking a significant qualitative change. Similarly, scientists have observed the phenomenon of emergence in various AI models (Wei et al., 2022). For instance, when queried about the possibility of inserting a stick into human ears, GPT surprised researchers by responding affirmatively. It justified its answer by referencing the "golden cudgel" from the Chinese literary work Journey to the West, suggesting that if the stick possessed magical properties described in the literature, it could be inserted into ears. This unexpected connection made by GPT exemplifies emergence, where its extensive training data results in a qualitative change, enabling GPT to draw novel connections and demonstrate intelligence beyond its initial programming (OpenAI, 2023).

Differences Between Infant Learning and AI

Despite the remarkable similarities between the learning processes of human infants and AI, a fundamental distinction arises concerning cognition and its application, primarily in the utilization of statistical techniques (Dupoux, 2018). In other words, the underlying mechanisms of learning for infants and classical AI diverge. However, it is noteworthy that advancements in AI are increasingly aligning with the problem-solving approaches characteristic of human intelligence, similar to those employed by infants.

The process of classical machine learning represented by artificial neural networks (ANN) is strongly supervised, which, as Dupoux (2018) argues, marks a considerable difference from infant learning. Infant learning can be likened to unsupervised learning, where infants are not explicitly provided with a one-to-one correspondence between inputs and outputs but are instead tasked with processing vast amounts of information to derive understanding. This is somewhat analogous to unsupervised machine learning, such as clustering, in which a model is presented with large datasets and tasked with organizing them into clusters in an unspecified manner (Goodfellow et al., 2016). Operating without explicit instructions or labels guiding the model toward a specific task, unsupervised machine learning reflects a process closer to infant language acquisition.

As a result, in language acquisition, infants do not have the advantage that machines do in receiving and comprehending clear and explicit labels (Goodfellow et al., 2016). Machine learning models already excel in the realm of binary language; their task lies in mapping human language to machine language under close

supervision (Lake et al., 2016). In contrast, infants are not born with a pre-existing language, nor do they have another language to refer to when acquiring their first language. Essentially, infants are "on their own" in understanding a new language; they must develop their understanding independently, without pre-existing labels to reference—similar to unsupervised machine learning processes. This process underscores the unique intelligence of human infants.

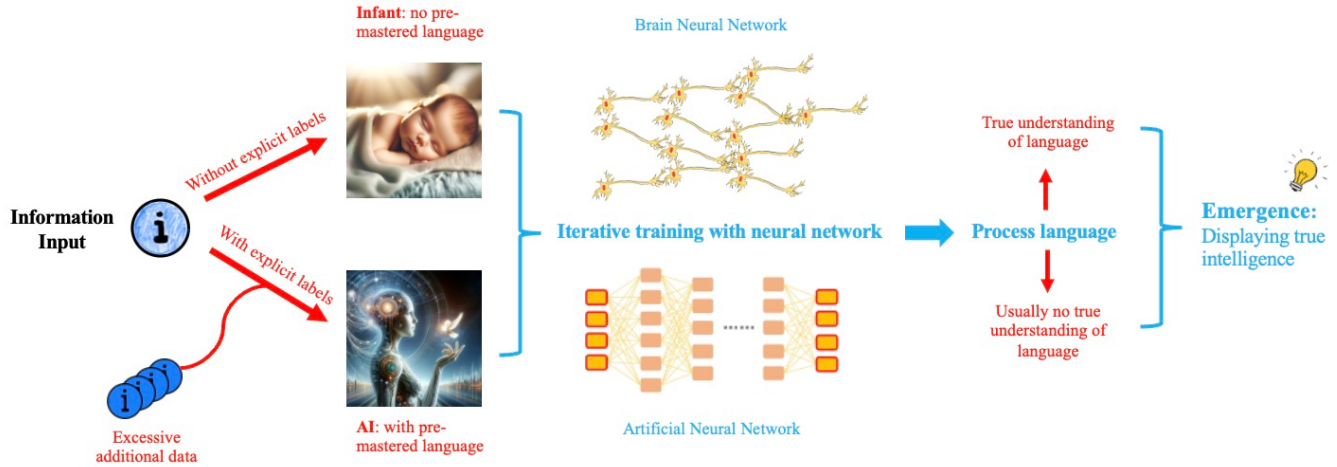
Another notable distinction between human and machine learning processes lies in the volume of data required for learning. Dupoux (2018) highlights that machines typically require around 10,000 hours of learning to accomplish a task that a four-year-old child achieves within 700-4,000 hours. Frank (2023) also observes that the amount of language input for training GPT-3 was on the order of 101210^{12}1012 words, whereas the upper bound of word input for human learners is below 10910^9109. While human learners begin with mere probabilities, their understanding evolves beyond correlations over time, gradually developing a deeper comprehension of language (Denison et al., 2013). In contrast, even after extensive learning, both primitive and advanced AI models often rely on coefficients to determine regressions, lacking the nuanced understanding observed in human learners. Figure 4 summarizes the overall similarities and differences between infant learning and AI training discussed in this section.

It is worth noting that AI faces the daunting task of comprehending a language crafted by human intellect. Dupoux (2018) eloquently outlines three key aspects of human language: its inherent complexity, flexible representations, and capacity as a finite computational system capable of generating an infinite array of expressions—a framework finely attuned to the human brain. These attributes are ubiquitous across all human languages, posing a formidable challenge for AI, which operates within the confines of binary logic. Thus, it is unsurprising that AI might take longer to wrap its circuits around this complex linguistic labyrinth. Attempting to navigate the intricacies of language through brute-force regression in hyperspace would likely prove as futile as trying to teach an infant calculus before they have mastered counting blocks.

Conclusions

This review has examined the nuanced interplay between infant learning and artificial intelligence, focusing particularly on the role of probabilistic learning methods and their implications for language acquisition.

Figure 4.
Comparative Analysis of Language Acquisition in Infants and AI



Note. This figure contrasts the learning pathways of infants and artificial intelligence in language development. Red arrows and text highlight the differences, emphasizing the absence of pre-mastered language and the reliance on unsupervised learning in infants, contrasted with the binary language comprehension and supervised learning in AI. Blue arrows and text indicate similarities, including the iterative nature of learning via neural networks and the emergence of complex language understanding. This graphical representation underscores the intricacies and nuances in the learning processes of two fundamentally different yet strikingly similar learners.

Our analysis reveals that both infants and artificial intelligence systems fundamentally rely on probabilistic mechanisms to process complex stimuli and learn from their environments. This shared approach underscores a fundamental similarity in how learning potentially occurs across biological and artificial systems, suggesting that AI can indeed mimic certain aspects of human cognitive development.

As the boundaries between human cognition and artificial intelligence continue to blur, we stand on the brink of a new era where the synergistic relationship between these fields could catalyze significant breakthroughs. The convergence of infant learning strategies and AI technologies provides not only a unique perspective on the future of cognitive development but also positions AI as a powerful tool that could reshape our understanding of learning and intelligence. This emerging dynamic has profound implications for both cognitive science and technology, setting the stage for future innovations that could redefine the very fabric of human-AI interaction.

It is important to recognize that the theories and methodologies discussed in this paper represent a selection of widely accepted approaches within the fields of infant learning and artificial intelligence. Both fields, however, are rich with diverse and sometimes competing theories that may offer different perspectives on the mechanisms of learning and cognition. While this review has focused on probabilistic learning due to its significant prevalence and support in recent research, the exclusion of alternative theories is not an assertion of their irrelevance but rather a necessary delimitation given the specific scope of this paper. Examining opposing views and alternative theories in detail remains beyond the scope of this review but represents a valuable avenue for future exploration.

Implications and Future Directions

To forecast the future of AI development, it is essential to examine the current landscape of advanced models and discern the direction of their progression. Successful models like GPT have notably departed from solely relying on classical supervised machine learning, as previously discussed. GPT incorporates unsupervised learning components, enabling it to operate autonomously, akin to human infants, without explicit guidance (Brown et al., 2020). Moreover, GPT's remarkable advancement lies in its adeptness at understanding context. Enhanced with various structural components, such as CNN and RNN, and more advanced tools like reservoirs and transformers (Goodfellow et al., 2016), GPT exhibits a capacity to comprehend context effectively, similar to human logic. These features, analogous to human reasoning, enable GPT to demonstrate intelligence when faced with novel challenges (Vaswani et al., 2017). As a result, advanced AI models aspire (perhaps unintentionally) to emulate human learners by developing capabilities such as understanding context and autonomously finding insights without explicit labels (Vaswani et al., 2017). Thus, we see a trend moving from supervised to unsupervised learning, from brute-force regression toward methods resembling human logic, and from purely data-driven approaches toward incorporating elements akin to human intuition.

AI development is currently focused on two key directions. The first direction involves augmenting model complexity by integrating more sophisticated components, allowing the model to address problems with "logical" reasoning similar to that of humans. These extensions enable AI models to gain a better understanding of context and enhance problem-solving capabilities. The second direction is advancing toward unsupervised learning, analogous to infant learning. With greater autonomy, the phenomenon of emergence—a hallmark of intelligence—is more likely to occur. For instance, the latest OpenAI model, Sora, can generate a video of two pirate ships fighting on a cup of coffee (OpenAI, 2024), a scenario that could not have been directly learned from any existing data. This video clip demonstrates both intelligence and emergence during Sora's training. However, the precise mechanisms and conditions for such qualitative changes in learning, and the amount of data required, remain open questions (Wei et al., 2022). Nevertheless, having more high-quality data for training can facilitate this transformation, with the hope that an increase in quantity will eventually lead to a qualitative shift.

As AI advances in language processing, the possibility emerges for AI to provide feedback to human learners, particularly infants. By emulating the reasoning processes of infants, or humans more broadly, certain AI advancements—such as emergence—may offer insights into infant language acquisition. Consequently, developments in AI could deepen our understanding of human behavior and cognition. Future research will likely extend beyond improving AI models that learn from humans to utilizing AI as a tool to study human behavior (Tan et al., 2024). This trend suggests that AI could play a dual role: as a means to enhance our knowledge of human behavior and as a potential aid in optimizing language acquisition methods for infants. This reciprocal feedback loop implies that infants could learn more efficiently from AI, while AI, in turn, gains insights from observing and understanding the language learning

processes of infants. The symbiotic relationship between AI and infant learning holds promise for advancing our understanding of language acquisition and potentially enhancing strategies for infant language learning.

While AI development offers exciting possibilities, it also raises significant ethical concerns. The advent of AI could profoundly reshape societal structures, such as displacing workers by automating tasks more cheaply and efficiently than humans can perform. More profoundly, if AI models were to significantly enhance human learning capabilities, they could lead to feelings of alienation, as the tools we created begin to surpass us in intelligence and instruct us on optimal behaviors and strategies. Such a shift could challenge our understanding of our role and status on this planet. Historically, similar ethical dilemmas have arisen; for example, the cloning of Dolly the sheep sparked a global debate on the limits of scientific exploration, leading to stringent regulations on cloning technologies. The ethical concerns posed by AI are even more consequential, with potential implications for human significance and existence itself. These considerations underscore the necessity for prudence and foresight in AI development, ensuring that technological advancements do not outpace our ability to manage their impacts responsibly.

Despite the ethical concerns associated with AI's advancement, an optimistic perspective might be more productive, viewing AI as a novel lens that reveals a dynamic panorama of possibilities (Epstein et al., 2023). As we enter a new AI era, our vision expands beyond mere technological advancement. In the unfolding narrative of technological evolution, we witness the inception of AI autonomy. Much as humans rely on intuition and insight to navigate life's complexities, AI systems independently parse, assimilate, and analyze data, making decisions based on patterns they discover. This emerging autonomy is likely to grow as AI sophistication increases. Although the nature of human and AI intelligence differs fundamentally, both share an inherent autonomy in engaging with the world. Thus, it is with profound insight that we recognize our partnership with AI in the grand journey of learning and civilization, signifying that humanity, in its quest for knowledge, no longer walks alone.

References

Abbate, F. (2023). Natural and artificial intelligence: A comparative analysis of cognitive aspects. *Minds and Machines*, *33*(4), 791-815. https://doi.org/10.1007/s11023-023-09646-w

Anastasi, A. (1986). Evolving concepts of test validation. *Annual Review of Psychology*, *37*(1), 1–15. https://doi.org/10.1146/annurev.ps.37.020186.000245

Arciuli, J., & Monaghan, P. (2009). Probabilistic cues to grammatical category in English orthography and their influence during reading. *Scientific Studies of Reading*, *13*(1), 73–93. https://doi.org/10.1080/10888430802633508

Asimov, I. (2004). *I, Robot*. Bantam Books.

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., … Amodei, D. (2020). *Language models are few-shot learners*. ArXiv. https://doi.org/10.48550/arXiv.2005.14165

Copeland, B. (2023). Artificial Intelligence. In *Encyclopedia Britannica*. https://www.britannica.com/technology/artificial-intelligence

Denison, S., Reed, C., & Xu, F. (2013). The emergence of probabilistic reasoning in very young infants: Evidence from 4.5- and 6-month-olds. *Developmental Psychology*, *49*(2), 243–249. https://doi.org/10.1037/a0028278

Dupoux, E. (2018). Cognitive science in the era of artificial intelligence: A roadmap for reverse-engineering the infant language-learner. *Cognition*, *173*, 43–59. https://doi.org/10.1016/j.cognition.2017.11.008

Epstein, Z., Hertzmann, A., Herman, L.M., Mahari, R., Frank, M.R., Groh, M., Schroeder, H., Smith, A., Akten, M., Fjeld, J., Farid, H., Leach, N., Pentland, A., & Russakovsky, O. (2023). Art and the science of generative AI. *Science*, *380*(6650), 1110–1111. https://doi.org/10.1126/science.adh4451

Estes, K. G. (2009). From tracking statistics to learning words: Statistical learning and lexical acquisition. *Language and Linguistics Compass*, *3*(6), 1379–1389. https://doi.org/10.1111/j.1749-818X.2009.00164.x

Faraboschi, P., Frachtenberg, E., Laplante, P., Milojicic, D., & Saracco, R. (2023). Artificial general intelligence: Humanity's downturn or unlimited prosperity. *Computer*, *56*(10), 93–101. https://doi.org/10.1109/mc.2023.3297739

Frank, M. C. (2023). Bridging the data gap between children and large language models. *Trends in Cognitive Sciences*, *27*(11), 990–992. https://doi.org/10.1016/j.tics.2023.08.007

Gervain, J., de la Cruz-Pavía, I., & Gerken, L. (2020). Behavioral and imaging studies of infant artificial grammar learning. *Topics in Cognitive Science*, *12*(3), 815–827. https://doi.org/10.1111/tops.12400

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*.

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). *Generative adversarial networks*. ArXiv. https://doi.org/10.48550/arXiv.1406.2661

Hickok, G., & Poeppel, D. (2015). Neural basis of speech perception. *Handbook of Clinical Neurology*, *129*, 149–160. https://doi.org/10.1016/B978-0-444-62630-1.00008-1

Johnson, M., Albizri, A., Harfouche, A., & Fosso-Wamba, S. (2022). Integrating human knowledge into artificial intelligence for complex and ill-structured problems: Informed artificial intelligence. *International Journal of Information Management*, *64*, Article 102479. https://doi.org/10.1016/j.ijinfomgt.2022.102479

Kim, Y. (2014). *Convolutional neural networks for sentence classification*. ArXiv. https://doi.org/10.48550/arXiv.1408.5882

Lake, B. M., Ullman, T. D., Tenenbaum, J. B., & Gershman, S. J. (2016). Building machines that learn and think like people. *Behavioral and Brain Sciences*, *40*, Article e385. https://doi.org/10.1017/S0140525X16001837

Lipton, Z. C., Berkowitz, J., & Elkan, C. (2015). *A critical review of recurrent neural networks for sequence learning*. ArXiv. https://doi.org/10.48550/arXiv.1408.5882.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). *Efficient estimation of word representations in vector space*. ArXiv. https://doi.org/10.48550/arXiv.1301.3781

Miller, G. A. (1963). *Language and communication* (Revised edition).

Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of Child Language*, *17*(2), 357–374. https://doi.org/10.1017/S0305000900013817

OpenAI. (2023). *ChatGPT* [Large language model]. https://chat.openai.com

OpenAI. (2024). *Sora* [Artificial intelligence model]. https://openai.com/index/sora/

Pothos E. M. (2007). Theories of artificial grammar learning. *Psychological Bulletin*, *133*(2), 227–244. https://doi.org/10.1037/0033-2909.133.2.227

Powell, J. (2019). Trust me, I'm a chatbot: How artificial intelligence in health care fails the Turing test. *Journal of Medical Internet Research*, *21*(10), Article e16222. https://doi.org/10.2196/16222

Saffran, J. R. (2003). Statistical Language Learning: Mechanisms and Constraints. *Current Directions in Psychological Science*, *12*(4), 110-114. https://doi.org/10.1111/1467-8721.01243

Saffran, J. R. (2020). Statistical language learning in infancy. *Child Development Perspectives*, *14*(1), 49–54. https://doi.org/10.1111/cdep.12355

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928. https://doi.org/10.1126/science.274.5294.1926

Shevlin, H., Vold, K., Crosby, M., & Halina, M. (2019). The limits of machine intelligence. *EMBO Reports*, *20*(10), Article e49177. https://doi.org/10.15252/embr.201949177

Shultz, T. R., & Nobandegani, A. S. (2022). A computational model of infant learning and reasoning with probabilities. *Psychological Review*, *129*(6), 1281–1295. https://doi.org/10.1037/rev0000322

Tan, A. W. M., Yu, S., Long, B., Ma, W. A., Murray, T., Silverman, R. D., Yeatman, J. D., & Frank, M. C. (2024). *DevBench: A multimodal developmental benchmark for language learning*. ArXiv. https://doi.org/10.48550/arXiv.2406.10215

Turing test. (2024, December 30). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Turing_test&oldid=1266226188

Vaswani, A., Shazeer, N.M., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. *Proceedings of the Thirty-First Annual Conference on Neural Information Processing Systems: Advances in Neural Information Processing Systems*, *30*, 1–11. https://doi.org/10.48550/arXiv.1706.03762

Wang, H., Fu, T., Du, Y., Gao, W., Huang, K., Liu, Z., Chandak, P., Liu, S., Van Katwyk, P., Deac, A., Anandkumar, A., Bergen, K., Gomes, C. P., Ho, S., Kohli, P., Lasenby, J., Leskovec, J., Liu, T.-Y., Manrai, A., … Zitnik, M. (2023). Scientific discovery in the age of artificial intelligence. *Nature*, *620*(7972), 47–60. https://doi.org/10.1038/s41586-023-06221-2

Wei, J., Tay, Y., Bommasani, R., Raffel, C., Zoph, B., Borgeaud, S., Yogatama, D., Bosma, M., Zhou, D., Metzler, D., Chi, E. H., Hashimoto, T., Vinyals, O., Liang, P., Dean, J., & Fedus, W. (2022). *Emergent abilities of large language models*. ArXiv. https://doi.org/10.48550/arXiv.2206.07682

Wojcik, E. H., & Saffran, J. R. (2015). Toddlers encode similarities among novel words from meaningful sentences. *Cognition*, *138*, 10–20. https://doi.org/10.1016/j.cognition.2015.01.015

Xu, F., & Garcia, V. (2008). Intuitive statistics by 8-month-old infants. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(13), 5012–5015. https://doi.org/10.1073/pnas.0704450105

Zhang, C., Bengio, S., Hardt, M., Recht, B., & Vinyals, O. (2016). *Understanding deep learning requires*

rethinking generalization. ArXiv. https://doi.org/10.48550/arXiv.1611.03530

Author Note
This work was inspired by the research and insights of Dr. Jenny Saffran and her lab. We extend our gratitude to Dr. Saffran for her encouragement and invaluable feedback. We appreciate not only her insights but also her contributions to editing this paper.
We have no known conflict of interest to disclose.

Correspondence concerning this article should be addressed to Rong Fan, University of Wisconsin-Madison, Madison, WI 53706, United States (443) 498-3238. Email: rfan44@wisc.edu

Journal of Psychological Inquiry
2025, Vol. 29, No. 2, pp. 62-71

DO THE SEMANTIC ASSOCIATIONS OF VISUAL ITEMS INFLUENCE VISUAL SEARCH PERFORMANCE?

ZOE C. FRANCIS AND KENITH V. SOBEL

UNIVERSITY OF CENTRAL ARKANSAS

Abstract – In traditional visual search experiments, participants search through displays that contain a single target item and varying numbers of nontarget distractors. Targets and distractors had traditionally been defined by distinct visual features such as their color and shape, but eventually researchers began to wonder whether search items' semantic associations, such as an A that is a letter and an 8 that is a number, could affect search performance. Although the initial attempt to do so demonstrated that category membership (i.e., numbers versus letters) influenced visual search, repeated failures to replicate the initial effect discouraged further attempts for the next two decades. One major problem with such studies was that any effects of semantic associations could also be explained as resulting from shape differences: an A is in a different category than an 8, but they also have different shapes. A breakthrough study managed to find that visual search was more efficient when the distractors' shapes had the same semantic identity (i.e., B and b) than when they had different identities (i.e., B and p), while controlling for the search items' shapes. The author of this breakthrough argued that his results undermined the classical perception-cognition divide, whereby higher level cognitive processes such as knowledge and expectation have traditionally been presumed not to affect basic perceptual processing. Nonetheless, in this review we argue that this study as well as later studies that built on its success can be interpreted to be consistent with the classical perception-cognition divide. Specifically, we argue that search items' semantic associations affect the role of cognitive processing in visual search rather than the basic perceptual processing.

Keywords: visual search, semantic associations

Much of life is spent searching for things: looking for your car in a mall parking lot, searching for a friend's familiar face in a crowded room, or trying to find your favorite cereal in the packed grocery store aisle. In each of these scenarios there are several items to search through and just one item that is the target, whether that be a car, a friend, or a box of Cheerios. What guides visual searches for these targets? How do people process what they are seeing and discover where the target is? A reasonable first guess might claim that search begins by picking one visual location at random, then checking to see if the target is at the chosen location, and if not, randomly selecting another location. But the visual search literature shows that searches rely on more efficient strategies than this kind of random selection. When looking for a specific target, searchers begin with the target's defined visual features, and limit the range of search to just the items that have that feature (Bacon & Egeth, 1997). Distractor items that do not have the target feature are eliminated from consideration. By using target features to limit the range of search items, these features are said to be guiding features (Wolfe, 1998).

This reflects the long-held view that visual search performance is generally presumed to be influenced by the visual features of the visual display items, and that any meanings associated with those items have no bearing on search performance. Indeed, in the seminal paper that laid the groundwork for the bulk of the visual search literature (Treisman & Gelade, 1980), the target letter T is distinct from the distractor letter X due to their shapes, rather than the fact that they have different roles in written language. The notion that perceptual and cognitive processing operate independently was later formalized (Pylyshyn, 1999; Firestone & Scholl, 2015) as the perception-cognition divide. The perception-cognition divide asserts that while perception can

undoubtedly influence aspects of cognition such as thoughts and beliefs, the information flows in just one direction from perception to cognition but not the other direction, because thoughts and beliefs cannot alter the way the world looks, sounds, and feels. Cognition can play a role in the way we interpret perceptions, but can't alter basic perceptual appearances. This idea of the perception-cognition divide was later applied in the context of visual search, by Wolfe and Horowitz (2004), who reviewed the visual search literature to find what attributes influence visual search performance. The authors found that visual features such as color and shape definitely guide search. Previous attempts to find effects of semantic associations, however, such as whether an item was a letter versus a number, had been unsuccessful, leading Wolfe and Horowitz to doubt that these kinds of attributes guide search. The distinction between visual attributes that definitely guide search and semantic attributes that are doubted to guide search supports the idea of the perception-cognition divide; color and shape work at the perceptual level and semantic associations work at the cognitive level.

The arguments presented by Wolfe and Horowitz (2004) suggested that any attempts to find an effect of semantic associations on visual search would be misguided. Nevertheless, Lupyan (2008) took their conclusions as a challenge, and found a new way to reveal the role of semantic associations in visual search, thereby opening the door for others to find new ways in which semantic associations might influence visual search. In this review, we begin with early attempts that failed to find any effect of semantic associations on visual search, then continue with Lupyan's groundbreaking work that opened the door, then highlight the work of other researchers who built on the foundation he laid. For example, in searches that use numerical targets, the distance on the number line between the target's and distractors' positions affects visual search efficiency, as does the congruity between the target's numerical and physical size (i.e., font size). The Stroop effect, in which incongruities of word meaning and color slow responses, had a similar effect on visual search performance. These effects demonstrate that semantic associations *do* influence visual search. Nevertheless, we will argue that all of these results can be interpreted as consistent with the perception-cognition divide, because semantic associations influence mechanisms such as working memory that inhabit the cognitive side of the divide.

Visual Features Versus Semantic Associations

In their review of the kinds of features that can guide search, Wolfe and Horowitz (2004) argued that color and shape are undoubtedly visual features that can be used to guide search, a conclusion backed by decades of research. During visual search a guiding feature is used in the searching process to distinguish a target item from nontarget distractor items. For example, if you were partial to green Skittles candy, you would like to pick some out from a bowl filled with M&Ms and Skittles all mixed together. The green skittle is your target item, so you could start your search by removing all the candies that are not green from consideration. Human visual systems naturally have this ability to differentiate colors in search. Once you reduce the range of items to just those with the target color, locating the items with the target color that are labeled with an "S" shape would lead you to the target. Studies involving undoubted attributes such as color and shape have shown response time (RT) patterns indicating that participants can limit their search to just the items that display a target color or shape while ignoring the items that have a different color and shape (Egeth et al., 1984).

In contrast to visual features like color and shape, Wolfe and Horowitz (2004) argued that the semantic associations of a search item are unlikely to be useful in guiding search. Semantic associations refer to when a word or symbol has a meaning assigned to it, such as the numeral 4 that has a distinct shape, but also has an associated meaning. Students learn in grade school that each numeral has its own numerical size and corresponding location on a number line. Suppose participants were asked to search for a 7 or an 8 among several 3s, and if performance is different for 7 compared to 8, it would be tempting to conclude that the difference is attributable to numerical size, such that a 7 is closer to the 3s on the number line than 8. But Wolfe and Horowitz argued that there are two problems with this conclusion. First, while the human visual system can readily distinguish items on the basis of their color and shape, it is unimaginable that innate visual processing mechanisms would organize visual items by numerical size. Second, the 7 has a shape that is much more distinct from the shape of the 3s relative to 8, so any effect of numerical size could be driven by the visual feature of shape rather than semantic associations. Wolfe and Horowitz based their argument on experiments that tried to find effects of semantic associations on visual search, which began with Jonides and Gleitman (1972).

In an early attempt to investigate whether semantic associations can affect visual search performance, Jonides and Gleitman (1972) asked participants to search for a circle target; in one condition, participants were instructed to search for the "zero", and in another condition to search for the "oh". A zero is a number and an oh is a letter, so they belong to different

semantic categories even though they look the same. The circular target was presented among various numbers of nontarget distractors that were either in the same semantic category or a different category. For example, an oh letter target among the numbers 5, 6, 8 and 9 would be different target and distractor categories, whereas a zero number target would be the same category as these four number distractors. When the targets and distractors were from the same semantic category, RTs increased with the number of search items, so in other words responses were slower when searching among six items than when searching among two items, but when targets and distractors were from different categories, RTs were flat across increasing numbers of search items. This distinction between steeply increasing RTs as a function of the number of search items versus flat RT is a classic way to indicate the efficiency of search (Wolfe, 1998), where flat RT functions indicate relatively efficient search compared with tasks that elicit steep RT functions. These results, in which the RT functions are steeper when targets and distractors are the same category is analogous to the role of color in search: when target and distractor colors are similar, RT functions are steeper than when target and distractor colors are distinct (D'Zmura, 1991). The results from Jonides and Gleitman provided early support for the argument that semantic associations can affect visual search performance.

Duncan (1983) tried and failed to replicate the Jonides and Gleitman (1972) oh-zero effect in his own experiment. In Duncan's attempt to replicate Jonides and Gleitman's effect, he designed an experiment using the same set of stimuli and instructions. Distractor items were either letters or numbers and participants were instructed that the target was either an oh or zero. Duncan's results led him to conclude that there is no effect from different target and distractor categories. Even though the number zero and letter oh are from different categories, Duncan argued that the visual feature of shape is more important in visual search than the semantic associations of category membership.

Krueger (1984) built on Duncan's (1983) failure to replicate Jonides and Gleitman (1972) by expanding the target and distractor sets, and controlling the shapes of the items in the numerical and alphabetical categories. Participants were instructed to search for a target that was a 5, 6, S, or G. The target shapes were designed so the 5 was similar to the S, and the 6 was similar to the G. The numerical distractors were 1, 2, 3, 4, 7, 9 and the letter distractors were L, Z, B, K, J or P. As with the targets, the distractors were matched for appearance so the 1 and L were similar, as was the 2 and Z, 3 and B, 4 and K, 7 and J, and 9 and P. Krueger found that when letters and digits had physically similar appearances, the target being in a different category from distractors (e.g., a 5 target among Z, B, K, J, and P distractors) did not lead to an advantage in RTs. These results supported Duncan's argument that distinguishable shapes rather than category membership drives visual search performance. Taken together, these studies (Duncan, 1983; Jonides & Gleitman, 1972; Krueger, 1984) support the idea of the perception-cognition divide. Jonides and Gleitman (1972) provided early evidence challenging the perception-cognition divide, but then both Duncan (1983) and Krueger (1984) failed to replicate these findings, suggesting that cognitive processing in the form of semantic associations has no role in visual search performance.

Breakthrough: Letters That Are the Same or Different

Based on the results of Duncan (1983) and Krueger (1984), Wolfe and Horowitz (2004) argued that the ability for semantic associations of alphanumeric characters to guide search is doubtful. It seemed as though the question was settled: the semantic associations of alphanumeric characters do not guide search. This conclusion is consistent with the classical perception-cognition divide. As described by Pylyshyn (1999), higher level cognitive processes such as knowledge and expectation do not influence lower level perceptual processing, or in other words, perception is cognitively impenetrable. But where Pylyshyn provided evidence about numerous different kinds of visual processing, Wolfe and Horowitz's claims are specific to visual search performance.

Lupyan (2008) was not satisfied with this consensus and worked to reintroduce the idea of semantic associations having a direct effect on visual processing. The results from Duncan (1983) and Krueger (1984) showed that category membership was unlikely to affect search performance, so Lupyan used search items whose semantic similarity was more specific than category membership. Furthermore, Lupyan knew that he needed to carefully control the shapes of the search items. To do so, Lupyan had participants search for a character whose shape looks to modern eyes like a blend of a "b" and "p", but in the Old English alphabet, it was called the thorn (þ). Most modern English readers do not recognize the thorn, but the thorn's visual similarities to both letters b and p could be expected to activate participants' visual memories for both letters. Participants were instructed to localize the þ target from among (upper case) B and (lower case) p nontarget distractors in one condition, and among (upper case) B and (lower case) b distractors in the other condition. In the B and p condition, the distractors were not just two

different shapes but were also semantically distinct because the B and p are two different letters, but in the B and b condition, the distractors were the same letter. Visual search was less efficient when the distractors were two different letters (i.e., B and p) than when distractors were the same letter (i.e., B and b). By using specific shapes, rather than general categories, and controlling the shapes of the target and distractors, Lupyan managed to find an effect of semantic associations on visual search where Duncan (1983) and Krueger (1984) had failed. Further, he argued that the conceptual relationship between distractors (i.e., different letter identities for B and p versus the same identity for B and b) penetrated the perceptual processing needed to carry out the search task. That is, Lupyan argued that his results undermine Pylyshyn's argument that perception is cognitively impenetrable.

However, an alternative explanation of Lupyan's (2008) results that is consistent with Pylyshyn's (1999) claim that perception is cognitively impenetrable (later amplified by Firestone & Scholl, 2015), is that semantic associations affect the role of working memory in visual search (Kane et al., 2006) rather than visual processing per se. According to the working memory account, when search items such as those in Lupyan's experiment occupy the same conceptual identity in working memory (i.e., B and b), the mental representation is simpler than when they occupy different conceptual identities (i.e., B and p). As a result, the semantic similarity between items affects search efficiency because remembering what distractors need to be ignored requires less mental effort when the distractors occupy the same identity (B and b) than when they occupy different identities (B and p). Thus, Lupyan's results are consistent with both the cognitive penetrability and working memory accounts. While the interpretation of Lupyan's results is debatable, his primary contribution to the visual search literature was to reopen the door that had been closed by Duncan (1983) and Krueger (1984), which encouraged other experimenters to look for effects of semantic associations on visual search. Later studies built on Lupyan's breakthrough by looking at how numerical size affects visual search performance.

The Effect of Number Line Position

Schwarz and Eiselt (2012) were the first to look for effects of numerical distance on visual search efficiency. They varied the numerical distance between targets and distractors across visual displays. In an attempt to control the influence of shape on participants' search in the experiments, Schwarz and Eiselt used the digit 5 as a target in all their displays. They found that search was more efficient when the mean numerical distance between the target digit (5) and distractor digits was large (e.g., a 5 target among distractor digits 1, 2, 8, and 9) than when target-distractor numerical distance was small (e.g., a 5 target among distractor digits 3, 4, 6, and 7). This influence of number line position on visual search is equivalent to the influence of visual features such as color and shape, in which search efficiency increases with the target-distractor visual distinctiveness (Duncan & Humphreys, 1989). One limitation of the Schwarz and Eiselt study is that the target was always the same in all displays. Although the decision to use a single target item allowed Schwarz and Eiselt to control the target shape, their results may be specific to the chosen target digit of 5 and may not generalize to other target digits. Furthermore, the distractors that were numerically close to the target on the number line had different shapes than the numerically distant distractors, so the role of shape remains a plausible explanation for their results. Later studies looking at visual search for numerals developed further techniques to control for shape while broadening the target set to include more than one digit.

Godwin et al. (2014) expanded on the work of Schwarz and Eiselt (2012) by using multidimensional scaling to create a two-dimensional map of visual similarity between all the digits from 0 to 9. In the initial phase of their study, participants ranked the similarity in shape between every pair of digits in a Verdana font, and the researchers generated a two-dimensional map in which visually similar items (e.g., 2 and 5) were closer to each other on the map than visually distinct items (e.g., 4 and 8), which were farther away from each other. Godwin et al. thus had two maps, the one they built depicting visual similarity in two-dimensions, and the one-dimensional mental number line that most people develop in the course of their basic mathematical education (Pinhas et al., 2013). Thus equipped with a visual similarity and numerical similarity mental maps, Godwin et al. were able to tease apart the effects of visual and numerical similarity. They found effects of numerical size that were consistent with the results in Schwarz and Eiselt (2012) for targets other than the digit 5: search was more efficient when the numerical distance between the target and distractor digits was large than when numerical distance was small. By controlling for the visual feature of shape, the work of both Schwarz and Eiselt and Godwin et al. demonstrate a role for numerical distance in visual search that cannot be attributed to shape differences. Furthermore, numerical size is analogous to visual similarity, such that visual search efficiency increases with the target-distractor numerical distance.

Sobel et al. (2015) used a different method to control for shape differences. Instead of creating a two-dimensional map of the visual similarity of digits rendered in the font (i.e., Verdana) used in Schwarz & Eiselt (2012) and Godwin et al. (2014), they created digits from line segments as on the faces of digital clocks, depicted in Figure 1. Furthermore, the target digits in the first condition were the distractor digits in the second condition, and the distractor digits in the first condition were the target digits in the second condition. And finally, they used digits that were as similar as possible in shape (Cohen, 2009); the target digits 5 and 6 in the first condition only had one line segment difference between them (the lower left vertical line segment present in the 6 was absent in the 5), and the distractor digits 3 and 8 had two line segment differences (the left vertical segments present in the 8 was absent in the 3). Participants viewed displays that each contained one of the target digits and varying numbers of the distractor digits. When the target digits were 5 and 6 and distractors were 3 and 8, the target digits were next to each other on the number line and distractors were separated from each other; this was called the inner targets condition (see Figure 2 for target and distractor positions on the number line). When the target digits were 3 and 8 and the distractors were 5 and 6, the target digits were separated from each other and the distractor digits were next to each other on the number line; this was called the outer targets condition. Even though both conditions used the same four digits, Sobel et al. found that search was more efficient for the inner targets condition than for the outer targets condition. This indicates that participants were inclined to group targets by their number line position; when the target digits were beside each other in a number line, it allowed participants to create a simpler target set than when the target digits were separated by the distractor digits.

Figure 1

Digits rendered in a "digital clock" font to control shape differences as in Krause et al., 2017, and Sobel et al., 2016

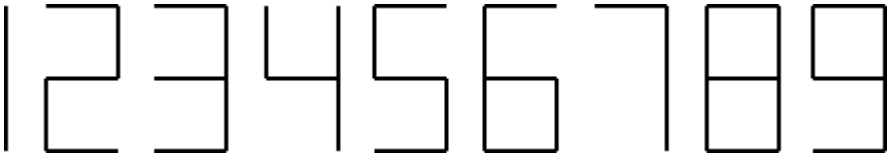
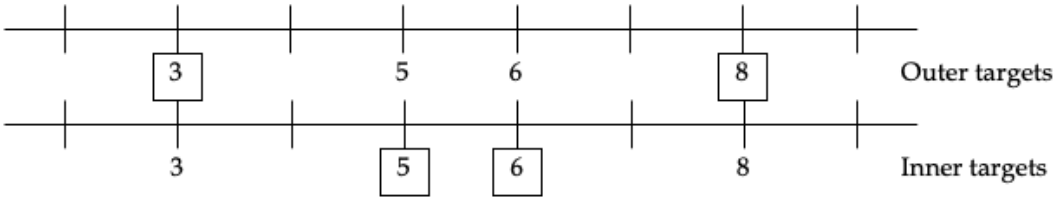


Figure 2

Number line positions for targets and distractors in Sobel et al., 2016



While these three studies (Godwin et al., 2014; Schwarz & Eiselt, 2012; Sobel et al., 2015) demonstrated that number line position can affect visual search performance, all three can be interpreted as consistent with the classical cognition-perception divide (Firestone & Scholl, 2015; Pylyshyn, 1999; Wolfe & Horowitz, 2004) because number line position affects the cognitive aspects of visual search as described in the following argument. Hout & Goldinger (2015) argued that when searching for a target item, visual items are compared to the objects stored in a mental target template to identify whether the visual items are targets. These target templates are used more in everyday life than one might expect. When planning to make a salad, you might find yourself in the produce section of the grocery store. Looking at all of the produce, the specific ingredients needed for a salad would be held in your mind by envisioning lettuce, carrots, tomatoes, and cucumber as your eyes scan the shelves. When you examine some particular produce item (e.g., asparagus), you compare the visual image of the asparagus to each item in the target template, eventually deciding that it is not one of the target items. Thus, searching for salad ingredients is not just a search through the external visual landscape, but also proceeds through an internal target template. This kind of task, in which searches proceed through both the external visual field and the internal target template is called hybrid search.

For the studies in which the numerical distance between target and distractor digits was varied (Godwin et al., 2014; Schwarz & Eiselt, 2012), comparisons to the target template can be carried out more quickly when the target and selected item are numerically distant than

when numerically close (Moyer & Landauer, 1967). When target digits are adjacent to each other on the number line (Sobel et al., 2015), maintaining the target template in working memory requires less effort than when target digits are separated from each other, which frees up mental resources and allows for efficient comparisons between selected items and target digits. With the role for numerical distance in visual search firmly established, researchers turned to the question of whether the interaction between a target's numerical size and its physical size (i.e., font size) can affect visual search performance).

Size Congruity Effect

The interaction between a digit's numerical size and physical size has traditionally been investigated using size congruity experiments. In the original size congruity experiments, participants were asked to examine two different numbers and select the larger or smaller number out of the pair based on their numerical value (Besner & Coltheart, 1979). In these experiments, the physical size (i.e., font size) was also manipulated. For the numbers 9 and 2, for example, 9 would be considered numerically larger and 2 numerically smaller. The font size of the number 2 could be increased (2, 9) to make it physically larger than 9. Therefore, in some size congruity experiments participants are asked to select the physically larger number (Henik & Tzelgov, 1982). Because both numerical and physical size can be manipulated, the numerical and physical size can be either congruent or incongruent with each other. In congruent trials the numbers shown to participants were considered either large or small both in terms of physical size and numerical size. The number 9 for example, is numerically larger than 2, so in congruent trials it would be presented in a large font size as well. There were incongruent trials as well, where a large numerical size, such as 9 was expressed in a small font size. In a typical size congruity experiment, RTs are faster in the congruent trials, where numerical size and physical size are the same, than in incongruent trials.

Risko et al. (2013) sought to understand how attention might be implicated in size congruity effects. Participants viewed two numbers and indicated the location of the numerically larger number, as in typical numerical comparison tasks such as Besner and Coltheart (1979). In control trials, both numbers appeared at the same time, but in experimental trials either the physically smaller number or physically larger number was presented first, then 100 ms later the other number appeared. The authors predicted that when one digit appeared before the other (experimental trials), the first item to appear would grab attention. As expected, the size congruity effect, as indicated by the RT difference between congruent and incongruent trials, was greater when the physically larger item appeared first than when the two items appeared simultaneously. Apparently, the item's larger *physical* size grabbed participants' attention even though they were instructed to pay attention to *numerical* size. This result supports Risko et al.'s claim that visual attention is implicated in size congruity experiments. However, Arend and Henik (2015) identified a limitation in Risko et al.: their experimenters had only asked participants to locate the numerically larger target. Arend and Henik argued that if Risko et al. had instructed participants to localize not just the numerically larger target but also the numerically smaller target, the effect of attention might have disappeared.

Nevertheless, even with the limitation identified by Arend and Henik (2015), Risko et al.'s (2013) finding that the size congruity effect depends on attention suggests that the visual attention required for visual search (Wolfe, 1998) may be affected by size congruity. After all, although Risko et al. had intended their task to be a numerical comparison, it was analogous to a visual search in which participants searched through visual displays containing just two items to localize the single target item. Emboldened by the results from Risko et al., later researchers recast their task as a visual search task containing several items (Krause et al., 2017, Sobel et al., 2016). They predicted that the size congruity effect found in numerical comparison would also arise in visual search tasks. Further, they addressed Arend and Henik's critique by asking participants not just to localize the numerically larger target, but also the numerically smaller target. The results indicated that in both cases, when participants were asked to find the numerically large target or numerically small target, there was a size congruity effect for the visual search; RTs were slower when the target's numerical and physical size were incongruent than when they were congruent.

Sobel and Puri (2018) explained this size congruity effect in terms of the target template in visual short term memory. In their first experiment, participants were instructed to search for one of two target items: one group of participants searched for numerically small targets (2 or 3), and another group searched for numerically large targets (8 or 9). Displays consisted of one of the two target digits among varying quantities of distractor digits. In Experiment 1, RTs were faster when the target's numerical and physical size were congruent than when they were incongruent, replicating the previous studies (Krause et al., 2017; Sobel et al., 2016). As in these previous studies, the distractor digits

were interleaved within a block, so participants who searched for a numerically large target could never tell whether a given display would contain an 8 or a 9 and they would need to maintain a range of digits (i.e., greater than 5) in their target template.

In a follow-up experiment, the target digits were presented in blocks rather than randomly interleaved. Thus participants who searched for a numerically large target would see an 8 in every display in one block and a 9 in every display in the other block. Presenting target digits in blocks eliminated the size congruity effect; RTs for congruent and incongruent targets were not different. Sobel et al. argued that in the first experiment, when the target digits were randomly interleaved, participants needed to load their target template with a range of numerical values (e.g., greater than 5), but in the follow-up experiment, when the target digits were arranged in blocks, participants could load their target template with the *shape* of the single digit they would see in every trial in the block. Thus, in the first experiment the physical and numerical size of the target item had the opportunity to interact in the target template, but in the follow-up experiment, targets were stored in the target template in terms of their shape, *not* their numerical size, so there was no opportunity for the target's physical and numerical size to interact in the target template. These experiments suggest that the effect of size congruity in visual search occurs primarily in the target template that is stored in visual working memory. After one of the search items has been selected, it is compared to the target template to determine whether search should be terminated as successful, or it should continue by visually selecting another search item. In Experiment 1, when the target was defined by its relative numerical size (e.g., greater than 5), any selected items with an incongruent (e.g., small) physical size delayed this decision making process. As with the effect of number line position on visual search, size congruity affects aspects of visual search processing (e.g., working memory) that occur on the cognition side of the perception-cognition divide, rather than the aspects (e.g., visual selection) that occur on the perceptual side of the perception-cognition divide.

Stroop Effect

The size congruity effect is sometimes called the numerical Stroop effect (e.g., Ativan et al., 2024; Shichel & Goldfarb, 2022), to emphasize the interference between semantic (i.e., numerical size) and perceptual features (i.e., physical size) that were first demonstrated by Stroop's (1935) seminal experiments. The fact that interference between numerical size and physical size (i.e., size congruity effect) can influence visual search performance implies that the interference between a word's meaning and ink color (i.e., Stroop effect) can similarly affect search. In modern versions of the Stroop task, participants view color words on a computer screen that are written with pixels that are either congruent with the target word's meaning (e.g., "Red" written in red pixels) or incongruent (e.g., "Red" written in blue pixels). In a Stroop condition participants identify the target's pixel color, and in a reverse Stroop condition participants identify the target's meaning instead of its pixel color. The modern day Stroop capitalizes on the advancements of technology by using computers to capture trial-by-trial RTs, something the original Stroop study could not accomplish when RTs were measured with handheld timekeepers. The difference in RTs between incongruent and congruent trials are typically larger when participants attend to the target's color (Stroop) than when they attend to the target's meaning (reverse Stroop); in other words, the Stroop effect is typically larger than the reverse Stroop effect. This asymmetry is so common that it is often referred to as the classic Stroop asymmetry (Melara & Algom, 2003). With the asymmetry being so common, some researchers have tried to find out if the asymmetry can be inverted; that is, can the Stroop effect be smaller than the reverse Stroop effect?

The answer is yes: using a task he called a matching task, Durgin (2000) inverted the classical Stroop asymmetry. In his experiment participants were shown a color word on a computer display. The meaning of the target word ("Red", "Green", "Blue", or "Yellow") would either be congruent or incongruent with its pixel color. The target word was surrounded by four colored (red, green, blue, and yellow) patches. In the Stroop task participants were instructed to attend to the target's pixel color, and to move their cursor from the target word to the color patch that matched the target word's pixels. For example, if the target word was "Red" in blue pixels, the correct response would be to move their cursor to the blue patch. In the reverse Stroop task participants were instructed to move their cursor from the target word to the color patch that matched the target word's meaning. For the target word "Red" in blue pixels, the correct response would be to move their cursor to the red patch. The results from these Stroop and reverse Stroop experiments found an inverted asymmetry; the Stroop effect was smaller than the reverse Stroop effect. Durgin attributed the inversion of the classic asymmetry to the role of translation in such a matching task. Translation refers to the participants processing the meaning of the target word, and then translating it into a visual color that can be matched to one of the color patches surrounding the target in the display. Durgin claimed

that the reverse Stroop effect was larger than the Stroop effect because translation was required in the reverse Stroop condition but was not required in the Stroop condition. Durgin's inversion of the classic Stroop asymmetry has been replicated in other matching tasks (Miller et al., 2016; Song & Hakoda, 2015; Yamamoto et al., 2016).

Sobel et al. (2020) noted that what Durgin and others (Durgin, 2000; Miller et al., 2016; Song & Hakoda, 2015; Uleman & Reeves, 1971; Yamamoto et al., 2016) called a matching task used displays containing several items, and participants were instructed to select one of the display items, which is a kind of visual search, even though the authors didn't explicitly refer to their tasks as "visual search". Sobel et al. argued that traditional Stroop experiments required identification (i.e., identify the target's color or meaning), whereas these matching tasks required localization (i.e., indicate the location of an item defined by its color or meaning), and that identification primarily depends on verbal processing but localization primarily depends on visual processing. Thus, the simple fact that matching tasks require localization should be sufficient to invert the classical Stroop asymmetry from traditional experiments that require identification, even without any need for translation.

To support their argument that the distinction between identification and localization is sufficient to invert the Stroop asymmetry without the need for translation, Sobel et al. (2020) manipulated the task requirements and eliminated the need for translation. In Experiment 1 participants viewed target color words with congruent or incongruent pixel colors, and identified the pixel color of the target word (Stroop condition) in one block and the meaning of the word (reverse Stroop condition) in the other block. They identified the target's color or meaning by pressing one of four keys on the keyboard. The Stroop effect was larger than the reverse Stroop effect, which replicated the classic Stroop asymmetry. In Experiment 2 participants were presented with a display containing a cue in the center that was either a patch of color (Stroop condition) or a color word written in black pixels (reverse Stroop condition). Surrounding the cue were four color words ("Red", "Green", "Blue", or "Yellow") each written with pixels that were one of four colors (red, green, blue, or yellow), arranged on an imaginary circle centered on the cue. The search items' meanings and pixel colors were either all congruent or incongruent. Participants pressed one of two keys to indicate whether the target was on the right or left side of the display. Even though no translation was required, the visual processing required for localization was sufficient to invert the classic Stroop asymmetry: the Stroop effect was smaller than the reverse Stroop effect. Smith et. al (2022) showed that the inversion of the classical Stroop asymmetry generalizes to a diverse sample with experimental materials delivered remotely via the internet.

We have argued that even though semantic associations can influence visual search performance, all the effects discussed prior to the word-color Stroop effect (thorn targets among B, b, and p, numerical size and grouping on the number line, size congruity effect) are nevertheless consistent with the classical perception-cognition divide (Firestone & Scholl, 2015; Pylyshyn, 1999; Wolfe & Horowitz, 2004). We are inclined to believe the same about the effect of Stroop interference on visual search, but must admit that further research will be needed to definitively support such an interpretation. Nevertheless, we can imagine something along the following lines. Eimer (2015) argues that visual search proceeds in different processing stages: preparation, guidance, selection, and recognition. Visual processing is implicated in the guidance and selection stages, and working memory is implicated in the recognition stage. As with the previously discussed effects of semantic associations on visual search, we believe that Stroop incongruity affects the recognition stage of visual search, only after visual processing has selected one of the search items. Therefore, we believe that these results are consistent with the perception-cognition divide. Future research will determine whether we are correct.

Conclusion

Visual search performance has long been presumed to be driven by visual features such as color and shape. Early attempts to find a role for the semantic associations of search items such as numbers and letters were equivocal, and a repeated failure to replicate the earliest effect of alphanumeric categories on visual search had a chilling effect on further investigations for the next two decades. But Lupyan (2008) had the courage to take another crack at it, and the insight to design target and distractor items that allowed him to show a clear effect of letter identity on visual search efficiency. And while Lupyan seemed to argue that his results sounded the death knell for the classical perception-cognition divide, his results could also be interpreted in a manner that is consistent with the divide. He also opened the door for other researchers who found effects of numerical distance, number line grouping, and numerical size congruity on visual search. These studies that extended on Lupyan further supported the claim that semantic associations affect the role of cognition in visual search rather than basic perceptual processing itself. Recent studies have found that the classic Stroop asymmetry can

be inverted in visual search localization tasks, but further research is needed to demonstrate whether this effect is attributable to the cognitive or perceptual processing in visual search.

References

Arend, I., & Henik, A. (2015). Choosing the larger versus choosing the smaller: Asymmetries in the size congruity effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(6), 1821–1830. https://doi.org/10.1037/xlm0000135

Avitan, A., Marom, D., & Henik, A. (2024). Numerical values modulate size perception. *Attention, Perception, & Psychophysics*, 1-8. https://doi.org/10.3758/s13414-024-02875-w

Bacon, W. F., & Egeth, H. E. (1997). Goal-directed guidance of attention: Evidence from conjunctive visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *23*(4), 948–961. https://doi.org/10.1037/0096-1523.23.4.948

Besner, D., & Coltheart, M. (1979). Ideographic and alphabetic processing in skilled reading of English. *Neuropsychologia*, *17*, 467–472. https://doi:10.1016/0028-3932(79)90053-8

Cohen, D. J. (2009). Integers do not automatically activate their quantity representation. *Psychonomic Bulletin & Review, 16*, 332–336.

Duncan, J. (1983). Category effects in visual search: A failure to replicate the oh-zero phenomenon. *Perception & Psychophysics, 34*, 221–232. https://doi.org/10.3758/BF03202949

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*(3), 433–458. https://doi.org/10.1037/0033-295X.96.3.433

Durgin, F. H. (2000). The reverse Stroop effect. *Psychonomic Bulletin & Review*, *7*(1), 121-125. https://doi.org/10.3758/BF03210730

D'Zmura, M. (1991). Color in visual search. *Vision research*, *31*(6), 951-966. https://doi.org/10.1016/0042-6989(91)90203-H

Egeth, H. E., Virzi, R. A. & Garbart, H. (1984). Search for conjunctively defined targets. *Journal of Experimental Psychology: Human Perception & Performance, 10,* 32–39. https://doi.org/10.1037/0096-1523.10.1.32

Eimer, M. (2015). EPS Mid-Career Award 2014. The control of attention in visual search: Cognitive and neural mechanisms. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 68*(12), 2437–2463. https://doi.org/10.1080/17470218.2015.1065283

Firestone, C., & Scholl, B. J. (2015). Cognition does not affect perception: Evaluating the evidence for 'top-down' effects. *Behavioral and Brain Sciences*. https://doi.org/10.1017/S0140525X15000965

Godwin, H. J., Hout, M. C., & Menneer, T. (2014). Visual similarity is stronger than semantic similarity in guiding visual search for numbers. *Psychonomic Bulletin & Review*, 21, 689–695. http://dx.doi.org/10.3758/s13423-013-0547-4.

Henik, A., & Tzelgov, J. (1982). Is three greater than five: The relation between physical and semantic size in comparison tasks. *Memory & Cognition*, *10*, 389–395. https://doi:10.3758/BF03202431

Hout, M. C., & Goldinger, S. D. (2015). Target templates: The precision of mental representations affects attentional guidance and decision-making in visual search. *Attention, Perception, & Psychophysics*, *77*(1), 128–149. http://dx.doi.org/10.3758/s13414-014-0764-6.

Jonides, J., & Gleitman, H. (1972). A conceptual category effect in visual search: O as letter or as digit. *Perception & Psychophysics*, *12*, 457–460. https://doi.org/10.3758/BF03210934

Kane, M. J., Poole, B. J., Tuholski, S. W., & Engle, R. W. (2006). Working memory capacity and the top-down control of visual search: Exploring the boundaries of "executive attention." *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(4), 749–777. https://doi.org/10.1037/0278-7393.32.4.749

Krause, F., Bekkering, H., Pratt, J., & Lindemann, O. (2017). Interaction between numbers and size during visual search. *Psychological Research*, *81*(3), 664–677. https://doi.org/10.1007/s00426-016-0771-4

Krueger, L. E. (1984). The category effect in visual search depends on physical rather than conceptual differences. *Perception & Psychophysics*, *35*(6), 558–564. https://doi.org/10.3758/BF03205953

Lupyan, G. (2008). The conceptual grouping effect: Categories matter (and named categories matter more). *Cognition, 108*, 566-577. https://doi.org/10.1016/j.cognition.2008.03.009

Melara, R. D., & Algom, D. (2003). Driven by information: A tectonic theory of Stroop effects. *Psychological Review*, *110*(3), 422–471. https://doi-org/10.1037/0033-295X.110.3.422

Miller, H. C., Kubicki, S., Caffier, D., Kolski, C., & Naveteur, J. (2016). The Stroop and reverse Stroop effects as measured by an interactive tabletop. *International Journal Of Human-Computer Interaction*, *32*(5), 363-372. https://doi.org/10.1080/10447318.2016.1150642

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgements of numerical inequality. *Nature*, *215*(5109), 1519–1520. https://doi.org/10.1038/2151519a0

Pinhas, M., Pothos, E. M., & Tzelgov, J. (2013). Zooming in and out from the mental number line:

Evidence for a number range effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(3), 972–976. https://doi.org/10.1037/a0029527

Pylyshyn, Z. (1999). Is vision continuous with cognition? The case of impenetrability of visual perception. *Behavioral and Brain Sciences*, *22*(3), 341–423. https://doi.org/10.1017/S0140525X99002022

Risko, E. F., Maloney, E. A., & Fugelsang, J. A. (2013). Paying attention to attention: Evidence for an attentional contribution to the size congruity effect. *Attention, Perception, & Psychophysics*, *75*, 1137-1147. https://doi.org/10.3758/s13414-013-0477-2

Shichel, I., & Goldfarb, L. (2022). The effect of proportion manipulation on the size-congruency and distance effects in the numerical Stroop task. *Memory & Cognition*, *50*(7), 1578–1589. https://doi.org/10.3758/s13421-022-01292-4

Schwarz, W., & Eiselt, A. K. (2012). Numerical distance effects in visual search. *Attention, Perception, & Psychophysics*, 74, 1098–1103. http://dx.doi.org/10.3758/s1314-012-0342-8.

Smith, M. E., Smith, M. D., & Sobel, K. V. (2022). The classic Stroop asymmetry in online experiments. *Psi Chi Journal of Psychological Research*, *27*(2). https://doi.org/10.24839/2325-7342.JN27.2.133

Sobel, K. V., & Puri, A. M. (2018). Size congruity influences visual search via the target template. *Acta Psychologica*, *183*, 66–74. https://doi.org/10.1016/j.actpsy.2018.01.003

Sobel, K. V., Puri, A. M., & Faulkenberry, T. J. (2016). Bottom-up and top-down attentional contributions to the size congruity effect. *Attention, Perception, & Psychophysics*,78, 1324–1336. http://dx.doi.org/10.3758/s13414-016-1098-3.

Sobel, K. V., Puri, A. M., & Hogan, J. (2015). Target grouping in visual search for multiple digits. *Attention, Perception, & Psychophysics*, 77, 67–77. http://dx.doi.org/10.3758/s13414-014-0761-9.

Sobel, K. V., Puri, A. M., & York, A. K. (2020). Visual search inverts the classic Stroop asymmetry. *Acta Psychologica*, *205*. https://doi.org/10.1016/j.actpsy.2020.103054

Song, Y., & Hakoda, Y. (2015). An fMRI study of the functional mechanisms of Stroop/reverse-Stroop effects. *Behavioural Brain Research*, *290,* 187-196. https://doi.org/10.1016/j.bbr.2015.04.047

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643–662. https://doi.org/10.1037/h0054651

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*(1), 97–136. https://doi.org/10.1016/0010-0285(80)90005-5

Uleman, J. S., & Reeves, J. (1971). A reversal of the Stroop interference effect, through scanning. *Perception & Psychophysics*, *9*(3-A), 293-295. https://doi.org/10.3758/BF03212651

Wolfe, J. M. (1998). What can 1 million trials tell us about visual search? *Psychological Science*, *9*(1), 33–39. https://doi.org/10.1111/1467-9280.00006

Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience*, 5, 495–501. https://doi.org:10.1038/nrn1411

Yamamoto, N., Incera, S., & McLennan, C. T. (2016). A Reverse Stroop Task with Mouse Tracking. *Frontiers in Psychology*, *7*, 670. https://doi.org/10.3389/fpsyg.2016.00670

Journal of Psychological Inquiry
2025, Vol. 29, No. 2, pp. 72-82

SYSTEMATIC REVIEW OF FEMALE-FOCUSED ATTENTION DEFICIT HYPERACTIVITY DISORDER GENETICS STUDIES

LAILA KEWALRAMANI AND GRACE A. SULLIVAN

WESLEYAN UNIVERSITY

Abstract – The narrowness of the current research in attention-deficit hyperactivity disorder (ADHD) emphasizes the need for broader demographics in terms of sex and ethnic background, more specific molecular research especially in female biological systems, and to stray away from generalizations about ADHD as a behavioral disorder. Biological females may experience ADHD differently than biological males due to their different reproductive organs, hormonal pathways, and likely socialization experiences as well, which signifies the need to look more closely at the current research centering ADHD in females. This literature review will highlight what studies have been conducted that look at ADHD genetics studies exclusively in women or females. Gene-association studies found that in females with seasonal affective disorder (SAD), variation in the 5-HT2A receptor gene is associated with impulsivity and other ADHD symptoms in childhood and SAD following into adulthood. A variant of the DRD4 gene is associated with a history of childhood ADHD in probands with bulimia nervosa (BN). Understanding the genetic variation behind symptom profiles in females with ADHD may help in understanding how the behavioral disorder presents in females, and evaluating the methodology of these studies highlights considerations about trends in female-focused research, which are discussed in context of a brief review of male-focused results using the same search process.

Attention-deficit hyperactivity disorder (ADHD) is the most commonly diagnosed neurodevelopmental psychological disorder in children, with continuance into adolescence and adulthood (Cabral et al., 2020). Considering frequent issues with attention span that present even in neurotypical children, it is crucial to establish healthy and productive practices in young people to aid in learning at these important developmental stages (Francis et al., 2021). This requires understanding obstacles such as ADHD that affect learning, processing, and memory. Understanding the genetics underlying ADHD may help individuals and their families with familial genetic risk predict their needs, at school as well as at home, as well as provide insight for more individualized treatment efforts.

Although among the extensive research on ADHD a gender disparity in the prevalence of ADHD has frequently been identified, recent studies show that in adults with ADHD, the gender disparity is reduced to the point of insignificance (Attoe & Climie, 2023). More boys than girls receive treatment for ADHD, but both are similar in the genetic variability and inheritance of ADHD (Wilens et al., 2021). There are some differences in ADHD presentation identified between men and women; it is unclear if these differences are biological or social, as the literature does not tend to distinguish between sex (male vs. female) or gender (men vs. women). For this review, we will be using these terms with the acknowledgement that sex and gender are being conflated in many of these studies due to a lack of transgender and non-binary participants.

Boys receive higher scores on the scale of disruptive behavior than girls do (Slobodin & Davidovitch, 2019), but there are insufficient genetic studies of ADHD centering female populations, so perhaps disruptiveness is not a common feature in girls with ADHD. Girls with ADHD are under-diagnosed compared to boys during childhood and adolescence, and this may be because they are more likely to have symptoms of inattention instead of impulsivity and hyperactivity like boys. Oftentimes, girls will shield their symptoms to conform to standards of their parents and teachers as well as of societal norms (Mowlem et al., 2019). Males, specifically White males, are the most common participants in these genetics studies of ADHD. The narrowness of the current research in ADHD emphasizes the need for broader demographics in terms of sex and ethnic background, more specific molecular research especially in female biological systems, and to stray away from generalizations about ADHD as a behavioral disorder (Young et al., 2020).

Due to differences in biological development that become more profound at puberty, it is reasonable to investigate biological contributions to ADHD separately for males and females. Contributions of genetic variation interact with other biological factors such as hormones that differ between the sexes, so it is likely important to understand the variety of genes and biological factors involved, as well as to consider their interactions with the environment in people with ADHD. In addition to the biological sex differences that can create experiences that interact differentially with genetic factors, the socialization experiences that differ across sex and gender identities may have profound effects that also relate to genetic variation (Farhane-Medina et al., 2022). This is an area of interest for research; research using model systems has demonstrated that gene expression in the brain differs notably between the sexes even before the differentiation of the gonads (Farhane-Medina et al., 2022). This implies that from early on in growth and development, human males and females neurobiologically contrast.

Looking at genetics specifically in females is important as there is evidence that hormone-related factors may relate to ADHD, therefore genetics associated with biological differences in these hormone systems or systems that interact with them may have distinct expressions between males and females. A study examining the prevalence of symptoms of premenstrual dysphoric disorder (PMDD), episodes of postpartum depression symptoms (PPD) after the first-time giving birth, and climacteric mood symptoms in ADHD in biological females found that the fluctuating hormone levels throughout the menstrual cycle may also impact severity of ADHD symptoms (Dorani et al., 2020). This study is unique, focusing specifically on processes exclusive to biological females that may have significant interactions with ADHD diagnosis.

This literature review will highlight what studies have been conducted that look at genetic ADHD exclusively in women or females. The mechanism for genetic inheritance is the same for males and females, so aside from some differences in X-linked variants, polymorphisms have comparable prevalence. However, the impact of these genetic variants may differ based on sex or gender because of hormone levels, socialization, or experiences, thus it is important for genetics studies to make sure they are appropriately representing women and/or females. Because one of the purposes of this review is to evaluate how female-focused ADHD genetics researchers approach the studies, we did not include studies that investigate males and females simultaneously.

The purpose of this review was to highlight what studies have been conducted that investigate ADHD genetics studies exclusively in women or females. However, upon conducting the review, there is not currently sufficient published evidence for molecular genetics studies exclusively in this population. Therefore, the studies reviewed in this paper include a variety of heritability studies and genetic studies of other disorders that included ADHD comorbidities. The evidence discussed in this review emphasizes the current lack of molecular genetic studies focused on women or females.

In order to frame the amount of research, type of research, and specific findings applicable to the genetics of ADHD on female populations, we also include a review of the research on exclusively male populations to compare. Although the best way to understand any sex differences that exist in genetic contributions to ADHD is with studies including both males and females, evaluating the work that is done on the sexes separately may reveal important biases in research methodology or framing of results.

Methods

Search for Genetics of ADHD in Females

Using the PubMed database, we conducted a search on February 9th, 2023 using the terms “ADHD AND gene* AND (women OR female*)”. This yielded 7,828 results, but a majority of these articles were not focused solely on ADHD in female populations. Articles that included both male and female participants and articles that were outside the scope of the review were excluded. Inclusion criteria required that the articles included a measure of genetic variation, included an exclusively biological female population, and that these represented participants received ADHD diagnoses at a point in their lives. The variability of methods across the studies that meet the search criteria (only studies with genetics measure, ADHD diagnosis, and exclusively a female/woman population) will be discussed in the results.

Search for Genetics of ADHD in Males

For the sake of contextualizing our results of female-specific studies, we repeated the search investigating male-specific studies. Thus, using the PubMed database, we conducted a search on March 2nd, 2023 using the terms “ADHD AND gene* AND (men OR male*)”. This yielded 8,519 results, but a majority of these articles were not focused solely on ADHD in male populations either. Articles that included both male and female participants and articles that were outside the scope of the review were excluded. Inclusion criteria required that the articles included a measure of genetic

variation, included an exclusively male population, and that these represented participants received ADHD diagnoses at a point in their lives. The variability of methods across the studies that meet the search criteria (only studies with genetics measure, ADHD diagnosis, and exclusively a male/man population) will be discussed in the results and compared with the female-focused search results that are the primary focus of this review.

Results

Methodological Variation Across Female-Focused Studies

Using our primary set of search terms, six studies fit the inclusion criteria, which is that the studies must involve genetics, participants in the study must have ADHD diagnosis, and feature exclusively a female/woman population. Of these six studies, four pertain to heritability and familial risk factors (Biederman et al., 2009, Faraone et al., 2000, Faraone et al., 2001, Todd et al., 2001) and two test specific genetic variants comorbid with other psychiatric disorders (Levitan et al., 2002, Yilmaz et al., 2012).

Table 1 illustrates the sample sizes, ADHD measures/diagnoses, genetic variables, and races of participants across the six studies found that meet the inclusion criteria. All of the participants had to meet the full Diagnostic and Statistical Manual of Mental Disorders (DSM) criteria for ADHD and/or another mental disorder (Biederman et al., 2009, Faraone et al., 2000, Faraone et al., 2001, Levitan et al., 2002, Todd et al., 2001, Yilmaz et al., 2012). Most of the studies contained either all White participants (Yilmaz et al., 2012) or a majority of the participants were White (Biederman et al., 2009, Faraone et al., 2000, Todd et al., 2001). Sample sizes of participants ranged from 66 female participants (Levitan et al., 2002) to 4,036 female participants (Todd et al., 2001) and the types of genetic variable measures across these studies were familial and heritability studies (Biederman et al., 2009, Faraone et al., 2000, Faraone et al., 2001, Todd et al., 2001) as well as investigating specific genes that play a role in ADHD such as HTR2A and DRD4 (Levitan et al., 2002, Yilmaz et al., 2012).

Results Across the Studies

In a report assessing the familial transmission of ADHD in families found through girls, interviews were given to 140 girls with ADHD and their 417 first-degree relatives and to 122 girls without ADHD and their 369 first-degree relatives. The increased risk for ADHD among the relatives of girls with ADHD was true for the inattentive and combined subtypes but not for the hyperactive-impulsive subtype, which is more often seen in boys (Faraone et al., 2000). The relatives of the girls with ADHD had significantly higher rates of mood, antisocial, substance abuse, and anxiety disorders than the relatives of the girls without ADHD (Faraone et al., 2000). Performing genetic testing on relatives of girls with risk of ADHD prior to school-age could help families and individuals understand their diagnoses as well as seek any accommodations to prevent suffering and struggling at home and in school. With family members showing genetic variation of potential risk factors of ADHD, preparing for potential phenotypic outcomes could help at-risk girls with school life as well as life at home.

A study investigating whether variation in the 5-HT2A receptor gene (HTR2A) is associated with symptoms of ADHD in childhood in adult women with SAD found that in females with SAD, the variation of the 5-HT2A receptor gene plays a role in a biological process that contributes to impulsivity and core ADHD symptoms in childhood, followed by seasonal mood change in adulthood (Levitan et al., 2002).

Another study investigating the association between genetic variants and ADHD in women detected an association of the DRD4 gene with a record of attention-deficit/hyperactivity disorder in women with bulimia nervosa (BN) (Yilmaz et al., 2012). This study assessed transmission of DRD4 alleles in families of patients with BN and explored the potential impact of the DRD4 gene in childhood ADHD history in women with BN. Results showed that the BN probands and controls were not different in DRD4 allele frequency, and there was no evidence found for “preferential transmission” of any variants of DRD4 to the impacted offspring with BN. The variants of DRD4 are associated with a history of childhood ADHD in BN probands; this insight may be helpful in the treatment of BN (Yilmaz et al., 2012).

A study examining 140 girls with ADHD, 122 non-ADHD comparisons and their 786 first degree relatives aimed to test hypotheses about the familial relationship between ADHD and bipolar disorder in girls (Faraone et al., 2001). It found that relatives of both ADHD subgroups were at much higher risk for ADHD than the relatives of non-ADHD controls. A heightened risk for bipolar disorder was also found among relatives when the proband child had BPD but not ADHD by itself. There was also inadequate evidence of a discrimination between ADHD and BPD (Faraone et al., 2001).

The familiality and heritability of ADHD subtypes in a population sample of adolescent female twins was examined and latent-class analysis was used on data from parents on the DSM-IV ADHD symptoms in female twins from ages 13-23 (Todd et al., 2001). It was

Table 1
Methodology Across Studies

Citation	Sample Size	ADHD Measure/Diagnosis	Genetic Variable	Race of Participants
Biederman et al., 2009	123 probands, 403 relatives and 112 probands, 359 relatives	Met DSM-III-R diagnostic criteria for ADHD at the time of the clinical referral	Familial risk analysis	"Since the sample consisted of largely Caucasian* subjects [96% of the sample], our findings may not generalize to other minority or ethnic groups."
Faraone et al., 2000	140 girls with ADHD and their 417 first-degree relatives and to 122 girls without ADHD and their 369 first-degree relatives	Met the full DSM-III-R diagnostic criteria for ADHD at the time of the clinical referral	Family study	95% Caucasian*, 3% African American, 2% Other
Faraone et al., 2001	140 ADHD probands and 122 non-ADHD comparisons. These groups had 417 and 369 first-degree biological relatives, respectively, who provided data	Met full DSM-III-R diagnostic criteria for ADHD at the time of the clinical referral	Familiality and heritability of subtypes of ADHD	Does not say
Levitan et al., 2002	66 women	DSM-III-R for major depression/winter seasonal pattern, Wender-Utah Rating Scale for ADHD	Genomic DNA extracted, single nucleotide polymorphism (SNP) in *HTR2A* gene	Does not say
Todd et al., 2001	4,036 adolescent female twins in 1,127 monozygotic pairs and 891 dizygotic pairs	DSM-IV diagnoses were derived by computer algorithm, latent class analysis	Familiality and heritability of subtypes of ADHD	"African Americans constituted 13.3% of the total sample. The majority of the sample was European American; less than 1% were Hispanic, Asian American, Native American, or of other self-identified ancestries."
Yilmaz et al., 2012	243 women of European Caucasian descent with current or past BN purging subtype	Current or past BN purging subtype, Structural Clinical Interview for DSM-IV for Axis I Disorders, Eating Disorder Examination, Wender Utah Rating Scale	Blood lymphocyte DNA was extracted, polymerase chain reaction (PCR), looking at *DRD4* gene	All participants were European Caucasian*

Note. *Although this language is recognized as outdated and problematic by many, we include it here for accuracy in the self-identified races of the participants in the studies.

found that the primarily inattentive and combined subtypes of ADHD co-clustered within families, but the primarily hyperactive/impulsive subtype and the individual latent-class analysis subtypes did not co-cluster (Todd et al., 2001). This demonstrates that unlike the subtypes of ADHD as shown in the DSM-IV, ADHD subtypes from latent class analysis show to be independently transmitted in families. Although heritability via twin studies is typically not sex-dependent unless the trait is X-linked, heritability rates can differ between genders likely due to genetic factors interacting with socialized environmental factors that differ between genders. This can lead to amplification of genetic contribution (Ludeke et al., 2013).

In female adolescents, familial risk analysis of the association between ADHD and psychoactive substance abuse disorder (PSUD) was studied in a controlled clinical trial, as literature findings at the time could not generalize to females (Biederman et al., 2009). It found that ADHD in the proband highly increased ADHD risk in relatives independently of the comorbidity with PSUD in relatives with ADHD or not. Trends in analysis of familial risk propose that these disorders are transmitted separately (Biederman et al., 2009).

Methodological Variation Across the Male-Focused Studies

As only six articles met the inclusion criteria for the primary female-focused purpose of this review, it made sense to do another search that is exclusive to genetics studies with only male participants with ADHD diagnosis to see how many articles would meet this inclusion criteria. Only three articles met the inclusion criteria for the male-only search (Bellgrove et al., 2005, Schulz et al., 1998, Torrioli et al., 2010). Table 2 illustrates the sample sizes, ADHD measures/diagnoses, genetic variables, and races of participants across the six studies found that meet the inclusion criteria. All of the participants had to meet the full Diagnostic and Statistical Manual of Mental Disorders (DSM) for ADHD and/or another mental disorder (Bellgrove et al., 2005, Schulz et al., 1998, Torrioli et al., 2010) although the diagnostic criteria may vary across editions. Sample sizes of participants ranged from 10 male participants (Torrioli et al., 2010) to 40 male participants (Schulz et al., 1998) and the types of genetic variable measures across these studies were genetic association studies (Bellgrove et al., 2005, Schulz et al., 1998, Torrioli et al., 2010) as well as investigating specific genes/alleles that play a role in ADHD such as 5-HT and DAT1 (Bellgrove et al., 2005, Schulz et al., 1998).

Results Across the Male-Specific Studies

One of the studies found that sons of alcoholic fathers have significantly greater cortisol responses to fenfluramine (FEN) challenge procedure relative to the sons of nonalcoholic fathers Schulz et al., 1998). This discrepancy may be a result of the different mechanisms that underlie their 5-HT stimulated release, suggesting that in boys with ADHD, those at familial risk for alcoholism may be different from those who are not at risk in 5-HT function (Schulz et al., 1998). It was also discovered in another study that treatment with valproic acid improved ADHD symptoms in fragile X syndrome boys (Torrioli et al., 2010).

The final article found that the high-risk dopamine transporter genotype (DAT1) ADHD group (meaning they have a 10-repeat DAT1 allele) displayed greater response variability on the sustained attention to response task (SART) than either the low-risk DAT1 group or healthy controls, whereas the latter two groups did not differ (Bellgrove et al., 2005). The high-risk DAT1 group showed weakened spatial asymmetry, relative to the low-risk DAT1 ADHD group. The 10-repeat DAT1 allele may mediate neuropsychological impairment in ADHD (Bellgrove et al., 2005).

Differences in Female-Specific vs. Male-Specific Search Results

There were a limited number of studies investigating either female-specific or male-specific associations between ADHD and genetics. This is likely in part to a tendency for studies to include both sexes and stratify by sex if they are interested in sex-based differences. There was a wide variety of different valuable results across all of the studies, but the variability in methods and focuses of the studies was striking. The types of comorbidities that researchers focus on in the studies differ for female-specific studies compared to male-specific studies. The female-specific studies included comorbidities such as BPD (Faraone et al., 2001), SAD (Levitan et al., 2002) and BN (Yilmaz et al. 2012), which are both diagnosed at higher rates in women and may be attached to gender stereotypes in women, whereas male-specific studies included a focus on alcoholism (Torrioli et al., 2010), which is diagnosed at higher rates in men and may be associated with gender stereotypes in men.

Discussion

The studies yielded from both searches may have similar or different methodological problems based on which sex or gender the study focuses on. As much research exists as comparison studies and that is likely the best way to understand differences between male and

Table 2
Methodology Across Studies

Citation	Sample Size	ADHD Measure/Diagnosis	Genetic Variable	Race of Participants
Bellgrove et al., 2005	Twenty two right-handed ADHD participants	DSM-IV criteria for ADHD, WISC-III	Genetic association study of low-risk and high-risk DAT1 allele	Does not say
Schulz et al., 1998	40 7-to 11-year-old prepubertal boys who met DSM-III-R criteria for ADHD	DSM-III-R criteria for ADHD, Wechsler Intelligence Scale for Children–Revised (WISC-R), Wide-Range Achievement Test–Revised (WRAT-R), Family History-Research Diagnostic Criteria (FH-RDC)	Relationship of 5-HT function to aggression	30% (n = 12) Caucasians*, 25% (n = 10) African-Americans, and 45% (n = 18) Hispanics.
Torrioli et al., 2010	10 male patients aged between 7 and 16 years old	FXS diagnosis, diagnosis of ADHD according to the DSM-IV, WISC-R, Conner's Teacher Rating Scale, VABS—Survey Form, CGI-S ADHD, SNAP-IV—Parent Rating Scale, K-SADS PL	Fragile X syndrome is due to expansion and methylation of the CGG sequence at the 5′ UTR of the FMR1 gene	All Caucasian participants*

Note. *Although this language is recognized as outdated and problematic by many, we include it here for accuracy in the self-identified races of the participants in the studies.

female specifics, by focusing on only one sex at a time, we can discuss some of the factors associated with sex or gender that can be contributing to how ADHD is perceived in males and females. The genetic variants identified may have different impacts across sexes due to biological and social factors, specifically, relatives of girls with ADHD are at increased risk for the inattentive and combined subtypes but not for the hyperactive-impulsive subtype, which is more common in boys (Faraone et al., 2000), and also had higher rates of mood, antisocial, substance abuse, and anxiety disorders than the relatives of girls without ADHD (Faraone et al., 2000). Gene-association studies found that in females with seasonal affective disorder, variation in the 5-HT2A receptor gene affects biological processes that lead to impulsivity and other ADHD symptoms in childhood, SAD following into adulthood, and the variant of the DRD4 gene are linked to a history of childhood ADHD in probands with BN (Yilmaz et al., 2012). There was also inadequate evidence of a discrimination between ADHD and bipolar disorder (Faraone et al., 2001).

The main findings from the studies reviewed confirm the heritability of ADHD and some genetic variants associated with symptom profiles in females with ADHD. This knowledge may help in understanding how the behavioral disorder presents in females. There are certainly distinctions between the approaches in female-specific studies versus in male-specific studies. Three of the ADHD genetics studies from this search term studied females with ADHD with a comorbid disorder (Faraone et al., 2001, Levitan et al., 2002, Yilmaz et al., 2012,). These comorbidities happen to be more prevalent in this particular gender. The same general trend showed up for the studies focused on exclusively males; they chose comorbidities that are more common in men to focus on (e.g. alcoholism; Torrioli et al., 2010).

The gender binary in this review is the standard for most research on ADHD, however, sex and gender are separate constructs, and neither are strictly binary. Sex-based biological variables may be more adequately investigated based on hormone levels or other intermediate phenotypes (Martin, 2024). In terms of

gender identity, the binary completely misses the complexity of the experiences nonbinary people have. Understanding the genetics of ADHD is more complex than boys compared to girls, and this is not emphasized in the existing research. This is a new level of investigation that may help better understand ADHD as a behavioral disorder that affects all kinds of people in different ways. There is some recent evidence investigating ADHD in nonbinary folk (Cheng et al., 2022, Conway et al., 2022, Goetz 2022, Salciunas 2021), but these studies were beyond the scope of the current review; however, as more nuanced and diverse measurements of sex and gender hopefully become more prevalent in the literature, future reviews can disentangle these factors and represent a wider portion of the human population. As nonbinary and gender nonconforming people with mental illness likely face additional adversity due to their gender fluidity, it is pertinent to investigate ADHD and other mental disorders in nonbinary people.

Our inability to find any studies that directly addressed our initial focus on molecular studies of ADHD in women or females highlights the complete lack of studies of this type. We urge future researchers to fill in this gap by focusing on understanding ADHD including the crucial molecular genetic components in specifically female populations of a wide and inclusive demographic.

Importance of Studying ADHD Specifically in Females

Studying females in isolation from males allows more focus on the female bodily, mental, and endocrine changes which differ between the sexes in ways that may be critical to understand with respect to ADHD (Eng et al., 2024). There is evidence that there are sex differences in the severity and presentation of ADHD symptoms, conduct problems, and learning problems in males and females with and without clinically recognized ADHD (Mowlem et al., 2019). In females, hyperactivity/impulsivity and conduct problems were stronger predictors of clinical diagnosis and prescription of pharmacological treatment; females with ADHD may be more easily discarded in the process of ADHD diagnosis and less likely to be administered medication unless they have prominent externalizing issues (Mowlem et al., 2019). From a consecutive cohort of Child and Adolescent Psychiatric (CAP) outpatients, emotional symptoms were more common as a cause for referral to CAP among girls; for boys, prevalence of neurodevelopmental disorders was more common for reasons for referral (Klefsjo et al., 2020).

This is consistent with results from a study comparing women with and without ADHD. Results showed that women with ADHD had three times the presence of insomnia, chronic pain, ideation of suicide, sexual abuse in their childhood, and generalized anxiety disorder as women without ADHD. Women with ADHD also have two times the prevalence of substance abuse, smoking habit, depressive issues, poverty, and childhood physical abuse compared to women without ADHD (Fuller-Thompson et al., 2016). Intensity of ADHD symptoms decreases in both genders as age increases, but women with childhood ADHD are more likely than males with ADHD to have symptoms later into adulthood (Fuller-Thompson et al., 2016). It is possible that due to the differences in reproductive organs and associated hormonal pathways, ADHD is affecting males and females differently in this regard, as the endocrine system plays a large role in mood patterns of females, specifically estrogen (Roberts et al., 2018).

Changes in estrogen can affect the intensity and presentation of ADHD symptoms at different times during lifespan, such as puberty, monthly menstrual cycle, perimenopause, and menopause. A more female-centered approach to the process of ADHD diagnosis may be the key to saving many from receiving inadequate or late diagnosis. To do so, physicians may see adopting a lifespan model of care as beneficial, to model the complex transitions and changes females (and males) undergo developmentally as clinical presentations and social circumstances also change (Young et al., 2020).

Research on pharmaceutical or therapeutic intervention options for ADHD is currently working on considering variation of the 5-HT2A receptor gene, among other therapeutic targets, as it plays a role in a biological process that contributes to impulsivity and core ADHD symptoms in childhood (Nazarova et al., 2022). 5-HT2A receptor gene drug targets may help in easing symptoms of ADHD. Variants of DRD4 are associated with several neuropsychiatric diseases such as ADHD, schizophrenia, substance abuse, mood disorders, and eating disorders such as BN (Botticelli et al., 2020), so looking more into the findings from these studies may provide researchers with important neuronal and hormonal pathways that affect ADHD in all people, not specific to a particular race or gender binary.

The approach to the current review was valuable, including providing insight into research approaches in ADHD genetic studies focusing on just one gender or sex, but the tradeoff was not including many of the excellent studies that compare gender differences by including both males and females. The limitations of this literature review include the fact that since this review only included articles exclusively with female participants, some key evidence may have been missed regarding the genetics of ADHD in females. Studies that focus on sex

differences by comparing results in both males and females are likely to be even more informative about the unique qualities of ADHD in females than those that study females alone. Many studies include both boys and girls but might do an analytic strategy that looks at association separately, thus allowing for further clarification about differences that may occur between the sexes.

The limitations of the studies in this literature review are that in the majority of the studies cited, the participants included in the studies using the specific search term were of European descent (Biederman et al., 2009, Faraone et al., 2000, Todd et al., 2001, Yilmaz et al., 2012) unless the authors did not describe racial or ethnic demographics (Faraone et al., 2001, Levitan et al., 2002). This means that conclusions about ADHD from these types of studies are still limited and narrow, as these studies include samples that do not adequately represent every person affected by ADHD and therefore cannot be generalized to all females with ADHD. Another limitation is that many of the studies found through the search term are from at least a decade ago, so some of the findings may not be as up to date with current knowledge.

Future research should represent people with ADHD who are from a variety of ethnic backgrounds, genders, races, and socioeconomic experience so that the conclusions drawn from the studies explored in this literature review can be generalized to non-White gender-varying populations. The way researchers focus on ADHD in females in these studies (and the way the researchers focused on ADHD in males) reveals how attitudes about gender shape research and limit our understanding of disorders like ADHD. It may be useful in a future review to include all studies that include gender differences, to provide a broader perspective for how researchers examine sex and gender differences in ADHD genetics. Results from the female-only studies draw suggestions for how we should do genetics studies of ADHD in females to be inclusive of females who have ADHD, taking into account the processes that are specific to biological females, such as menstruation, gestation, and menopause, as well as life experiences that are impacted by their societies and environments.

References

Attoe, D. E., & Climie, E. A. (2023). Miss. Diagnosis: A Systematic Review of ADHD in Adult Women. *Journal of attention disorders*, *27*(7), 645–657. https://doi.org/10.1177/10870547231161533

Babinski, D. E., Pelham, W. E., Molina, B. S., Waschbusch, D. A., Gnagy, E. M., Yu, J., Sibley, M. H., & Biswas, A. (2011). Women with childhood ADHD: Comparisons by Diagnostic Group and gender. *Journal of Psychopathology and Behavioral Assessment*, *33*(4), 420-429. https://doi.org/10.1007/s10862-011-9247-4

Bralten, J., Franke, B., Waldman, I., Rommelse, N., Hartman, C., Asherson, P., Banaschewski, T., Ebstein, R. P., Gill, M., Miranda, A., Oades, R. D., Roeyers, H., Rothenberger, A., Sergeant, J. A., Oosterlaan, J., Sonuga-Barke, E., Steinhausen, H.-C., Faraone, S. V., Buitelaar, J. K., & Arias-Vásquez, A. (2013). Candidate genetic pathways for attention-deficit/hyperactivity disorder (ADHD) show association to hyperactive/impulsive symptoms in children with ADHD. *Journal of the American Academy of Child & Adolescent Psychiatry*, *52*(11), 1204-1212. https://doi.org/10.1016/j.jaac.2013.08.020

Bellgrove, M. A., Hawi, Z., Kirley, A., Gill, M., & Robertson, I. H. (2005). Dissecting the attention deficit hyperactivity disorder (ADHD) phenotype: Sustained attention, response variability and spatial attentional asymmetries in relation to dopamine transporter (DAT1) genotype. *Neuropsychologia*, *43*(13), 1847–1857. https://doi.org/10.1016/j.neuropsychologia.2005.03.011

Biederman, J., Petty, C. R., Monuteaux, M. C., Mick, E., Clarke, A., Ten Haagen, K., & Faraone, S. V. (2009). Familial risk analysis of the association between attention-deficit/hyperactivity disorder and psychoactive substance use disorder in female adolescents: A controlled study. *Journal of Child Psychology and Psychiatry*, *50*(3), 352–358. https://doi.org/10.1111/j.1469-7610.2008.02040.x

Botticelli, L., Micioni Di Bonaventura, E., Del Bello, F., Giorgioni, G., Piergentili, A., Romano, A., Quaglia, W., Cifani, C., & Micioni Di Bonaventura, M. V. (2020). Underlying Susceptibility to Eating Disorders and Drug Abuse: Genetic and Pharmacological Aspects of Dopamine D4 Receptors. *Nutrients*, *12*(8), 2288. https://doi.org/10.3390/nu12082288

Brikell, I., Burton, C., Mota, N. R., & Martin, J. (2021). Insights into attention-deficit/hyperactivity disorder from recent genetic studies. *Psychological Medicine*, *51*(13), 2274–2286. https://doi.org/10.1017/s0033291721000982

Cabral, M. D. I., Liu, S., & Soares, N. (2020). Attention-deficit/hyperactivity disorder: diagnostic criteria, epidemiology, risk factors and evaluation in youth. *Translational Pediatrics*, *9*(Suppl 1), S104–S113. https://doi.org/10.21037/tp.2019.09.08

Castle, L., Aubert, R. E., Verbrugge, R. R., Khalid, M., & Epstein, R. S. (2007). Trends in medication treatment for ADHD. *Journal of Attention Disorders*, *10*(4), 335–342. https://doi.org/10.1177/1087054707299597

Cheng, A., Illango, A., El-Kadi, A., Ma, H., Mitchell, N., Rajapakse, N., & Robb, C. (2022). Factors Impacting Gender Diagnostic Differences in ADHD: A Review. *The Child Health Interdisciplinary Literature and Discovery Journal*, *1*(1), 25-29. https://doi.org/10.15173/child.v1i1.3126

Conway, F., Ensink, K., Farsang, M., Lyon, S., & Burger-Calderon, M. (2022). The ADHD compassion project case consultation with Dr. Karin Ensink: Using reflective functioning in the treatment of a gender non-binary ADHD child. *Journal of Infant, Child, and Adolescent Psychotherapy*, *21*(2), 208-216. https://doi.org/10.1080/15289168.2022.2045465

Corominas, J., Klein, M., Zayats, T., Rivero, O., Ziegler, G. C., Pauper, M., Neveling, K., Poelmans, G., Jansch, C., Svirin, E., Geissler, J., Weber, H., Reif, A., Arias Vasquez, A., Galesloot, T. E., Kiemeney, L. A., Buitelaar, J. K., Ramos-Quiroga, J.-A., Cormand, B., Ribasés, M., Hveem, K., Gabrielsen, M. E., Hoffmann, P., Cichon, S., Haavik, J., Johansoon, S., Jacob, C. P., Romanos, M., Franke, B., & Lesch, K.-P. (2020). Identification of ADHD risk genes in extended pedigrees by combining linkage analysis and whole-exome sequencing. *Molecular Psychiatry*, *25*(9), 2047–2057. https://doi.org/10.1038/s41380-018-0210-6

Davies, W. (2014). Sex differences in attention deficit hyperactivity disorder: Candidate genetic and endocrine mechanisms. *Frontiers in Neuroendocrinology*, *35*(3), 331–346. https://doi.org/10.1016/j.yfrne.2014.03.003

Derks, E. M., Dolan, C. V., Hudziak, J. J., Neale, M. C., & Boomsma, D. I. (2007). Assessment and etiology of attention deficit hyperactivity disorder and oppositional defiant disorder in boys and girls. *Behavior Genetics*, *37*(4), 559–566. https://doi.org/10.1007/s10519-007-9153-4

Derks, E. M., Hudziak, J. J., & Boomsma, D. I. (2007). Why more boys than girls with ADHD receive treatment: A study of Dutch twins. *Twin Research and Human Genetics*, *10*(5), 765–770. https://doi.org/10.1375/twin.10.5.765

Dorani, F., Bijlenga, D., Beekman, A. T. F., van Someren, E. J. W., & Kooij, J. J. S. (2021). Prevalence of hormone-related mood disorder symptoms in women with ADHD. *Journal of Psychiatric Research*, *133*, 10–15. https://doi.org/10.1016/j.jpsychires.2020.12.005

Eng, A. G., Nirjar, U., Elkins, A. R., Sizemore, Y. J., Monticello, K. N., Petersen, M. K., Miller, S. A., Barone, J., Eisenlohr-Moul, T. A., Martel, M. M. (2024). Attention-deficit/hyperactivity disorder and the menstrual cycle: Theory and evidence. *Hormones and Behavior*, *158*, 105446. https://doi.org/10.1016/s0006-3223(01)01146-5

Faraone, S. V., Biederman, J., Mick, E., Doyle, A. E., Wilens, T., Spencer, T., Frazier, E., & Mullen, K. (2001). A family study of psychiatric comorbidity in girls and boys with attention-deficit/hyperactivity disorder. *Biological Psychiatry*, *50*(8), 586–592. https://doi.org/10.1016/s0006-3223(01)01146-5

Faraone, S. V., Biederman, J., Mick, E., Williamson, S., Wilens, T., Spencer, T., Weber, W., Jetton, J., Kraus, I., Pert, J., & Zallen, B. (2000). Family study of girls with attention deficit hyperactivity disorder. *American Journal of Psychiatry*, *157*(7), 1077–1083. https://doi.org/10.1176/appi.ajp.157.7.1077

Faraone, S. V., Biederman, J., & Monuteaux, M. C. (2001). Attention deficit hyperactivity disorder with bipolar disorder in girls: Further evidence for a familial subtype? *Journal of Affective Disorders*, *64*(1), 19–26. https://doi.org/10.1016/s0165-0327(00)00213-5

Farhane-Medina, N.Z., Luque, B., Tabernero, C., & Castillo-Mayén, R. (2022). Factors associated with gender and sex differences in anxiety prevalence and comorbidity: A systematic review. *Science Progress*, *105*(4), 1-30. doi:10.1177/00368504221135469

Francis, K., Karantanos, G., Al-Ozairi, A., & AlKhadhari, S. (2021). Prevention in Autism Spectrum Disorder: A Lifelong Focused Approach. *Brain sciences*, *11*(2), 151. https://doi.org/10.3390/brainsci11020151

Fuller-Thomson, E., Lewis, D. A., & Agbeyaka, S. K. (2016). Attention-deficit/hyperactivity disorder casts a long shadow: Findings from a population-based study of adult women with self-reported ADHD. *Child: Care, Health and Development*, *42*(6), 918–927. https://doi.org/10.1111/cch.12380

Goetz, T. G., & Adams, N. (2022). The transgender and gender diverse and attention deficit hyperactivity disorder nexus: A systematic review. *Journal of Gay & Lesbian Mental Health*, *28*(1), 2-19. https://doi.org/10.1080/19359705.2022.2109119

Greene, Ross W., Biederman, Joseph, Faraone, Stephen. V., Monuteaux, Michael C., Mick, Erick, DuPre, Emily P., Fine, Catherine S., & Goring, Jennifer C. (2001). Social impairment in girls with ADHD: Patterns, gender comparisons, and correlates. *Journal of the American Academy of Child & Adolescent Psychiatry*, *40*(6), 704–710. https://doi.org/10.1097/00004583-200106000-00016

Hou, Y.-wei, Xiong, P., Gu, X., Huang, X., Wang, M., & Wu, J. (2018). Association of serotonin receptors with attention deficit hyperactivity disorder: A systematic review and meta-analysis. *Current*

Medical Science, *38*(3), 538–551. https://doi.org/10.1007/s11596-018-1912-3

Kanarik, M., Grimm, O., Mota, N. R., Reif, A., & Harro, J. (2022). ADHD co-morbidities: A review of implication of gene × environment effects with dopamine-related genes. *Neuroscience & Biobehavioral Reviews*, *139*, 104757. https://doi.org/10.1016/j.neubiorev.2022.104757

Kitsune, G. L., Kuntsi, J., Costello, H., Frangou, S., Hosang, G. M., McLoughlin, G., & Asherson, P. (2016). Delineating ADHD and bipolar disorder: A comparison of clinical profiles in adult women. *Journal of Affective Disorders*, *192*, 125–133. https://doi.org/10.1016/j.jad.2015.12.024

Klefsjö, U., Kantzer, A. K., Gillberg, C., & Billstedt, E. (2020). The road to diagnosis and treatment in girls and boys with ADHD – gender differences in the diagnostic process. *Nordic Journal of Psychiatry*, *75*(4), 301–305. https://doi.org/10.1080/08039488.2020.1850859

Levitan, R. D., Masellis, M., Basile, V. S., Lam, R. W., Jain, U., Kaplan, A. S., Kennedy, S. H., Siegel, G., Walker, M. L., Vaccarino, F. J., & Kennedy, J. L. (2002). Polymorphism of the serotonin-2a receptor gene (HTR2A) associated with childhood attention deficit hyperactivity disorder (ADHD) in adult women with seasonal affective disorder. *Journal of Affective Disorders*, *71*(1-3), 229–233. https://doi.org/10.1016/s0165-0327(01)00372-x

Ludeke, S., Johnson, W., McGue, M., & Iacono, W. G. (2013). Genetic amplification and the individualization of the parent-child relationship across adolescence. *Psychological Medicine*, *43*(2), 413–422. https://doi.org/10.1017/S0033291712001201

Mahone, E. M., & Denckla, M. B. (2017). Attention-deficit/hyperactivity disorder: A historical neuropsychological perspective. *Journal of the International Neuropsychological Society*, *23*(9-10), 916–929. https://doi.org/10.1017/s1355617717000807

Martin, J. (2024). Why are females less likely to be diagnosed with ADHD in childhood than males? *The Lancet Psychiatry*, *11*(4), 303-310. https://doi.org/10.1016/S2215-0366(24)00010-5

Mowlem, F. D., Rosenqvist, M. A., Martin, J., Lichtenstein, P., Asherson, P., & Larsson, H. (2019). Sex differences in predicting ADHD clinical diagnosis and pharmacological treatment. *European Child & Adolescent Psychiatry*, *28*(4), 481–489. https://doi.org/10.1007/s00787-018-1211-3

Mowlem, F., Agnew-Blais, J., Taylor, E., & Asherson, P. (2019). Do different factors influence whether girls versus boys meet ADHD diagnostic criteria? Sex differences among children with high ADHD symptoms. *Psychiatry Research*, *272*, 765–773. https://doi.org/10.1016/j.psychres.2018.12.128

Nazarova, V. A., Sokolov, A. V., Chubarev, V. N., Tarasov, V. V., & Schiöth, H. B. (2022). Treatment of ADHD: Drugs, psychological therapies, devices, complementary and alternative methods as well as the trends in clinical trials. *Frontiers in pharmacology*, *13*, 1066988. https://doi.org/10.3389/fphar.2022.1066988

Nussbaum, N. L. (2012). ADHD and Female Specific Concerns: A Review of the Literature and Clinical Implications. *Journal of Attention Disorders*, 16(2), 87–100. https://doi.org/10.1177/1087054711416909

Pinkhardt, E. H., Kassubek, J., Brummer, D., Koelch, M., Ludolph, A. C., Fegert, J. M., & Ludolph, A. G. (2009). Intensified testing for attention-deficit hyperactivity disorder (ADHD) in girls should reduce depression and smoking in adult females and the prevalence of ADHD in the longterm. *Medical Hypotheses*, *72*(4), 409–412. https://doi.org/10.1016/j.mehy.2008.11.025

Reimherr, F. W., Marchant, B. K., Gift, T. E., & Steans, T. A. (2017). ADHD and anxiety: Clinical significance and treatment implications. *Current Psychiatry Reports*, *19*(12), 109. https://doi.org/10.1007/s11920-017-0859-6

Rhee, S. H., Waldman, I. D., Hay, D. A., & Levy, F. (1999). Sex differences in genetic and environmental influences on DSM–III–R attention-deficit/hyperactivity disorder. *Journal of Abnormal Psychology*, *108*(1), 24–41. https://doi.org/10.1037/0021-843x.108.1.24

Ribasés, M., Ramos-Quiroga, J. A., Hervás, A., Bosch, R., Bielsa, A., Gastaminza, X., Artigas, J., Rodriguez-Ben, S., Estivill, X., Casas, M., Cormand, B., & Bayés, M. (2007). Exploration of 19 serotoninergic candidate genes in adults and children with attention-deficit/hyperactivity disorder identifies association for 5HT2A, DDC and Maob. *Molecular Psychiatry*, *14*(1), 71–85. https://doi.org/10.1038/sj.mp.4002100

Roberts, B., Eisenlohr-Moul, T., & Martel, M. M. (2018). Reproductive steroids and ADHD symptoms across the menstrual cycle. *Psychoneuroendocrinology*, *88*, 105–114. https://doi.org/10.1016/j.psyneuen.2017.11.015

Rommel, A.-S., Kitsune, G. L., Michelini, G., Hosang, G. M., Asherson, P., McLoughlin, G., Brandeis, D., & Kuntsi, J. (2016). Commonalities in EEG spectral power abnormalities between women with ADHD and women with bipolar disorder during rest and cognitive performance. *Brain Topography*, *29*(6), 856–866. https://doi.org/10.1007/s10548-016-0508-0

Salciunas, L. (2021). *Gender Differences in Cognitive Distortions in Adults with ADHD*. (Publication

No. 558) [Doctoral dissertation, Philadelphia College of Osteopathic Medicine Psychology. Digital Commons. https://digitalcommons.pcom.edu/psychology_dissertations/558

Schimmelmann, B. G., Hinney, A., Scherag, A., Pütter, C., Pechlivanis, S., Cichon, S., Jöckel, K. H., Schreiber, S., Wichmann, H. E., Albayrak, Ö., Dauvermann, M., Konrad, K., Wilhelm, C., Herpertz-Dahlmann, B., Lehmkuhl, G., Sinzig, J., Renner, T. J., Romanos, M., Warnke, A., Lesch, K. P., Reif, A., & Hebebrand, J. (2013). Bipolar disorder risk alleles in children with ADHD. *Journal of Neural Transmission*, *120*(11), 1611–1617. https://doi.org/10.1007/s00702-013-1035-8

Schote, A. B., Bonenberger, M., Pálmason, H., Seitz, C., Meyer, J., & Freitag, C. M. (2016). Glucocorticoid receptor variants in childhood attention-deficit/hyperactivity disorder and comorbid psychiatric disorders. *Psychiatry Research*, *246*, 275–283. https://doi.org/10.1016/j.psychres.2016.10.001

Schulz, M.A., K. (1998). Serotonin function and risk for alcoholism in boys with attention-deficit hyperactivity disorder. *Neuropsychopharmacology*, *18*(1), 10–17. https://doi.org/10.1016/s0893-133x(97)00091-2

Slobodin, O., & Davidovitch, M. (2019). Gender Differences in Objective and Subjective Measures of ADHD Among Clinic-Referred Children. *Frontiers in Human Neuroscience*, *13*(441), 1-14. https://doi.org/10.3389/fnhum.2019.00441

Thapar, A., Cooper, M., Eyre, O., & Langley, K. (2012). Practitioner review: What have we learnt about the causes of ADHD? *Journal of Child Psychology and Psychiatry*, *54*(1), 3–16. https://doi.org/10.1111/j.1469-7610.2012.02611.x

Thapar, A., Langley, K., Asherson, P., & Gill, M. (2007). Gene–environment interplay in attention-deficit hyperactivity disorder and the importance of a developmental perspective. *British Journal of Psychiatry*, *190*(1), 1–3. https://doi.org/10.1192/bjp.bp.106.027003

Thapar, A., Langley, K., Owen, M. J., & O'Donovan, M. C. (2007). Advances in genetic findings on attention deficit hyperactivity disorder. *Psychological Medicine*, *37*(12), 1681–1692. https://doi.org/10.1017/s0033291707000773

Todd, R. D., Rasmussen, E. R., Neuman, R. J., Reich, W., Hudziak, J. J., Bucholz, K. K., Madden, P. A. F., & Heath, A. (2001). Familiality and heritability of subtypes of attention deficit hyperactivity disorder in a population sample of adolescent female twins. *American Journal of Psychiatry*, *158*(11), 1891–1898. https://doi.org/10.1176/appi.ajp.158.11.1891

Torrioli, M., Vernacotola, S., Setini, C., Bevilacqua, F., Martinelli, D., Snape, M., Hutchison, J. A., Di Raimo, F. R., Tabolacci, E., & Neri, G. (2010). Treatment with valproic acid ameliorates ADHD symptoms in fragile X syndrome boys. *American Journal of Medical Genetics Part A*, *152A*(6), 1420-1427. https://doi.org/10.1002/ajmg.a.33484

Vainieri, I., Adamo, N., Michelini, G., Kitsune, V., Asherson, P., & Kuntsi, J. (2020). Attention regulation in women with ADHD and women with bipolar disorder: An ex-gaussian approach. *Psychiatry Research*, *285*, 112729. https://doi.org/10.1016/j.psychres.2019.112729

Wang, Y., Wang, T., Du, Y., Hu, D., Zhang, Y., Li, H., & Pei, W. (2021). Polygenic risk of genes involved in the catecholamine and serotonin pathways for ADHD in children. *Neuroscience Letters*, *760*, 136086. https://doi.org/10.1016/j.neulet.2021.136086

Wilens, T. E., Biederman, J., Faraone, S. V., Martelon, M. K., Westerberg, D., & Spencer, T. J. (2009). Presenting ADHD symptoms, subtypes, and comorbid disorders in clinically referred adults with ADHD. *The Journal of Clinical Psychiatry*, *70*(11), 1557–1562. https://doi.org/10.4088/jcp.08m04785pur

Yilmaz, Z., Kaplan, A. S., Levitan, R. D., Zai, C. C., & Kennedy, J. L. (2012). Possible association of the DRD4 gene with a history of attention-deficit/hyperactivity disorder in women with bulimia nervosa. *International Journal of Eating Disorders*, *45*(4), 622–625. https://doi.org/10.1002/eat.20986

Young, S., Adamo, N., Ásgeirsdóttir, B. B., Branney, P., Beckett, M., Colley, W., Cubbin, S., Deeley, Q., Farrag, E., Gudjonsson, G., Hill, P., Hollingdale, J., Kilic, O., Lloyd, T., Mason, P., Paliokosta, E., Perecherla, S., Sedgwick, J., Skirrow, C., Tierney, K., van Rensburg, K., & Woodhouse, E. (2020). Females with ADHD: An expert consensus statement taking a lifespan approach providing guidance for the identification and treatment of attention-deficit/ hyperactivity disorder in girls and women. *BMC Psychiatry*, *20*(1), 404. https://doi.org/10.1186/s12888-020-02707-9

Journal of Psychological Inquiry
2025, Vol. 29, No. 2, pp. 83-89

A REVIEW OF RECENT QUALITATIVE RESEARCH ON ASEXUALITY

STEPHANIE REDMAN AND WIND GOODFRIEND

BUENA VISTA UNIVERSITY

Abstract – While there is an abundance of research on human sexuality in general, the topic of asexuality is largely ignored. The first comprehensive review of empirical articles was not published until 2022 (Guz et al., 2022), and that paper only covered 48 articles produced between 2004 and mid-2021. The present article provides an update, focusing on nine qualitative articles on sexuality in the last few years. This review serves as a companion to Guz et al. (2022), noting (1) themes that have broadened our understanding of the diversity of asexual individuals, (2) a need for more research, and (3) a lack of education about asexuality in the general population. Implications are discussed, including that still more research is needed—especially in areas such as physical and mental health, relationships between asexual individuals, language and stigma related to asexuality, and intersectionality.

Keywords: asexuality, review, stigma, qualitative research

Asexuality is a sexual identity that does not receive much attention. It is commonly defined as a type of sexuality in which an individual has no sexual attraction toward others (Asexuality.org, 2024). It is difficult to estimate how common asexuality is in the general population, due to issues such as stigma and lack of education that will be discussed further in this paper. The most often cited statistic is that about 1% of people may be asexual—but it may be as high as 5.5% (Kaur, 2019; Lech et al., 2024). By any account, this is a small portion of the population, so it makes sense that there is less research available than for other communities. However, asexuality is still an independent identity with its own nuances that deserve attention.

There are relatively few research articles from the field of psychology regarding asexuality compared to other identities like homosexuality. Even though most researchers believe that asexuality has existed for a long time, most studies on this topic are relatively new. One important paper is Guz et al. (2022). This comprehensive paper discussed 48 articles considered to be all of the empirical asexual literature from 2004 to mid-2021. The authors found many themes regarding how samples of different studies were gathered and the definition of asexuality; we summarize it in the next section. The purpose of our paper is to add to what people know about asexuality by reviewing qualitative literature review on asexuality from 2021 to 2024 and analyzing whether knowledge on asexuality is growing in thematic ways.

Summary of Guz et al. (2022)

In Guz et al. (2022), the authors sought to provide the first systematic review of asexuality research. They included articles published from 2004 (when research on asexuality generally began) to mid-2021. They found and narrowed possible articles down to a list of 48. The authors completed online searches with ProQuest Dissertation, PsycINFO, Academic Search Complete, PsycARTICLES, and LGBT Life to find possible articles for their review, limiting their choices to only articles that were empirical, peer-reviewed, and focused on humans.

The authors came to two major conclusions—one regarding research methodology on asexuality, and one regarding the nature of asexuality itself. In terms of methodology, literature was both quantitative and qualitative. Demographics varied across studies, but most samples comprised cis, White women. Importantly, most researchers used a specific convenience sample of asexual people from the Asexual Visibility & Education Network (AVEN). AVEN is a place for the asexual community to gather online and have safe, honest conversations. (Asexuality.org, 2024). AVEN is a useful and popular website for the asexual community. However, when researchers use it as the only source of participants for their studies, it restricts diversity and representation of the larger population of interest. For example, AVEN is an online community, so people without good access to the internet or those who are not internet savvy will likely not be involved in studies recruiting from the AVEN

community. Likewise, those with schedules too busy to use AVEN for social media or who simply have not heard of the community are likely to be excluded from research samples.

Another key conclusion Guz et al. (2022) emphasized in their review was in regard to the nature of asexuality. Both researchers and participants seem to agree that asexuality means low or no sexual attraction. However, there is disagreement about whether asexuality should be further defined as (1) a theoretical continuum of sexual attraction—implying an accompanying complex identity and community, similar to other communities in the LGBTQIA+ world, or (2) the lack of a sexual identity. Either way, overall, Guz et al. (2022) noted that researchers focused on asexuality agree that it is not a disorder; over the past few decades, fewer scholars view asexuality from a pathological lens. Guz and her colleagues suggest further exploring the political and social lives of asexual people.

This Review

Guz et al. (2022) suggest that more research should be done on asexuality, to better understand its nuances and the needs of asexual people. The current paper focuses on qualitative papers published on asexuality since early 2021—essentially, an update since the Guz et al. (2022) paper was published. We exclusively reviewed qualitative studies because the nature of qualitative work promotes diverse ideas and perspectives (Stutterheim & Ratcliffe, 2021). Qualitative research allows for open-ended discussion with the participant talking freely, which permits them to use their own words to describe their experiences (Flavin, 2001). In other words, qualitative research is important as it connects readers directly to participants' voices without filtering them through any possible bias or the second-hand interpretation of the researchers (Charmaz, 2006; Stutterheim & Ratcliffe, 2021).

Qualitative research also prioritizes community involvement and engagement. Stutterheim and Ratcliffe (2021) point out that qualitative studies require researchers to better understand the language and terms the transgender community prefers. Using the proper language when referring to groups is important as it demonstrates respect for the community of interest. Additionally, qualitative research can enhance participants' connection to their community through learning about shared experiences (Opsal et al., 2016). Participants may feel not as alone after learning about what others in their community have experienced. This means that qualitative research benefits both researchers and individuals within the relevant community being investigated.

This literature review seeks to answer a few different questions:

1. What themes are in cutting-edge asexual research, and are there any important limitations in current understanding of the sexual community? This question will address both limitations noted by the researchers who have published the articles being reviewed and by the authors of the current study.

2. How has asexual literature and research changed, if at all, in the years since Guz et al.'s (2022) review article? This question will help provide a more thorough understanding of asexuality.

3. How does research on asexuality contribute to the community, and what further research is needed? This question will serve as a way to identify how expansive asexual literature is, whether the research helps or hurts asexual individuals themselves, and where future research should take the topic of asexuality.

Method

We started identifying articles relevant to this paper with an initial Google Scholar search to find qualitative articles related to asexuality. We chose to include any article for the current review that fit the following criteria: (a) focused on asexuality specifically, (b) in English, (c) in a peer-reviewed journal, and (d) empirically qualitative.

During our initial search for research on asexuality, we encountered the Guz et al. (2022) article, which thoroughly reviewed empirical research on asexuality from 2004 to mid-2021. After reading it, we realized we only needed to review work published after their important paper. However, we started by contacting them to request a list of the 18 articles they eliminated from their paper. We completed this step to see if we wanted to include any of the articles they excluded, just to be thorough. They provided 16 out of the 18 articles originally mentioned in their paper, and we agreed that they were justly excluded for various reasons (e.g., they were non-empirical or were not peer reviewed; see Table 1).

We then used Google Scholar and PsycINFO to identify additional articles published after mid-2021, when Guz et al. (2022) submitted their article for publication. We identified 21 additional articles from Google Scholar and 37 additional articles from PsycINFO for a total of 58 articles (again, see Table 1). We then started to eliminate articles for the reasons stated earlier. Elimination of articles was as follows: three were not in English; seven were dissertations; 33 were irrelevant; and 5 were review or non-empirical. That meant a total of 49 of the 58 articles were eliminated (plus the 16 from Guz et al., 2022), leaving nine articles to review. We read,

Table 1
Summary of Found and Eliminated Articles in This Review

	Article Source			
	Guz et al.	Google Scholar	PsycINFO	Total
Total identified	16	21	37	74
Reasons for elimination:				
Non-English	0	3	0	3
Dissertation	2	6	1	9
Irrelevant	2	3	30	35
Non-Empirical	12	5	1	18
Total eliminated	16	17	32	65
Remaining (used)	0	4	5	9

Note. Guz et al. (2022) eliminated several articles from their review; we double checked and agreed that they were either irrelevant to asexuality or non-empirical. Using Google Scholar and PsycInfo, we identified several additional articles for possible inclusion in our review, published since mid-2021. We eliminated all but nine for the reasons shown above.

summarized, and analyzed the nine remaining articles individually based on the three research questions presented above.

Results and Discussion

Themes and Limitations in Existing Asexuality Research

Three themes stood out from an examination of recent research, all of which indicate areas of limited understanding of the asexual community. One of the biggest themes is in regards to how to define asexuality. From the outside perspective, it may seem that the only distinction between asexual and allosexual individuals is the presence or absence of sexual attraction. However, there is more variety in people's attractions and experiences, reflecting diversity within the asexual community.

Recent research has highlighted subcategories of asexuality; two examples are demisexual and graysexual. Demisexual refers to people who only experience sexual attraction after forming an emotional connection or bond, while graysexual refers to people who experience sexual attraction some of the time (Boot-Haury & Cusick, 2023; Glass, 2022; Sheppard et al., 2024; Yang, 2023). The same studies, plus Brandley and Dehnert (2023), explain the split attraction model, where romantic and sexual attraction are separated. This means that someone can have romantic attraction, but not sexual attraction (asexual), or they could have sexual attraction, but not romantic attraction (known as "aromantic"). A person could also lack both kinds of attraction, which the community refers to as "aroace." This wide variety of subcategories shows that asexuality can be hard to define, encompasses a wide variety of individuals, and is more nuanced than it may first appear to outgroup members. Further, asexual individuals vary on how they feel about sex, with some being sex positive, others being sex negative, and others being sex neutral.

This variety can create different experiences for different asexual individuals. That said, they still appear more united than divided, as many people in the community feel a disconnect from allosexual individuals. The psychological schism can come in the form of not being understood by allosexual individuals, but also not understanding allosexual individuals. Many asexual individuals will flip the narrative from being the "weird" and misunderstood ones to labeling allosexuals in the same way (Brandley & Dehnert, 2023; Hart-Brinson et al., 2023; Peters, 2022). This attempt to regain agency helps mitigate some of the effects of discrimination that asexual individuals may face, as it allows for an asexual community.

A second major theme was the ubiquitous limitation regarding the recurring need for more research on asexuality. Supporting this idea was the fact that many researchers used the same articles for citations in their literature review, and that many studies—four out of nine—used AVEN as their source for participants (see Table 2, the column labeled "Sample Source"; Boot-Haury & Cusick, 2023; Brandley & Dehnert, 2023; Higginbottom, 2024; and Yang, 2024); this limitation is discussed further in the next section. In short, while gay, lesbian, bisexual, and even (to some extent) transgender people are getting increasing attention and acceptance in mainstream culture and media, the "A" in LGBTQIA+ initiatives is often still ignored.

Interestingly, Hart-Brinson et al. (2023) asked what the "A" should stand for in the LGBTQIA+ abbreviation. They found a general consensus among

Table 2

Articles Reviewed in this Paper, Source of Samples, and Topics

Author(s)	Article Title	Year	Sample Source	Topics
Boot-Haury & Cusick	Intersectional Asexual and Transgender and Gender Diverse Identity and Existential Concerns: A Thematic Analysis	2023	AVEN	Asexuality among gender diverse individuals
Brandley & Dehnert	"I am not a Robot, I am Asexual": A Qualitative Critique of Allonormative Discourses of Ace and Aro Folks as Robots, Aliens, Monsters	2023	AVEN & Reddit	Disconnect between asexual and allosexual individuals
Glass	Queering Relationships: Exploring Phenomena of Asexual Identified Persons in Relationships	2022	Facebook	Asexual relationships
Hart-Brinson, Tlachac, & Lepien	Contradictions in Experiences of Compulsory Sexuality and Pathways to Asexual Citizenship	2023	Personal contacts	"Becoming" an asexual citizen
Higginbottom	The Nuances of Intimacy: Asexual Perspectives and Experiences with Dating and Relationships	2024	Tumblr, Reddit, & AVEN	Asexual relationships
Mollet	"It's Easier Just to Say I'm Queer": Asexual College Students' Strategic Identity Management	2021	Campus wide email	Diversity of asexual individuals and intersectionality
Peters	Asexuality, Affect Aliens, and Digital Affect Cultures: Relationality with the happy Objects of Sexual and Romantic Relationships	2022	Instagram	Disconnect between asexual and allosexual individuals & asexual community
Sheppard, Mann, & Pfeffer	"Your Brain Isn't All Backwards": Asexual Young Women's Narratives of Sexual Healthism	2024	Facebook, Twitter, Instagram, & fliers	Disconnect between asexual and allosexual individuals & asexual community
Yang	Gender Uncoupled: Asexual People Making Sense of High School Sex Talk	2023	AVEN	The relationship between gender and asexuality and sex talk

Note. Articles are listed alphabetically by first author's last name.

their respondents that it should stand for asexuality, but some thought "ally" should be included too. If the latter were to occur, the asexual community would get even less attention than it does now. If the "A" refers to allies, this furthers the idea that asexual individuals do not get their own space and have to share it with others who may not even be in a sexual minority group themselves. It further pushes asexuality out of the focus, giving heterosexual allies a seat at the figurative table instead. However, neither Hart-Brinson and colleagues nor their participants posed that question. They also failed to bring up the idea of "agender," another valid identity under the LGBTQIA+ umbrella that starts with "A". It is hypocritical to talk about asexuality being excluded while simultaneously excluding another valid identity—agender—entirely.

A final point within this second theme is that some asexual participants described "gatekeeping" within the LGBTQIA+ community (Boot-Haury & Cusick, 2023; Brandley & Dehnert, 2023; Hart-Brinson et al., 2023; Mollet, 2021; Sheppard et al., 2024). Essentially, they perceived discrimination from people with other marginalized sexualities who may have implied that asexual individuals were not "queer" enough or just did

not fit in with the other identities within the LGBTQIA+ community. This perception led to many asexual individuals having mixed feelings about their connection to the LGBTQIA+ community.

A third theme within the articles was a lack of education about the asexual community from outsiders. Many asexual individuals recounted not sharing their identity because of the education they would be expected to bestow on others (Higginbottom, 2024; Mollet, 2021). For example, one asexual person explains, "As we often say in the asexual community, coming out is very demanding because it always feels like you also have to give an asexuality 101 class" (Foisy et al., 2022). When most people are not aware of what asexuality is, they feel the right to ask questions. Asexual people put in this position must then choose between accepting the burden of responsibility (i.e., educating those who are ignorant about the issue) and potentially answering deeply personal questions, or not answering and appearing rude. The general lack of education and conversation about asexuality also results in many asexual individuals not realizing who else is in their own community (i.e., they are not "coming out").

Lack of education also contributed to people not finding out about asexuality sooner. Many individuals explained that they found out about asexuality through social media. This was important as many people knew they felt different from allosexual individuals, but did not have the language and labels to describe themselves (Boot-Haury & Cusick, 2023; Glass, 2022; Hart Brinson et al., 2023; Peters, 2022; Sheppard et al., 2024). Tied to ignorance and misunderstanding was also the idea of discrimination. The discrimination people felt was mostly invalidation related to people not understanding asexuality. People make comments or judgments about asexuality without fully understanding what they were saying, which is hurtful (Boot-Haury & Cusick, 2023; Hart-Brinson et al., 2023; Higginbottom, 2024; Mollet, 2021; Peters, 2022; Sheppard et al., 2024).

How Asexuality Research Has Recently Changed

In the Guz et al. (2022) article, 48 articles from over 17 years were reviewed. About half of those were qualitative (a total of 22 – a little over one per year) and half were quantitative. In comparison, we found nine relevant qualitative articles published within the last 3 years. This suggests a rapid increase in recent qualitative research regarding human asexuality (from about one per year to three per year), which means more attention is being brought to the topic.

In the past, using AVEN as the primary source for finding a sample of asexual individuals was common among many studies on asexuality (Guz et al., 2022). This is true for more recent articles as well, although Reddit and other social media sites are also relatively common sources now. Using the same population across multiple studies is a problem because it means the data from one analysis to the next are no longer independent, an assumption for most statistical analyses (Dawson et al., 2017). Guz et al. (2022) found that 27 out of the 48 articles (about 56%) they analyzed used AVEN for recruiting. The present review found that only four out of the nine articles analyzed used AVEN, which is equal to 44%. Therefore, it seems that recent studies are using more sources to gather participants, potentially increasing the diversity and representativeness of asexual research samples.

Guz et al. (2022) also found that many articles suggested more research needs to be done, and on a wider variety of topics. Based solely on the number of articles on asexuality within the past three years, it seems that progress is being made. The last column of Table 2 shows the general topics of the research articles reviewed for the current analysis. Topics include: the diversity of asexuality, the disconnect between asexual and allosexual people, asexual relationships, and the relationship between gender and asexuality. With more articles comes more information and visibility. This information can help combat the discrimination that comes from misunderstandings, ideally decreasing the burden for asexual individuals to educate others and increasing their feelings of safety if they wish to come out.

Contributions and Future Directions of Asexuality Research

Guz et al. (2022) noted that a lot of early literature focused on what asexuality was (i.e., defining the concept and understanding individuals in the community). More recent research has branched out and focused on the diversity of asexuality, relationships of asexual individuals, accepting the asexual identity, and asexuality in the LGBTQIA+ community. That said, only one recent article focused on the health of asexual individuals and sexual education (Sheppard et al., 2024). This subject—physical and mental health in asexual people—could be a place for future research, expanding the knowledge base.

Additionally, a few articles looked into the relationships of asexual individuals, but those relationships were regarding an asexual individual paired with an allosexual individual (Glass, 2022; Hart-Brinson et al., 2024; Higginbottom, 2024). This leaves a gap in research on relationships between two asexual people. Many of the articles talked about their samples needing more diversity among participants (Brandley & Dehnert, 2023; Glass, 2022; Hart-Brinson et al., 2023;

Higginbottom, 2024; Peters, 2022; Sheppard et al., 2024; Yang, 2023). Adding this information in the future could help provide better and more accurate data on asexual individuals and their experiences.

A third area for future additional research is the language people use when referring to asexuality as a category and in regards to asexual people. Only one article (from a sociology and communication perspective, instead of a psychological perspective) focused on this specific question (Hart-Brinson et al., 2023). These authors interviewed asexual participants regarding their experiences feeling alienation from general society and the pressure to fit into a world with assumed allosexuality. Their participants noted a struggle with implicit language and stigma used when referring to the asexual community. This kind of implicit prejudice and microaggression could be an interesting topic for future investigation. While several studies in Guz et al. (2022) discussed the pathologizing of asexuality (i.e., that it is considered by some a problem or lack of a "natural" human instinct or tendency), implicit values and cultural norms about sexuality passed through language usage should be further explored.

Finally, some future directions that were not mentioned in the articles reviewed but that could be of importance are mental health and intersectionality (Crenshaw, 1989). LGBTQIA+ individuals face many mental health concerns, such as depression and anxiety (Mongelli et al., 2019). It would be helpful to know how and why these concerns affect asexual people, specifically. In addition, while Boot-Haury and Cusick (2023) covered gender diverse asexual individuals, there are many other identities that asexual individuals have (e.g., race, ability, SES). Understanding how intersectionality mitigates or aggravates the life experiences of asexual individuals provides new connections and a deeper understanding of this important subject.

Limitations

Like the papers reviewed here, the current analysis also has limitations. Perhaps most importantly, only qualitative papers published in the last few years were included. While a focus on qualitative research has benefits (Charmaz, 2006; Flavin, 2001; Opsal et al., 2016; Stutterheim & Ratcliffe, 2021), a more comprehensive review would have included both qualitative and quantitative studies. In addition, articles were only found through PsycINFO and Google Scholar, which led to most of the articles being psychology based. While Google Scholar is interdisciplinary, using other databases (e.g., PubMed) might have uncovered additional studies. These investigations might be particularly relevant to questions such as what topics are being studied in cutting-edge, current research in other disciplines. Finally, it has only been a few years since Guz et al.'s (2022) paper was published. More research on asexuality is being done every day, so additional reviews will be needed in the future.

Conclusion

This paper focused on qualitative research on asexuality within recent years. This research is important as asexuality is a topic that does not receive much attention. This paper aims to (1) increase education about the asexual community that is currently lacking and (2) decrease stigma against the community, which might then lead to reduced prejudice and discrimination. In short, the more people know about asexuality, the better everyone can serve asexual individuals and the larger community by increasing acceptance of this currently marginalized population.

References

Asexuality.org. (2024). *The asexual visibility and Education Network*. https://www.asexuality.org/

Boot-Haury, J. W., & Cusick, K. M. (2023). Intersectional asexual and transgender and gender diverse identity and existential concerns: A thematic analysis. *The Humanist Psychologist* [advance online publication]. https://doi.org/10.1037/hum0000346

Brandley, B., & Dehnert, M. (2023). "I am not a robot, I am asexual": A qualitative critique of allonormative discourses of ace and aro folks as robots, aliens, and monsters. *Journal of Homosexuality*, *71*(6), 1560-1583. https://doi.org/10.1080/00918369.2023.2185092

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. Sage Publications.

Crenshaw, K. (1989). Demarginalizing the intersection of race and sex: A black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. *University of Chicago Legal Forum*, 139-167.

Dawson, M., McDonnel, L., & Scott, S. (2017). Note on recruitment as an ethical question: Lessons from a project on asexuality. *International Journal of Social Research Methodology*, *20*(3), 255–261. https://doi.org/10.1080/13645579.2017.1287871

Flavin, J. (2001). Feminism for the mainstream criminologist: An invitation. *Journal of Criminal Justice*, *29*, 271-285. https://doi.org/10.1016/S0047-2352(01)00093-9

Foisy, C., Bosom, M., & Lima, I-A. (2022, April 8). *Coming out as asexual and carrying the burden*

of educating others. Club SexU. https://clubsexu.com/en/coming-out-as-asexual-and-carrying-the-burden-of-educating-others/

Glass, V. Q. (2022). Queering relationships: Exploring phenomena of asexual identified persons in relationships. *Contemporary Family Therapy*, *44*, 344-359. https://doi.org/10.1007/s10591-022-09650-9

Guz, S., Hecht, H. K., Kattari, S. K., Gross, E. B., & Ross, E. (2022). A scoping review of empirical asexuality research in social science literature. *Archives of Sexual Behavior*, *51*, 2135-2145. https://doi.org/10.1007/s10508-022-02307-6

Hart-Brinson, P., Tlachac, M. L., & Lepein, E. (2023). Contradictions in experiences of compulsory sexuality and pathways to asexual citizenship. *Sexuality & Culture*, *28*, 187-213. https://doi.org/10.1007/s12119-023-10110-1

Higginbottom, B. (2024). The nuances of intimacy: Asexual perspectives and experiences with dating and relationships. *Archives of Sexual Behavior*, *53*, 1899-1914. https://doi.org/10.1007/s10508-024-02846-0

Kaur, H. (2019, Oct. 20). *Asexuality isn't celibacy or abstinence. Here's what it is—and isn't*. CNN. https://www.cnn.com/2019/10/20/us/asexuality-explainer-trnd/index.html

Lech, S., Köppe, M. Berger, M., Alonso-Perez, E., Gellert, P., Herrmann, W., & Buspavanich, P. (2024). Depressive symptoms among individuals identifying as asexual: A cross-sectional study. *Scientific Reports*, *14*, Article 16120. https://doi.org/10.1038/s41598-024-66900-6

Mollet, A. L. (2021). "It's easier to just say I'm queer": Asexual college students' strategic identity management. *Journal of Diversity in Higher Education*, *16*(1), 13-25. https://doi.org/10.1037/dhe0000210

Mongelli, F., Perrone, D., Balducci, J., Sacchetti, A., Ferrari, S., Mattei, G., & Galeazzi, G. M. (2019). Minority stress and mental health among LGBT populations: An update on the evidence. *Minerva Psichiatricia*, *60*(1), 27-50. https://doi.org/10.23736/S0391-1772.18.01995-7

Opsal, T., Wolgemuth, J., Cross, J., Kaanta, T., Dickmann, E., Colomer, S., Erdil-Moody, Z. (2016). "There are no known benefits…": Considering the risk/benefit ratio of qualitative research. *Qualitative Health Research*, *26*(8), 1137-1150. https://doi.org/10.1177/1049732315580109

Peters, C. (2022). Asexuality, affect aliens, and digital affect cultures: Relationality with the happy objects of sexual and romantic relationships. *Western Journal of Communication*, *86*(5), 585-600. https://doi.org/10.1020/10570314.2022.2100473

Sheppard, A., Mann, E. S., & Pfeffer, C. A. (2024). "Your brain isn't all backwards": Asexual young women's narratives of sexual healthism. *The Qualitative Report*, *29*(1), 229-243. https://doi.org/10.46743/2160-3715/2024.5116

Stutterheim, S. E., & Ratcliffe, S.E. (2021). Understanding and addressing stigma through qualitative research: Four reasons why we need qualitative studies. *Stigma and Health*, *6*(1), 8-19. https://doi.org/10.1037/sah0000283

Yang, Y. (2023). Gender uncoupled: Asexual people making sense of high school sex talk. *Sexualities*, *26*(3), 372-387. https://doi.org/10.1177/13634607211033865

Author Note

Correspondence may be addressed to: Wind Goodfriend, Ph.D., Psychology Department, 610 W. 4th Street, Buena Vista University, Storm Lake, IA 50588, email goodfriend@bvu.edu, telephone 712-749-2108.

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.