2021

Volume 25, Number 2, Fall

Issn 1085-6641

Journal of Psychological Inquiry

Journal of Psychological Inquiry

Journal of
Psychological Inquiry

Volume 25 Number 2 Fall, 2021

Journal Contents

From the Editor's Desk

Time really does fly. Here we are again, at the end of another year, and what a way to end the year. The research in the Fall 2021 issue was truly unparalleled. I was incredibly impressed with each of these students' work. Their contributions to the field do not go unnoticed, and I can only imagine what we may see from them again in the future.

The future... for the last few years, the future has been difficult to imagine and bleak. Now, it seems that things are beginning to look up. As vaccination rates continue to rise and possible treatments for COVID-19 become available, the future begins to feel hopeful. It now feels as if there is a light at the end of the tunnel, and we will make it.

I am honored to be given the opportunity to help Dr. Ken Sobel copyedit the last three issues of this journal. It is a privilege to be a part of undergraduate psychological research. These students are the future of our field, and, because of that, the future's looking bright.

Julianne Wright

Kristen Julianne Wright
Graduate Assistant
University of Central Arkansas

Journal of Psychological Inquiry
2021, Vol.25, No. 2, pp. 5-14

Elizabeth A. Dahl, Ph.D., Award for Excellence in Undergraduate Research

#TRENDING: THE IMPACT OF TWITTER ON THE CREATION OF FALSE MEMORY

GRACE WASINGER, TREY HILL, AND JOHN RAACKE

FORT HAYS STATE UNIVERSITY

Abstract – With popular networking sites, such as Twitter, there comes the potential for the diffusion of erroneous information and accounts of events. The spread of information in this way may present a peril to memory representations in the form of false memories. Misleading post-event information can lead to the formation of false memories in individuals. In Experiment 1, we presented participants with two videos that differed in emotional content. Then, individuals were exposed to false information through an information feed that had either high or low resemblance to modern Twitter feeds. In Experiment 2, individuals were again exposed to videos that differed in valence; however, the feed type changed based on what account type (i.e., professional law enforcement or personal college student Twitter account) was spreading false information. We found that positive emotion yielded less vivid event memories in individuals. Additionally, we found that false memories fluctuated based on source type. Individuals exposed to feeds where police officers tweeted out false information had higher levels of false memories. Ultimately, when individuals are seeking news from Twitter, we suggest users apply the same scrutiny to information coming from professional Twitter accounts that they would peer accounts.

Twitter is a social media platform that allows individuals to share messages within a 280-character limit. Twitter's microblogging style encourages users to briefly share their thoughts on a global scale. Every month approximately 330 million individuals engage with Twitter (Twitter.com, 2020). Modern Twitter has extended beyond personal usage to provide individuals with a forum to receive hard news in real time. In recent years, the number of individuals receiving world news via social media has increased. In 2018, 20% of Americans reported that they often receive their news from social media sites. However, in 2019, that rate increased to 30%. Furthermore, 71% of individuals on Twitter use the site for news (Shearer & Grieco, 2019). Twitter features, such as the trending tab, allow for individuals to see what topics or current events are gaining the most attention. Twitter differs from other social media sites (e.g., Facebook, LinkedIn) in that its purpose is to connect individuals with topics and ideas on a large scale. With Twitter, the goal of diffusing topics and ideas is much different than other sites that provide individuals with opportunities to network or catch up with other individuals. However, the speed of news transmission — a strength of Twitter— is also a problem when the news being transmitted is factually incorrect. False information encountered by mass amounts of users can have real-world implications. For example, a fake tweet from the Associated Press, a perceived credible source, about an attack on the white house caused the stock market to drop 130 points (i.e., around 1%); this drop began just one minute after the Tweet was sent from a group of hackers (Memmott, 2013). Recent research has shown that false news stories spread faster on Twitter than true ones (Vosoughi et al., 2018).

Over the past four years, the spread of unverified or false information, specifically through social media, has become a significant problem, most notably during the 2016 U.S. presidential election. Beyond the speed and scope of transmitting potentially false information, research also shows that misleading post-event information can alter individuals' memory of events; this is referred to as the misinformation effect (e.g., Loftus, 1979; Loftus et al., 1978; for reviews, see Ayers & Reder, 1998; Loftus, 2005). An explanation for this effect is the memory-trace replacement hypothesis which states that misleading post-event information interferes with original event memory, preventing individuals from retaining information that was encoded during the

original event (Ayers & Reder, 1998). Despite Twitter's popularity as a social media platform, research on its "fake news" and possible subsequent false memory of users is not extensive. The present study is filling an important, yet neglected, gap in the literature concerning law enforcement professional accounts and susceptibility to false memory on Twitter.

Fenn et al. (2014) evaluated the effect that Twitter usage has on false memory creation. To examine the relationship between false memory and Twitter, participants completed three experimental phases: encoding, misinformation, and a confidence test. The encoding phase consisted of fifty images that showed a man robbing a car. A distractor task was placed after encoding to reduce rehearsal. Then participants viewed one of the three information feeds: control, Twitter-control, and Twitter. The feed, which narrated the pictures of the robbery that was previously shown, contained both true and false information. Researchers also manipulated the language style used in each feed. For example, in the control feed language was formal, while the Twitter condition included the most informal language (e.g., slang, hashtags), and the Twitter-control condition included more formal language. In the Twitter group, researchers found that confidence was lower for false items, meaning individuals had less false memory creation in the Twitter condition. Additionally, researchers found the medium of the message to be the most significant factor in modulating false memory, with less false memories resulting from Twitter than the control (Fenn et al., 2014).

In a follow-up study by Griffin et al. (2017), false memory and Twitter were studied with the added component of familiarity. Familiarity was introduced during the encoding phase when participants read a story about their home university (i.e., Michigan, high familiarity) or a university hours away (i.e., Iowa, low familiarity). During the misinformation phase, participants read through a Twitter feed which contained both true and false information. Additionally, the Tweeter also changed based on familiarity: Michigan or Iowa students. Researchers found higher false memory formation for the less familiar Iowa Twitter feed than the more familiar Michigan feed. Further, false memory formation was highest when Twitter feed and content were different (e.g., Michigan author and Iowa story). Thus, social media users take the author's expertise into account when encountering potentially false information online (Griffin et al., 2017).

Source Credibility of Information

Research shows that when individuals are told a source has low credibility or a likelihood to deceive, then post-event information is evaluated more carefully (for reviews, see Echterhoff et al., 2005; Pena et al., 2017; Wright et al., 2009). Perceived source credibility can have a powerful effect on the creation of false memory. For example, Echterhoff et al. (2005) found that when source credibility is low, false memory formation is also low. When a Twitter account is professional (e.g., law enforcement, legitimate news outlet), individuals may be likelier to believe false information as opposed to if the tweet containing false information comes from an individual Twitter user's personal account.

Law Enforcement and Twitter

Twitter has allowed police agencies around the world to keep citizens updated and engaged within their environments (Van de Velde et al., 2015). In addition to informing the public of relevant information, researchers have found that police departments can increase their perceived legitimacy simply by using Twitter as a medium to communicate with citizens (Grimmelikhuijsen & Meijer, 2015). Despite the promising interactions that can ensue between police department Twitter accounts and the public, there have also been negative interactions between these two parties. For example, in 2014, the New York Police Department Twitter account asked users to use the hashtag "#MyNYPD" to inform the NYPD of important events occurring in their policing areas. Responses to this hashtag varied; however, the majority of the interactions with this tweet were negative and disclosed the maltreatment of individuals by officers (Grimmelikhuijsen & Meijer, 2015).

Role of Emotion in False Memory Formation

In comparison to events that are neither positive nor negative in content, events with strong emotion are better recalled (Cahill & McGaugh, 1995, Damme & Smets, 2014). There is much variability in the formation of false memories; however, emotion plays a significant role in the formation of false memories (Bookbinder & Brainerd, 2016). Building on research by Porter et al. (2008), which examined the role that emotional content has in false memory creation, the present study will test the impact that a negatively or positively charged event has on false memory creation. The aforementioned study, Porter et al. (2008), showed a differential impact on memory recollection based on event type. In their 2008 study, individuals formed more vivid negative event memories and false negative memories than they did positive (Porter et al., 2008). This finding is supported by the Paradoxical Negative Emotion (PNE) hypothesis, which states that memories of negative events are better remembered, but also these memories are more susceptible to misleading post-event information than positively charged memories (Porter et al., 2008).

However, the literature concerning emotion and false memory has contradictory theories, and not all research literature supports the PNE hypothesis like Porter et al. (2008). For example, some literature finds support for positive emotion being more susceptible to false memory creation than negative; this positive emotional effect on false memory is explained by positive emotion leading to less detailed information processing than negative emotion, leading individuals to confuse details and misremember events. (for reviews, see Levine & Bluck, 2004; Kensinger & Shacter, 2006; Forgas, Laham, & Vargas, 2005).

Additionally, research investigating the influence that affective states have on false memory formation found that individuals in a negative mood were less susceptible to false memory effects (Storbeck & Clore, 2005). Moreover, Storbeck (2013) found that participants in negative mood conditions were particularly accurate in item-specific information and had less false memories than participants in positive conditions.

Hypotheses

To further explore the role misinformation can have on Twitter users, we conducted an initial experiment and a follow up experiment. In Experiment 1, participants viewed a video that was either positive or negative in emotional valence. A video was chosen, as opposed to pictures used in Fenn et al. (2014), because we hoped videos could potentially offer higher ecological validity and help to communicate the emotionality of the stimuli easier. False memory research has been successful in using visualization to create rich false memories (i.e., intense memories high in emotion and confidence for completely made up events) in participants (Loftus, 2003). Using video to modulate false memory has shown to be a reliable resource (e.g., Otgaar et al., 2013).

Additionally, participants were divided into feed type (e.g., Twitter feed, control feed). Based on aforementioned credibility research from Echterhoff et al. (2005), we hypothesized that participants exposed to the Twitter feed would show lower retention of false post-event information. This lower false memory rate is expected in part due to research from Fenn et al. (2014); researchers found that participants considered the medium of the message while viewing false post-event information. Participants were more skeptical of the Twitter condition, which resulted in less false memory formation. Additionally, we hypothesized that participants exposed to the positive emotional content would demonstrate more false memory. This hypothesis is due to research showing positive emotion is more susceptible to the inclusion of misleading information into memory representations (e.g., Levine & Bluck, 2004; Kensinger & Shacter, 2006; Forgas, Laham, & Vargas, 2005).

Experiment 1

Methods

Participants

Participants were recruited from a state comprehensive regional university in the Midwest and offered course credit for participation; 283 English speaking undergraduate students (171 women, 109 men, 3 non-report) between the ages of 18 and 22 (M = 19.3, SD = 2.1) participated. The majority of students identified as White, Non-Hispanic (n = 199, 70.3%), followed by Hispanic (n = 45, 15.9%), and African American (n = 13, 4.6%) as the next largest identifiers. A large majority of participants in the experiment had an active Twitter account (n = 195, 68.9%), with most having the account for over 2 years (n = 162, 83.1%).

Materials and Procedure

A 2 (emotional content: positive or negative) x 2 (narrative feed type: high Twitter resemblance or low Twitter resemblance) between-subjects design was used to look at the aforementioned research question. Participants completed three experimental phases (e.g., encoding, misinformation, and retrieval phase).

Encoding Phase

Participants watched a short video under two minutes in length that contained either positive or negative emotional content. In the positive video, a baby with a rare disorder that affects his vision reacts to seeing his mother for the first time due to the use of special glasses. In the negative video, reporters show live video from a SWAT standoff involving a murder-suicide outside a suburban Texas neighborhood. After the video was watched, a distractor task was used to prevent mental repetition of information used to enhance participants' memory of the video. During this task, participants were given a timed trigonometry test. Using only paper and pencil, participants were asked to complete as many mathematical problems as they could within the time limit of five minutes.

Misinformation Phase

The misinformation phase required participants to read from a news feed that had a high or low resemblance to Twitter. The feed narrated the events of the video and incorporated misleading post-event information. Participants were randomly assigned to either the control (n=135) or Twitter feed (n= 148). The Twitter feed did not contain user handles or pictures; however, the feed included the Twitter bird on each

Tweet as well as the blue and white color theme of the social networking site. The top of the Twitter feed read "Twitter Feeds," while the control feed read "Narrative of Events" at the top. The control feed had no color, and information appeared in bullet form. Both the control and Twitter feed showed four lines each for 30 seconds before new text appeared at the bottom. Both feeds contained true and false information. Each feed contained 20 lines of information about the respective video watched. For the feed that corresponded to the video concerning a murder-suicide, a misinformation example read, "the incident involved a young married couple." However, as stated by news reporters in the video, the couple was elderly. The language used was also changed based on the Twitter or control condition. For example, in the Twitter condition that same piece of information read, "I heard it was a young married couple with lots of kids #howcrazy." The Twitter condition, along with informal language, also contained emojis to better resemble Twitter. Additionally, participants were not given notice that the feed may contain some misinformation.

Retrieval

Participants were asked to answer questions about items remembered in the video that they watched using an undifferentiated scale from definitely-saw-in-the-video to definitely-did-not-see-in-the-video. The items consisted of 10 items that appeared in both the video and feed, 10 false items that appeared only in the feed, and 10 false items that appeared neither in the feed nor the video. Participants were also asked to rate how much attention they paid to the information feed using an undifferentiated scale from not-at-all to very-much. Additionally, participants were also asked to rate their level of trust in the information feed using an undifferentiated scale from not-at-all to very-much.

Results

To test the hypotheses, a series of three separate 2 (emotional content: positive or negative) x 2 (narrative feed type: high Twitter resemblance or low Twitter resemblance) between-subjects factorial ANOVAs were performed, one for each of the three dependent variables used to assess false memory: true information, implanted information, and completely false information. Also, the same analysis was performed on both an attention and a trust measure.

True Information

Results showed a main effect for emotion, F (1, 279) = 13.86, p <.001, η^2 = .06. There was also a statistically significant main effect for feed type, F (1, 279) = 6.27, p = .01, η^2= .06, with Twitter (M = 69.10, SD = 14.36) eliciting lower levels of false memory than the control group (M = 72.78, SD = 15.52). However, there was not a statistically significant interaction between feed type and emotion, F (1, 279) = 1.53, p = .218, η^2= .06).

Implanted Information

For implanted information, there was a statistically significant main effect for feed type, F (1, 279) = 4.01, p = .046, η2 = .02; similar to the results of true information, Twitter (M = 28.68, SD = 12.06) elicited lower false memory scores than the control condition (M = 31.74, SD = 13.49). There was not, however, a statistically significant main effect for emotion, F (1, 279) = .12, p= .73, η2 =.02. There was also not a statistically significant interaction between feed type and emotion, F (1, 279) = .73, p = .39, η2= .02.

Completely False Information

There was a statistically significant main effect for emotion, F (1, 279) = 122.93, p < .001, η^2 =.31. There was not a statistically significant main effect for feed type, F (1, 279) = 1.24, p = .267, η^2=.3, or a statistically significant interaction between feed type and emotion, F (1, 279) = 0.01, p = .918, η^2= .31.

Trust

Results showed no statistically significant main effect for feed type, F (1,260) = 1.06, p = .303, η2 < .01. There was not a statistically significant main effect for emotion, F (1, 260) = 0.13, p = .717, η2< .01. There was also not a statistically significant interaction between emotion and feed type, F (1, 260) = 0.75, p = .388, η2< .01.

Attention

There was not a statistically significant main effect for feed type, F (1, 261) = 1.02, p =.315, η^2 < .01. There was also no statistically significant main effect for emotion, F (1, 261) = 0.24, p =.627, η^2 < .01. However, there was a statistically significant interaction between feed and emotion, F (1, 261) = 4.20, p =.041, η^2 = .02.

Discussion

Feed type made a significant difference in the recall of true information in participants. For true information, participants assigned to the Twitter group had better recall than the control group. In Experiment 1, the findings for true and completely false information concerning emotion is in accordance with previous research (for reviews, see Damme & Smets, 2014; Porter et al., 2008) that found emotional content plays a role in false memory formation. However, additional research is needed to better understand the link between positively and negatively charged emotional videos on implanted information. Additionally, the present experiment echoes

the results of Fenn et al. (2014), finding that false memory formation is lower when the medium of the message is Twitter. This effect can be explained, in part, through trust of the medium. Participants rated their trust as higher for the control condition than participants in the Twitter condition. More research is warranted that seeks to better understand source credibility judgments on Twitter. The present results supported our first hypothesis that participants in the Twitter condition will show lower retention of false post-event information. Lastly, the second hypothesis that positively charged emotional content will result in higher retention of false post-event information was not supported. Emotional content was found to be significant in two out of the three dependent variables; this effect was not found for implanted information. Additional research is needed to understand the role of positively emotionally charged events on the recollection of implanted post-event information on Twitter.

Experiment 2

Little research has addressed the difference between professional and personal Twitter account type on the formation of false memories. Experiment 1 showed that individuals took into account the medium of the message when reading about a narrative of events, meaning participants were more skeptical of information coming from a Twitter feed than a standard information feed (i.e., significantly less false memory for Twitter than control). It is unclear if this credibility judgment will extend to specific Twitter profiles and create more false memory. However, it is known that individuals respond to authority cues on social media sites. Westerman et al. (2014) found that participants were more influenced by Tweets coming from the American Heart Association, due to authority cues being a robust credibility indicator. Experiment 2 aims to investigate the role that Twitter account type (e.g., credibility cues) has on the formation of false memory, while also seeking to better understand the role that emotionally charged events can play into a memory representation.

Hypotheses

We hypothesized that participants exposed to feeds in which law enforcement individuals were tweeting false post-event information would have higher retention of false information. In addition to source credibility, emotional content was again evaluated. However, in Experiment 2 both the positive and negative video concerned law enforcement and the videos were closer in duration. Additionally, the events of the videos selected were more emotionally charged than in the first experiment. In Experiment 2, the negative video showed footage of a police shooting of a 19-year-old, while the positive video showed live footage of police officers bonding and dancing with members of their community. With the changes made to the emotional stimuli in Experiment 2, we hypothesized that participants exposed to positive emotional content would demonstrate higher retention of false information. This hypothesis is largely due to extensive literature showing the interaction between positive emotion and integration of misleading post-event information (for reviews, see Levine & Bluck, 2004; Kensinger & Shacter, 2006; Forgas, Laham, & Vargas, 2005).

Methods

Participants

Undergraduate students from a state comprehensive regional university in the Midwest were offered course credit in return for their participation. The sample was composed of seventy-five English speaking undergraduate students (42 women, 30 males, 3 non-report) between the ages 18 and 22 (M = 19.3, SD = 1.3). The majority of the students identified as White (n = 54, 72%), followed by Hispanic (n = 9, 12%). Additionally, a large majority of participants in the experiment had an active Twitter account (n = 56, 74.7%), with most having the account for over 3 years (n = 29, 51.8%).

Materials and Procedures

For the second experiment, a 2 (emotional content: positive and negative) x 2 (false information source: false law enforcement or false peer) mixed-subjects design was developed to look at the aforementioned research question. As in Experiment 1, Experiment 2 consisted of three phases: encoding, misinformation, and confidence test. All participants were exposed to both positively and negatively emotionally charged videos, while the between-subjects factor was feed type.

Encoding Phase

Participants watched both the positively and negatively charged video; however, each video was seen one day apart from the other. Additionally, we counterbalanced the presentation order by having the two groups of participants watch a different video first (e.g., the first set of participants watched the positive video first, while the second set watched the negative video first). The positively charged video concerned community policing, while the negatively charged video involved an officer shooting a 19-year-old. The videos were more similar in theme (e.g., both videos concerned the subject of law enforcement) and duration (e.g., two minutes) than in the first experiment. After the encoding phase, participants were asked to complete a 5-minute mathematics test to reduce rehearsal of information.

Misinformation

Participants received a Twitter feed that either had misleading post-event information tweeted by professional (i.e., high credibility, law enforcement) or peer (i.e., low credibility, college student) accounts. Each feed integrated tweets from both professional and personal Twitter accounts, but feeds differed in what type of user was spreading false information. The group that received the Twitter feed in which professional sources (i.e., law enforcement accounts) tweeted the false post-event information was labeled false law enforcement, while the group of participants that read peer accounts (i.e., college students) tweeting false post-event information was labeled false peer. For example, the video clearly stated that the male was shot in the abdomen; however, Twitter feeds contained implanted information that he was shot in the chest and arm. Further, in the false law enforcement condition a Twitter user named "Trooper Holly" tweeted "Hammond was shot in the chest and arm." While in the false peer condition, a Twitter user named "Seri Heaton" tweeted "Hammond was shot in the chest and arm." The presence of law enforcement vs. peer accounts were designed to manipulate Twitter user authority cues leading participants to potentially make credibility judgments.

To manipulate credibility, we changed the Twitter circle icon (i.e., a place where Twitter users can upload an image of themselves) to a picture of a law enforcement individual or a badge. Additionally, the name next to the icon indicated that it was either a law enforcement account or a particular officer. Furthermore, for the low credibility peer accounts, Twitter users were made to look like college students, instead of professionals. Additionally, the Twitter feeds were more realistic than they were in the first experiment. The aesthetics of the feeds were matched to look nearly identical to the displays of Twitter at the time the experiment occurred. Unlike Experiment 1, each Tweet appeared by itself and remained there for 5 seconds until another appeared from the bottom. The feed was developed in this way to help mimic the Twitter experience of scrolling through a feed.

Retrieval

The confidence test included 25 items that appeared in both the video and feed, 10 false items that appeared only in the feed, 10 false items that appeared neither in the feed nor the video, and 5 items that appeared in the video only. Participants were asked to respond to each question on an undifferentiated scale from definitely-saw-in-the-video to definitely-did-not-see-in-the-video. Participants were also asked to rate their level of trust in the Twitter feed and how much attention they paid to the Twitter feeds on an undifferentiated scale from not-at-all to very-much. In an effort to ethically combine participants' responses from the two days, participants were asked to report the last 5 digits of their phone number at the top of the survey.

Results

In order to test the aforementioned hypotheses, we performed a series of 2 (emotional content: positive or negative) x 2 (false information source: false law enforcement or false peer) mixed-subjects factorial ANOVAs. One ANOVA was performed for each of the four dependent variables (e.g., true information, implanted information, completely false information, and information appearing only in the video). The same analysis was used to measure attention and trust.

True Information

Results showed a statistically significant main effect for emotional content, $F(1, 73) = 4.81$, $p = .03$, $\eta 2 = .06$. For feed type there was a statistically significant main effect, $F(1, 73) = 5.28$, $p = .024$, $\eta 2 = .07$, with false peer ($M = 5.22$, $SD = 0.83$) resulting in lower rates of false memory than false law enforcement ($M = 5.56$, $SD = 0.63$). There was also a statistically significant interaction between feed type and emotion, $F(1, 73) = 4.71$, $p = .03$, $\eta 2 = .06$.

Implanted Information

For information implanted into the feed, there was not a statistically significant main effect for emotion, $F(1, 73) = 1.58$, $p = .21$, $\eta 2 = .02$. There was also no main effect for feed type, $F(1, 73) = .58$, $p = .45$, $\eta 2 = .01$. However, there was a statistically significant interaction between emotion and feed type, $F(1, 73) = .56$, $p = .01$, $\eta 2 = .01$.

Video Only

For information left out of the feed and present in only the video, there was a main effect for emotion, $F(1, 73) = 16.20$, $p < .001$, $\eta 2 = .18$. There was also a main effect for feed type, $F(1, 73) = 5.51$, $p = .02$, $\eta 2 = .07$, with false peer ($M = 4.45$, $SD = 1.09$) showing lower false memory rates than false law enforcement ($M = 4.88$, $SD = 0.98$). Additionally, there was a significant interaction, $F(1, 73) = 4.33$, $p = .04$, $\eta 2 = .06$.

Completely False Information

For information that was completely made up, there was only a statistically significant main effect for emotion, $F(1, 73) = 67.78$, $p < .001$, $\eta 2 = .48$. All other effects were not statistically significant.

Attention

For attention to the video, emotion was not statistically significant, $F(1, 71) =.29$, $p =.59$, $\eta 2 =.00$.

Feed type was also not significant, F (1, 71) = .36, p = .55, η2 =.00. For emotion & feed there was a main effect found F (1, 71) =10.30, p =.00, η2 =.13. Attention to the negative video (M = 65.01, SD = 28.34) was higher than attention to the positive video (M = 64.32, SD = 28.23). For the negative video, false peer attention (M = 68.33, SD = 25.70) was higher than false law enforcement (M = 62.30, SD = 30.40). For the positive video, false peer attention (M = 57.12, SD = 30.81) was much lower than false law enforcement attention (M = 70.25, SD = 24.80).

Trust

For trust, emotion was statistically significant, F (1, 71) = 7.96, p = .01, η2 =.10. However, feed type was not statistically significant, F (1, 71) = .27, p = .61, η2 = .00. Emotion & feed was not statistically significant, F (1, 71) = 2.07, p = .15, η2 =.03. The positive video (M = 29.38, SD = 23.09) participants rated higher in trust than the negative video (M = 22.44, SD =19.17). For the negative video, false peer trust (M = 25.55, SD = 17.25) was higher than false law enforcement trust (M = 19.88, SD = 20.48). For the positive video false peer trust (M = 28.79, SD = 23.26) was slightly lower than false law enforcement trust (M = 29.88, SD = 23.22).

Discussion

When false information is tweeted by law enforcement individuals, false memory formation was found to be higher in participants. This finding supports our hypothesis that when law enforcement individuals tweet misinformation, participants would report higher retention of false post-event information. Participants in the condition where law enforcement tweeted the false information had higher rates of false memories than those in the condition where peer accounts tweeted the false information. Our findings were not consistent with our second hypothesis that participants exposed to positive emotional content would demonstrate higher retention of false information. More research is needed to understand how the recollection of Tweets is influenced by negatively and positively charged events.

General Discussion

Does the medium of the message influence false memory?

When individuals are presented with information from a Twitter feed as opposed to the more traditional information slideshow, individuals have lower rates of false memory formation. Fenn et al. (2014) also found this decrease in false memory to be true of participants in the Twitter condition. This effect may be due to participants evaluating Twitter with a higher level of scrutiny and perceiving Twitter sources as less credible.

Does the person Tweeting influence false memory?

Generally speaking, false information from law enforcement elicits more false memory. Literature concerning perceptions of police officers' credibility on Twitter is not yet extensive. However, one way of explaining these results is through credibility short-cuts. Heuristics are short-cuts that require little cognitive exertion. Individuals may use heuristics on Twitter to help users to evaluate source credibility quickly and with little effort. This is done by looking for information about a Twitter user's number of followers, reputation, and identity (Park & Kaye, 2019; Vaidya et al., 2019). When making credibility judgments on Twitter, accounts that seem to be close to the event are viewed as more credible (Murthy, 2018). In Experiment 2, both videos concerned events in which law enforcement officers were involved; because of this, law enforcement individuals tweeting about the event were perhaps seen as more credible than college student users.

How does attention & trust play into this?

In negative videos, people pay more attention to content when normal people are tweeting than when professional Twitter accounts tweet. However, when the video is positive, people pay more attention to law enforcement. For trust, when the emotional content of the video was positive, individuals trusted the feed more. Recent advances in credibility research show that negative emotion has the propensity to make individuals less trusting (Engelmann et al., 2019).

Limitations & Strengths

Because the sample was gathered from college aged individuals (e.g., Experiment 1: *M* = 19.3, Experiment 2: *M* =19) from a regional university in Kansas, results may be heavily influenced by geographical location and age. Twitter usage may be influenced by age. Researchers found that on Myspace younger individuals were more inclined to follow other users that were close in age, as opposed to older individuals that had much more variability (Pfeil et al., 2009). Moreover, on Twitter, college students' followers may be heavily composed of peers rather than professional Twitter accounts, whereas middle aged individuals may opt to follow more professional accounts. Griffin et al. (2017) found that false memory formation was higher in feeds that were less familiar. Thus, it may be particularly important to further investigate this relationship between familiarity and age on false memory formation for social media users. It is unclear whether these results extend beyond law enforcement and into other professional Twitter accounts, such as political figures or celebrities.

Additionally, results may be influenced by participants' relationship and history with law enforcement, particularly in Experiment 2 where law enforcement was more heavily emphasized. Future studies should aim to sample from the more generalizable public. Additionally, it is not well understood if a change in social media platform would yield the same results.

Fenn et al. (2014) and Griffin et al. (2017) provided the groundbreaking research concerning false memory and Twitter; however, this connection is still not very well understood as it relates to emotional content. The current study lacks the ability to compare the effects between positively and negatively valenced videos, because emotional content was manipulated using a positively and negatively valenced news video only (e.g., there was no neutral video).

The second experiment is unique in that it aims to understand how the dissemination of misleading post-event information from professional Twitter users may influence the creation of false memory in a way that was not yet examined. With law enforcement agencies and police officers utilizing Twitter more frequently, it is important to untangle the various ways that media users will be impacted by the shift from peer to professional influences. Ultimately, through this research, we caution users to be more attentive to sources that seem to be most credible.

Future Directions

Future studies concerning emotional content, false memory, and Twitter should remedy the limitation of the current study by using three levels to compare the effects between positive and negative emotion. Additionally, this effect should be examined using different visual/audio formats (e.g., pictures, videos, podcasts, and word-lists).

Future research should aim to understand the relationship between high-credibility sources (e.g., political figures, police officers) and language styles (e.g., formal, informal) utilized through Tweets. High profile users' language styles may have real-world implications on users' perceptions and memory representations.

Additionally, more research is needed to examine how users make decisions when sharing information on a wide scale, such as Twitter. For example, it may be particularly important to assess how individuals make credibility judgments while seeking news from Twitter.

References

Ayers, M. S., & Reder, L. M. (1998). A theoretical review of the misinformation effect: Predictions from an activation-based memory model. Psychonomic Bulletin & Review, 5(1), 1–21. https://doi.org/10.3758/BF03209454

Bookbinder, S. H., & Brainerd, C. J. (2016). Emotion and False Memory: The Context-Content Paradox. Psychological Bulletin, 142(12), 1315-1351. https://doi.org/10.1037/bul0000077

Cahill, L., & McGaugh, J. L., (1995). A novel demonstration of enhanced memory associated with emotional arousal. Consciousness and Cognition, 4(4), 410–421. https://doi.org/10.1006/ccog.1995.1048

Damme, I., & Smets, K. (2014). The power of emotion versus the power of suggestion: Memory for emotional events in the misinformation paradigm. American Psychological Association, 14(2), 310-320. https://doi.org/10.1037/a0034629

Echterhoff, G., Hirst, W., & Hussy, W. (2005). How eyewitnesses resist misinformation: Social postwarnings and the monitoring of memory characteristics. Memory & Cognition, 33(5),770–782. https://doi.org/10.3758/BF03193073

Engelmann, J., Meyer F., Ruff, C., & Fehr, E. (2019). The neural circuitry of affect-induced distortions of trust. Science Advances, 5(3), 1-16. https://doi.org/10.1126/sciadv.aau3413

Fenn, K., Griffin, M., Uitvlugt, N., & Ravizza, R. (2014). The effect of Twitter exposure on false memory formation. Psychonomic Bulletin & Review, 21(6), 1551-1556. https://doi.org/ 10.3758/s13423-014-0639-9

Forgas, J. P., Laham, S. M., & Vargas, P. T. (2005). Mood effects on eyewitness memory: Affective influences on susceptibility to misinformation. Journal of Experimental Social Psychology, 41, 574–588. https://doi.org/10.1016/j.jesp.2004.11.005

Griffin, N. R., Fleck, C. R., Uitvlugt, M. G., Ravizza, S. M., & Fenn, K. M. (2017). The tweeter matters: Factors that affect false memory from Twitter. Computers in Human Behavior, 77, 63-68. https://doi.org/ 10.1016/j.chb.2017.08.032

Grimmelikhuijsen, S., & Meijer, A. (2015). Does Twitter increase perceived police legitimacy? Public Administration Review, 75(4), 598-607. https://doi.org/10.1111/puar.12378

Kensinger, E. A., & Schacter, D. L. (2006). When the Red Sox shocked the Yankees: Comparing negative and positive memories. Psychonomic Bulletin & Review, 13, 757–763. https://doi.org/10.3758/BF03193993

Laney, C., & Loftus, E. F. (2013). Recent advances in false memory research. South African Journal of Psychology, 43(2), 137-146. https://doi.org/10.1177/0081246313484236

Levine, L. J., & Bluck, S. (2004). Painting with broad strokes: Happiness and the malleability of event memory. Cognition and Emotion, 18, 559–574. https://doi.org/10.1080/02699930341000446

Loftus, E. (1979). Eyewitness testimony. Harvard University Press.

Loftus, E. (2003). Make-believe memories. American Psychologist, 58(11), 867-873. https://doi.org/10.1037/0003-066X.58.11.867

Loftus, E. F. (2005). Planting misinformation in the human mind: A 30- year investigation of the malleability of memory. Learning & Memory, 12(4), 361–366. https://doi.org/10.1101/lm.94705

Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. Journal of Experimental Psychology: Human Learning & Memory, 4(1), 19–31. https://doi.org/10.1037/0278-7393.4.1.1

Loftus , E. F., & Pickrell, J. E. (1995). The formation of false memories. Psychiatric Annals, 25, 720-725. https://doi.org/10.3928/0048-5713-19951201-07

Memmott, M. (2013). AP Twitter Account Hacked, Tweet About Obama Shakes Market [Audio podcast transcript]. In All things considered. NPR. https://www.npr.org/sections/thetwo-way/2013/04/23/178620410/ap-twitter-account-hacked-tweet-about-obama-shakes-market

Murthy, D. (2018). Twitter. Medford, MA: Polity Press.

Otgaar, H., Howe, M. L., Peters, M., Sauerland M., & Raymaekers L. (2013). Developmental trends in different types of spontaneous false memories: implications for the legal field. Behavioral Sciences & the Law, 31(5), 662-682. https://doi.org/10.1002/bsl.2076

Park, C. S., & Kaye, B. K. (2019). Expanding Visibility on Twitter: Author and Message Characteristics and Retweeting. Social Media Society, 5(2), 1-10. https://doi.org/ 10.1177/2056305119834595

Patihis, L., Frenda, S., Leport, A., Petersen, N., Nichols, R., Stark, C., . . . Loftus, E. (2013). False memories in highly superior autobiographical memory individuals. Proceedings of the National Academy of Sciences of the United States of America, 110(52), 20947-20952. https://doi.org/ 10.1073/pnas.1314373110

Pena, M. M., Klemfuss, J. Z., Loftus, E. F., & Mindthoff, A. (2017). The effects of exposure to differing amounts of misinformation and source credibility perception on source monitoring and memory accuracy. Psychology of Consciousness: Theory, Research, and Practice, 4(4), 337–347. https://doi.org/10.1037/cns0000137

Pfeil, U., Arjan, R., & Zaphiris, P. (2009). Age differences in online social networking- A study of user profiles and social capital divide among teenagers and older users in MySpace. Computers in Human Behavior, 25(3), 643-654. https://doi.org/10.1016/j.chb.2008.08.015

Porter, S., Taylor, K., & Ten Brinke, L. (2008). Memory for media: Investigation of false memories for negatively and positively charged public events. Memory, 16(6), 658-666. https://doi.org/10.1080/09658210802154626

Shearer, E., & Grieco, E. (2019). Americans are wary of the role social media sites play in delivering the news. Pew Research Center, https://www.journalism.org/2019/10/02/americans-are-wary-of-the-role-social-media-sites-play-in-delivering-the-news/

Storbeck, J., & Clore, G. (2005). With sadness comes accuracy; with happiness, false memory: Mood and the false memory effect. Psychological Science, 16(10), 785-791. https://doi.org/10.1111/j.1467-9280.2005.01615

Storbeck, J. (2013). Negative affect promotes encoding of and memory for details at the expense of the gist: Affect, encoding, and false memories. Cognition and Emotion, 27(5), 800-819. https://doi.org/ 10.1080/02699931.2012.741060

Twitter.com. (2020). About Twitter, Inc [Website information]. Retrieved January 20, 2020, from about.twitter.com/company

Vaidya, T., Votipka, D., Mazurek, M., & Sherr, M. (2019). Does being verified make you more credible?: Account verification's effect on Tweet credibility. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, 1-13. https://doi.org/10.1145/3290605.3300755

Van de Velde, B., Meijer, A., & Homburg, V. (2015). Police message diffusion on Twitter: Analyzing the reach of social media communications. Behaviour & Information Technology, 34(1), 4-16. https://doi.org/10.1080/0144929X.2014.942754

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and fasle news online. Science, 359(6380), 1146-1151. https://doi.org10.1126/science.aap9559

Westerman, David, Spence, Patric R, & Van Der Heide, Brandon. (2014). Social media as information source: Recency of updates and credibility of information. Journal of Computer-mediated Communication, 19(2), 171-183. https://doi.org/10.1111/jcc4.12041

Wright, D. B., Memon, A., Skagerberg, E. M., & Gabbert, F. (2009). When eyewitnesses talk. Current Directions in Psychological Science, 18(3), 174–178. https://doi.org/10.1111/j.1467-8721.2009.01631.x

Journal of Psychological Inquiry
2021, Vol.25, No. 2, pp. 14-22

CROWDSOURCED VERSUS CRITERION-BASED CREATIVITY MEASUREMENTS OF DRAWINGS

HALEY SCHIELDS, DANIELLE WHITFIELD, AND ANDREW T. JOHNSON

PARK UNIVERSITY

Abstract – This study presents a comparison of rankings for 45 drawings (three sets of 15 pictures) based upon creativity scores from an Amazon MTurk sample and a separate criterion-based analysis (adapting three measures from the Torrance Tests of Creative Thinking - Figural Forms A and B). The results reveal positive correlations of MTurk versus Torrance scores for two of three picture sets and no differences in the rank orders. These results suggest that a criterion-based analysis is a viable and reliable strategy for determination of creativity in drawings.

Keywords: Assessing creativity, Torrance, Drawings

Creativity may be perceived as a central part of a person's identity, an essential component of jobs, and a desired characteristic of artists, entrepreneurs, and visionaries. As such it would be beneficial to have a comprehensive and effective method of measuring creativity. Assessing creativity is complex due to its multidimensional nature, as it cannot be described adequately through measuring a single characteristic or trait (Kim, 2006). This sentiment was expressed by Mel Rhodes in his 1961 analysis of creativity, "I had collected forty definitions of creativity and sixteen of imagination" (p. 306). Even the definition of creativity has been debated, and is often taken from multiple perspectives by both experts and individuals who have their own definition (Plucker & Runco, 1998; Runco & Jaeger, 2012). However, there are consistent criteria that appear in the standard definition of creativity, namely, originality and effectiveness. Originality, as described by Runco and Jaeger (2012), is essential to defining creativity, as commonplace or conventional work is not original, meaning it lacks creativity. Although originality is crucial to defining creativity, it alone is not sufficient. Effectiveness is needed in addition to originality for something to be described as creative. The standard definition suggests that creating something original is pointless unless it is effective in achieving something (Runco & Jaeger, 2012). In other words, something identified as creative must be uncommon, useful or purposeful, as well as appropriate or possessing the ability to function as intended and fulfill an intended purpose. Along with being useful and appropriate, effectiveness can also be measured by the assigned value, e.g, the monetary value, of the original creation or thought (Runco & Jaeger, 2012). However, while the standard definition used in creativity research in the past has been beneficial, some researchers have found that this definition is not as comprehensive and effective as it could be. Currently, creativity tests can be used to assess these criteria and determine the creativity of an individual. There are a range of creativity tests from subjective ones, such as the Picture Completion Task (creating a picture around a set stimulus), to objective ones, such as the Remote Associates Test (RAT) where a person is given three words like Cottage, Swiss, Cake, and are asked to indicate a fourth word that relates to all three words (in this case Cheese) (Mednick, 1968; Bowden & Jung-Beeman, 2003). There are even sophisticated tests of creativity, e.g., the Torrance Tests of Creative Thinking (TTCT). It has been proposed that the inclusion of additional criteria into the standard definition could allow for all aspects of the process of creativity to be measured more accurately through measures such as these tests.

Researchers have addressed the measurement of creativity from many different perspectives based on different definitions of creativity, in an attempt to standardize the criteria for creativity measurement to ensure a more accurate assessment (Runco & Jaeger,

2012; Glăveanu & Beghetto, 2020; Guilford, 1950; Smith & Smith, 2017; Stein, 1953; Simonton, 2018; Plucker & Runco, 1998). Runco and Jaeger (2012), clarified the basis of the standard definition of creativity in research by outlining past researchers who mentioned two major criteria of creativity: Originality and Effectiveness (Glăveanu & Beghetto, 2020). One of the first researchers to suggest that creativity could be studied scientifically was J. P. Guilford in 1950 (Runco & Jaeger, 2012). Guilford (1950) defined creativity as "the abilities that are most characteristic of creative people, which determine if the individual has the power to exhibit creative behavior to a noteworthy degree" (p.444). In other words, creativity is a part of the individual's personality (Guilford, 1950). It is not something that is exclusive only to some, and almost all individuals are capable of creativity to some degree (Guilford, 1950). Guilford also acknowledged that the idea of creativity is not the same in all situations, and that there is a need for creativity to accomplish things in society in areas such as the arts, science, and leadership (Guilford, 1950; Smith & Smith, 2017). Guilford's perspective was an early introduction to the current standard definition of creativity, because he also emphasized the importance of both the presence of the novelty or originality of ideas, as well as the realisticness and acceptableness or the effectiveness of an idea (Runco & Jaeger, 2012).

In addition to Guilford, Stein (1953) was also another researcher who applied similar criteria to those included in the standard definition to creativity research (Runco & Jaeger, 2012). Stein (1953) defined a creative work as something that is novel, which means it deviates from other previously existing works. This connects to the idea of originality present in the standard definition of creativity (Runco & Jaeger, 2012). Stein also introduced other important ideas such as the concept that creative works tend to benefit a group, which in turn allows for the presence of social judgment on the work (Runco & Jaeger, 2012). This group benefit fits within the criteria of effectiveness in the standard definition, as it suggests that the outcome or idea achieves something. He also mentioned that creativity stems from previous knowledge or materials in which inspiration is drawn, but the final work is something that contains new elements (Runco & Jaeger, 2012).

While the standard definition has been useful in creativity research, some researchers mention that it is limited and does not fully encompass all aspects of creativity (Csikszentmihalyi, 1999; MacKinnon, 1987). For instance, Glăveanu and Beghetto (2020) describe the standard use of the criteria of originality and effectiveness in creativity research as having little consideration for any other aspect of creativity other than the outcome. Having a standard definition that has a focus solely on the outcome of creativity can result in a loss of understanding about the full process and experience of creativity (Glăveanu & Beghetto, 2020). Specifically, there is a lack of understanding about the individual who is displaying creativity, such as their individual differences and process (Glăveanu & Beghetto, 2020). Glăveanu and Beghetto (2020) proposed a more comprehensive standard definition that includes the criteria of originality and effectiveness, but also has a focus on creativity as an experience. An extensive overview of this may be found in Csikszentmilhayli's 1996 book Creativity: Flow and the psychology of discovery and invention. This emphasis is based on the idea that creativity stems from an individual's experiences, situations, contexts, and lived events (Glăveanu & Beghetto, 2020). This is important because the impact of an individual's subjective and interpersonal experiences are often not thoroughly addressed in creativity research (Beghetto, 2016). Experiences challenge, interrupt the normal course of action, and present something which an individual has no predetermined answer, therefore allowing for creativity to take place (Glăveanu & Beghetto, 2020). Glăveanu and Beghetto (2020) state that a more comprehensive standard definition of creativity could consist of an individual having a novel experience of meaningful actions and interactions, which demonstrate the principles of open-endedness nonlinearity, pluri-perspective and future orientation.

Smith and Smith (2017) present a similar view to Glăveanu and Beghetto (2020) on the limitations of the standard definition of creativity, as well as describe a new measurement of creativity that includes some of the original aspects of the standard definition. Because the standard definition has a large focus on the product of creativity, there is also an emphasis on the success of the outcome or idea. Smith and Smith (2017) state that if success is required for something to be seen as creative, many ideas or outcomes that could potentially be seen as creative are lost, and research on creativity is limited. Instead of evaluating the success of the final product when determining what is creative or not, ideas should fit the criteria of having the potential to be successful and the focus should be on the process of creativity (Smith & Smith, 2017). The standard definition is not discarded entirely as there is still a focus on the criterion of novelty, but the process of the formation of ideas is most important. Smith and Smith (2017) instead proposed a 1.5 criteria model of creativity, in which the 1.0 criterion is novelty, and the 0.5 criterion is utility. These criteria are similar to those outlined by Runco and Jaeger (2012),

however they fit more closely to those described by Glăveanu and Beghetto (2020) because of the focus on the process rather than the outcome of the creative activity. The benefit of this is that including process more fully represents creativity (Stanley & Lehman, 2015).

While it is important to understand what makes something creative, Simonton (2018) emphasized that to understand creativity it may be useful to explore how to determine what is not creative as well as what is creative. According to Simonton (2018), it is difficult to have a comprehensive definition of creativity if the criteria that make something uncreative are not also outlined. For instance, one can not succeed in one of the criteria included in a definition of creativity, but could succeed in the other one or two aspects (Simonton, 2018). Also, if a definition has more than one criterion than the typical two seen in the standard definition, it is easier to see the shortcomings and failures regarding specific standards, making it easier to determine if something can be called creative (Simonton, 2018). Simonton (2018) describes ideas that are not creative as routine, fortuitous response based, having irrational perseveration, and problem finding. Simonton also proposes three components of creativity, of which he places in an equation to be used to develop a definition of creativity. The components are, the initial probability or response strength (p), final utility (u), and prior knowledge (v). The two first components are similar to the standard definition of creativity, with the initial probability referring to originality, final utility referring to usefulness or effectiveness of the idea, and prior knowledge referring to surprise or how much new knowledge an individual gained from creating the end result. These criteria are also very similar to the definition of creativity that is also utilized by the United States Patent Office, which defines creativity as new, useful, and nonobvious (Simonton, 2012). The formula to measure creativity, with "c" representing creativity, is expressed as, $c = (1-p)\ u\ (1-u)$ where $0 \leq c \leq 1$ (Simonton, 2018). Simonton (2018) explains that if $c = 0$ creativity is nonexistent, however, if $c = 1$ creativity is at its greatest.

In addition to the debate regarding the definition of creativity, some researchers also take issue with the methods used to measure creativity. Plucker and Runco (1998) mention that much of the research in creativity measurement is based on methods that are psychometric or developed due to the belief that there is not a fully accurate way to measure creativity. Because it is so difficult to accurately define creativity in its many forms, creativity research has attempted to measure creativity through both implicit theories of creativity consisting of individual definitions of creativity, as well as explicit ones that rely on expert definitions. Implicit theories of creativity can consist of qualities such as adventurous, artistic, and curious, while explicit theories of measurement rely on a specific definition described by an expert (Plucker & Runco, 1998). Focusing on implicit theories of measurement in the future of creativity research can be beneficial because when an individual demonstrates creativity, they are not typically thinking of the concept of being creative from an expert's perspective (Plucker & Runco, 1998). The individual most likely has their own definition of what creativity is, so finding a way to measure the individual's thought process and actions can help researchers to gain a better understanding of how to use these implicit theories in combination with explicit theories to accurately measure creativity. Examining different perspectives on how to define creativity, as well as adding new concepts into existing measurements, can help to strengthen them and increase their accuracy.

One of the measurement issues with assessing an artifact's creativity is whether individuals are making separate judgments of creativity and originality. Multiple sources, e.g., Runco and Jaeger (2012), include originality/ novelty in their definition of creativity. Knoche et al. (2016) explored this issue by having an online sample rate drawings on creativity, originality, and artistry. They used a crowdsourced sample from Amazon MTurk to rate a total of 45 drawings. They divided the 45 drawings into three sets (15 drawings each) and had a different sample rate for each set for Creativity, Originality, and Artistry on a 5-point Likert scale (Very high … Very low). The sample sizes for Drawing sets 1, 2 and 3 were 146, 152, and 153, respectively. Across all three sets the ratings for Creativity, Originality, and Artistry were significantly correlated with each other. In addition, their results indicated significant positive correlations between ratings of Creativity, Originality, and Artistry. As ratings increased for Creativity, so did Originality and Artistry. However, regarding their research question about whether Creativity, Originality, and Artistry ratings would significantly differ from each other (indicating separate or orthogonal judgments), they received a mixed set of results. For Drawing Set 1 there was a significant difference between creativity and artistry, and originality and artistry, but not between creativity and originality (although the result was approaching significance). For Drawing Set 2, there was only a significant difference between Originality and Artistry. Creativity judgments were not significantly different from Originality or Artistry. For Drawing Set 3, there was a significant difference between all three

judgments. This mixed set of results underscores the variability of samples, drawings, and judgments.

Measuring creativity in drawings is challenging. Creativity judgments can be influenced by multiple factors, e.g., color, humor, vivid imagery, storytelling, artistry, etc. Amezcua et al. (2018) examined whether there would be a difference in ratings for drawings and the rankings for drawings on creativity. They presented participants with a set of 25 creative drawings made by adults. Each participant rated the drawings for creativity, originality, and artistry. Following this, participants rank-ordered the drawings from most creative to least. The ratings for the drawings were used to rank order the drawings and then they were compared to the direct rank orders to observe any differences. They hypothesized that the rating order of the drawings would be similar to the ranking order. A Wilcoxon Matched-pair analysis did not reveal any statistical difference between the two rank orders. They concluded that they could have saved time by omitting one of the activities, e.g., rank ordering, and that rating and ranking lead to similar results. Having participants simply rank-order the drawings for creativity would be faster but lose a higher level of measurement, e.g., interval.

Is there a more efficient way to measure these drawings? Using diverse crowdsourced samples also brings in a greater potential of variability due to socio-cultural and affordance differences (Glaveanu, 2013; Glaveanu & Beghetto, 2020). The present study seeks to explore this issue through a modified replication of the Knoche et al. 2016 study. Instead of using a crowdsourced sample to judge the creativity of the 45 drawings, a criterion-based strategy will be used instead. We propose to adapt scoring measures from the Torrance Tests of Creative Thinking (TTCT) Figural Forms A and B. The TTCT is a type of divergent thinking test, in which a figural prompt is given to the participant who then uses it in their creative response (Puryear et al., 2019). The responses are then scored by assessing specific aspects of the responses (Puryear et al., 2019). The TTCT consists of five measurements: Fluency, Originality, Elaboration, Abstractness of Titles, and Resistance to Premature Closure (Scholastic Testing Services, 2017). The TTCT Figural Test is typically completed in three activities consisting of: picture construction, picture completion, and repeated figure of lines or circles (Kim, 2006). We propose to use three of the five TTCT measures (Originality, Resistance to Premature Closure, and Elaboration). The remaining two measures were not appropriate to include. Participants did not create a title for the drawing (Title) and did not create any drawings (Fluency). A set of raters will be trained on the adapted TTCT measures, and two raters will assess each drawing to allow for a reliability measure. The criterion-based measures will be combined into a composite score that will be used to rank order the drawings. The main research question we will be exploring is whether the crowdsourced rankings of Knoche et al. (2016) will be different from the criterion-based adapted TTCT rank orders.

Methods

Materials

The current study utilized the materials and pertinent MTurk data from the Knoche et al. (2016) study. This study received IRB approval from the sponsoring institution. Three sets of 15 drawings gathered from adult participants. More specifically, the drawings were created by college students taking an upper-level cognition course during discussion of creativity. The drawings sampled came from students who consented to give their instructor the drawings. No personal descriptors were on the drawings ensuring anonymity. The students were instructed to produce their most creative drawing on a standard-sized piece of copy paper. Students were instructed to embed a simple figure (resembles a capitalized letter L, with a capitalized letter C bisecting its horizontal line) at least once within their drawing.

Three individuals (one professor and two undergraduate research assistants) were trained on the Torrance Test of Creative Thinking for Originality, Resistance to Premature Closure, and Elaboration using the Torrance Tests of Creative Thinking: Streamlined score guide for Figural Forms A and B. (Scholastic Testing Service, 2017). Next, The three raters were trained adapting the three TTCT measures of Originality, Resistance to Premature Closure, and Elaboration, to drawings. The training consisted of discussing the measures, scoring some drawings together, and then scoring drawings independently. Each individual then scored two of the three drawing sets using an online coding sheet that gathered the entrant's last name, drawing identification number, and the scores for Originality, Resistance to Premature Closure, and Elaboration (see Appendix).

Each set of 15 drawings were previously assessed by crowdsourced participants on: How Difficult the embedded figure was to find; How Creative the drawing was; How Original the drawing was; and how Artistic the drawing was (each rated on a 5-point Likert scale) (Knoche et al., 2016). There were 146 MTurk participants for Drawing Set 1, 152 for Set 2, and 153 participants for Set 3. The scores for Creativity, Originality, and Artistry

were summed for a total score and then drawings in each set were rank ordered from highest to lowest scores.

Reliability was calculated using Pearson's r statistic. The outcomes yielded significant positive relationships. The reliability measures for each set are as follows: Set 1 ($r = 0.851$, $n = 15$, $p < 0.001$), Set 2 ($r = 0.875$, $n = 15$, $p < 0.001$), and Set 3 ($r = 0.884$, $n = 15$, $p < 0.001$). These outcomes indicate that the raters' scores were highly reliable.

Procedure

In the present study three individuals were trained on the Torrance Tests of Creative Thinking (TTCT) Figural Forms measures of Originality, Resistance to Premature Closure, and Elaboration. Two individuals scored each of the three sets of 15 drawings from the previous study. Originality was scored as a 1 or a 2, with a score of 1 being typical or an unoriginal use of the figure such as an arm, hand, flower, robot, circle, wheel, or glasses, and a 2 being an original use of the figure that was more embedded into the drawing. Resistance to Premature Closure was scored as a 0, 1, or 2. A score of 0 represented that the figure was closed quickly and directly, a score of 1 represented that the figure was closed less directly, and a score of 2 represented that the figure was not closed at all or that it was closed in a broad way. An Elaboration score was determined by a count of each unique feature presented in the drawing six categories: 1 = 0-5, 2 = 6-16, 3 = 13-19, 4 = 20-26, 5 = 27-33, 6 = > 33 unique features. The combined adapted TTCT measurements totaled to 10 points.

Results

Our main research question investigated whether there would be a difference in the creativity rankings of the same drawings assessed by the MTurk participants and scored by research assistants using three modified criterion from the Torrance Test of Creative Thinking Figural assessment. The means and standard deviations of the Crowdscourced and the Criterion-based outcomes for the three sets of drawings can be found in Table 1.

A Pearson product-moment correlation coefficient assessed the relationship between the Crowdscourced composite scores and the Criterion-based scores. There were significant positive correlations for the scores in Drawing Set 1($r = 0.583$, $n = 15$, $p = 0.022$), and Set 3 ($r = 0.691$, $n = 15$, $p = 0.004$), but no correlation for Set 2 ($r = - 0.013$, $n = 15$, $p = 0.964$). For two of the drawing sets the different assessments were related.

The drawings in each set were ranked based upon the crowdsourced scores and the criterion-based scores respectively. Then the two rankings for each drawing set were submitted to a Wilcoxon Signed-Ranks Test. The results yielded nonsignificant differences between the Crowdsourced rankings and the Criterion-based rankings for all three drawing sets ($t = 43.50$, $p > .001$, $t = 51.50$, $p > .001$, and $t = 39.00$, $p > .001$, respectively). This indicates that the two assessments produced similar rankings.

Table 1

Means and Standard Deviations for Crowdsourced and Criterion-based Scores Across Drawing Sets

Drawing Set	Crowdsourced		Criterion-based	
	M	SD	M	SD
Set 1	9.46	1.16	6.80	2.02
Set 2	9.08	1.02	5.53	2.26
Set 3	9.18	1.49	6.60	1.70

Note: The range of scores possible for the Crowdsourced scores were 3-15 and the range for the Criterion-based scores were 2-10.

Discussion

It has been over 70 years since J. P. Guilford issued the challenge for researchers to develop a greater focus on measuring creativity. Since that time, there have been significant successes, e.g., Torrance Test of Creative Thinking and the Remote Associates Test (Bowden & Jung-Beeman, 2003). However, these measures are lacking the authenticity that current researchers are calling for (c.f. Glăveanu, & Beghetto, 2020).

Assessing the creativity in drawings is especially challenging. On one hand, researchers can use crowdsourced responses with a hope that an accurate estimate of creativity is found at the midpoint of the score distribution. On the other hand, there is so much variability and subjectivity rendering a creativity judgment. We addressed these issues by adapting measures taken from an established and respected instrument - TTCT. We sought to investigate whether both assessment strategies ranked drawings in a similar way. We found that criterion-based rankings do not differ from crowdsourced ones. Given this outcome, essentially replicated in all three drawing sets, suggests that using the criterion-based approach is preferable because of the high reliability.

Moreover, for two of the three picture sets, there are strong positive correlations between crowdsourced scores and criterion-based scores. Drawings that were rated as higher in overall creativity received higher

criterion-based scores. However, Drawing Set 2 did not receive a similar outcome. In fact, there was no correlation between the crowdsourced and criterion-based scores. We examined the individual scores to observe any discernible patterns in the scores. For the criterion-based scores the drawings had scores spanning the entire range of possible scores. There was a cluster of scores in the middle. For the crowdsourced scores almost all of the scores were clustered in the middle of the range of possible scores and only separated by tenths and, in some instances, hundredths of a point. This restricted range of scores could be explained by regression to the mean or, more simply, that participants were being "nice" and did not want to give low ratings because creativity is in the eye of the beholder. This could be characterized as a demand characteristic effect. In comparing rank positions for the individual drawings in Set 2, the crowdsourced rank positions were lower for drawings with a rank difference of four or more positions except for one case. More specifically two of the drawings in Set 2 received the lowest scores possible indicating little creativity and what Simonton (2018) would describe as routine and reproductive. However, the crowdsourced scores for these drawings placed them in the middle of the range for scores. This type of outcome is a disadvantage of collecting judgments from a general population with no formal training in scoring creative works. Moreover, the results obtained for Drawing Set 2 underscore the advantage of using criterion-based measurements and trained scorers.

That stated, the criterion-based scores for Drawing Set 2 indicate a weakness that could be addressed. In this set eight of the 15 drawings received scores within two points of each other in the middle of the range. Adding more criteria would allow for a greater potential range of scores and variability to the outcomes. This study used three out of five criteria employed in the Torrance Test of Creative Thinking Figural test. Adding in more criteria would also balance out the effect of Elaboration scores. Drawings, while having a strong authentic aspect, have a challenge for measuring all of the variability possible. This also affects measurement and the reliability of scoring them. Having the Elaboration score account for six of the possible ten points a drawing can receive represents one of the stronger limitations in this study. Higher Elaboration scores can come from adding more features but also adding to the quality of the features as well. For example, additional points are earned by use of color, shading, and repetition and may appear to be peripheral to the main original idea. We observed some of these influences in Drawing Set 2. Specifically, for one of the drawings there was a scene depicting a neighborhood with a house and a driveway, a tree near the front of the drawing, and another house in the background of the drawing. A car, garage, stop sign, and two depictions of people were also included in this drawing. However, there is no shading or use of color, and everything is drawn as just linework. This drawing was scored highly on the criterion-based measurement largely based upon a high elaboration score, but was rated very low in the MTurk study. The incongruence in these outcomes may demonstrate a lack of representative characteristics for creative work in the standardized rating methods.

For future research in this area, we recognize that assessing the creativity in drawings is a challenge. Based upon discussions in the creativity literature, there is a growing emphasis on authentic processes, e.g., Beghetto (2016), Corazza (2016), Glăveanu, & Beghetto, (2020), and Runco (2019). The drawings that we assessed in this study represent authentic works and while this is in greater harmony with the current trend in the creativity field, there is a lot of variability in the drawings. While using a predetermined set of figures like that seen in the Torrance Test of Creative Thinking Figural test is easy to score, it is not very authentic. We think there is a middle ground using authentic drawings while utilizing the greatest strength of the TTCT - that is namely scoring criteria and trained scorers. Providing a tabula rasa with a figure to embed allows the drawers to play and evolve their work with no specific end objective in mind. This structure is central to the thesis of Stanley and Lehman's (2015) book Why greatness cannot be planned: The myth of the objective. Their work in evolutionary computer programming and artificial intelligence demonstrates that creativity emerges as a byproduct of intelligent serendipitous discovery rather than the objective. In other words, creative outcomes are not the prescribed destination and occur as a product of exploration.

Moving forward, it would be beneficial to add additional criterion-based measures to the composite measurement of Originality, Resistance to Premature Closure, and Elaboration. While these measures were found to be highly reliable, adding more measures could allow for a richer and more dimensional view of creativity. This could further differentiate more creative work from less creative work. For instance, these additions could have influenced the scores on Drawing Set 2, spreading out the scores and providing a clearer rank order. Possible limitations may include longer training sessions, longer rating durations for each drawing, and potential threats to reliability by adding in additional measurements.

Since both the Originality and Resistance to Premature Closure are scored up to two points each, the remaining six possible points come from the Elaboration score and this may skew the final score. For example, an image could score poorly in both Originality and Resistance to Premature Closure, but score highly in Elaboration simply by having a lot of features, e.g., color, shading, repetition, perspective, etc., but not contributing directly to the creativity of the outcome. Adding additional measurements could offset the larger contribution of the Elaboration score. Additional measures could be developed from the 13 Creative Strengths presented in the Torrance Tests of Creative Thinking Streamlined Scoring Guide for Figural Forms A and B. More specifically, this may include: Unusual Visualization, Richness of Imagery, or Storytelling Articulateness (Scholastic Testing Service, 2017). Unusual Visualization could be represented and measured in the drawings by the inclusion of an object or a situation that is being used in both an old way and a new way, or by utilizing a unique perspective beyond an upright or straight-on view of an object. Richness of Imagery could be represented by how much the individual can construct a story behind their drawing by adding liveliness and variety. Storytelling Articulateness could also be represented by how clearly the story of the image is communicated to others based on the context and environment included in the drawing (Scholastic Testing Service, 2017). In addition to balancing out the influences on the final score, these added measures each have a slightly different focus from the three used in the current study, therefore they could provide more representative characteristics for creative work.

Studies like the one presented here highlight the evolution of creativity research and more specifically how to measure the creativity expressed in drawings. While a sample of participants can provide ratings for how creative they perceive drawings to be, these will always be fraught with subjectivity and demand characteristics. We have demonstrated that criterion-based measurements can be an effective alternative, and there can be improvements to those. The Torrance Test of Creative Thinking Figural test provides a standard of reliability and accuracy, yet lacks the authenticity that current researchers are emphasizing. We believe giving individuals a simple figure to embed and the widest opportunity to produce their most creative work better addresses the issue of authenticity. Perhaps the future exploration of determining creativity in drawings will continue to evolve to maximize the advantages and minimize the disadvantages building upon the work of theorists such as E. Paul Torrance.

References

Amezcua, R., Foulks, H., Griffin, T., & Johnson, A. (2018, April). The fast and the furious: Ranking versus rating pictures for creativity. (Oral presentation at the 2018 Missouri Academy of Science, University of Missouri, Rolla, MO).

Beghetto, R. (2016). Creative learning: A fresh look. Journal of Cognitive Education and Psychology, 15, 6-23.

Bowdin, E. M., & Jung-Beeman, M. (2003). Normative data for 144 compound remote associate problems. Behavior Research Methods, Instruments, & Computers, 35, 634-639.

Csikszentmihalyi, M. (1996). Creativity: Flow and the psychology of discovery and invention. Harper/ Collins.

Csikszentmihalyi, M. (1999). Implications of a systems perspective for the study of creativity. In R. Sternberg (Ed.), Handbook of creativity (pp. 313–335). Cambridge University Press.

Corazza, G. (2016). Potential originality and effectiveness: The dynamic definition of creativity. Creativity Research Journal, 28, 258-267. https://doi.org/10.1080/10400419.2016.1195627

Glăveanu, V. P. (2013). Rewriting the language of creativity: The Five A's framework. Review of General Psychology, 17(1), 69–81. https://doi.org/10.1037/a0029528

Glăveanu, V. P., & Beghetto, R. A. (2020). Creative experience: A non-standard definition of creativity. Creativity Research Journal, 33(2), 75-80. https://doi.org/10.1080/10400419.2020.1827606

Guilford, J. P. (1950). Creativity. The American Psychologist, 5(9), 444–454. https://doi.org/10.1037/h0063487

Kim, K. H. (2006). Is creativity unidimensional or multidimensional? Analyses of the Torrance Tests of Creative Thinking. Creativity Research Journal, 18(3), 251-259. https://doi.org/10.1207/s15326934crj1803_2

Kim, K. H. (2006). Can we trust creativity tests? A review of the Torrance tests of creative thinking (TTCT). Creativity Research Journal 18(1), 3-14. https://doi.org/10.1207/s15326934crj1801_2

Knoche, K., Jaehn, A., Johnson, A. & Griffith, R. (2016, April). Using crowdsourcing to rate the creativity of pictures. (Oral presentation at the 2016 Missouri Academy of Science, Lincoln University, Jefferson City, MO).

MacKinnon, D. W. (1987). Some critical issues for future research in creativity. In S. G. Isaksen (Ed.), Frontiers of creativity research (pp. 119-130). Bearly.

Mednick, S. A. (1968). Remote associates test. Journal of Creative Behavior, 2, 213-214.

Plucker, J. A., & Runco, M. A. (1998). The death of creativity measurement has been greatly exaggerated: Current issues, recent advances, and future directions in creativity assessment. Roeper Review: A Journal on Gifted Education, 21(1), 36–39, https://doi.org/10.1080/02783199809553924

Puryear, J. S., Kettler, T., & Rinn, A. N. (2019). Relating personality and creativity: Considering what and how we measure. Journal of Creative Behavior, 53(2), 232–245.

Rhodes, M. (1961). An analysis of creativity. Phi Delta Kappan, 42, 305–311.

Runco, M. A. (2019). Creativity as a dynamic, personal, parsimonious process. In R. A. Beghetto & G.E. Corazza (Eds.), Dynamic perspectives on creativity (Vol. 4, pp. 181-188). Springer.

Runco, M., & Jaeger, G. (2012). The standard definition of creativity. Creativity Research Journal, 24, 92-96. https://doi.org/10.1080/10400419.2012.650092

Scholastic Testing Service (2017). Torrance Tests of Creative Thinking: Streamlined score guide for Figural Forms A and B. Scholastic Testing Service, Inc.

Simonton, D. K. (2018). Defining creativity: Don't we also need to define what is not creative? Journal of Creative Behavior, 52, 80–90. https://doi.org/10.1002/jocb.137

Simonton, D.K. (2012). Taking the US patent office creativity criteria seriously: A quantitative three-criterion definition and its implications. Creativity Research Journal, 24, 97–106. https://doi.org/10.1080/10400419.2012.676974

Smith, J. K., & Smith, L. F. (2017). The 1.5 criterion model of creativity: Where less is more, more or less. The Journal of Creative Behavior, 51(4), 281-284. https://doi.org/10.1002/jocb.191

Stanley, K. O., & Lehman, J. (2015). Why greatness cannot be planned: The myth of the objective. Springer.

Stein, M. I. (1953). Creativity and culture. The Journal of Psychology, 36(2), 311-322. https://doi.org/10.1080/00223980.1953.9712897

Author Note

Data collection and analysis were sponsored by the Park University Chapter of Psi Chi. Portions of these findings were presented as an oral presentation at the 2016 Missouri Academy of Science, Jefferson City, MO. We have no conflicts of interest to disclose.

Correspondence concerning this manuscript should be addressed to Haley Schields, Park University. Email: haley.schields@park.edu

Appendix

Data Entrant Information

Scoring information taken from the *Torrance Tests of Creative Thinking: Streamlined Scoring Guide for Figural Forms A and B* (2017). Scholastic Testing Service Inc.

Entrant Last Name: _____

Picture ID: _____

Torrance Score – Originality • 1 • 2

Common – arm/hand, glasses, circle/wheel, flower, robot

Torrance Score – Resistance to Premature Closure • 0 • 1 • 2

0 – figure is closed in quickest easiest way

1 – figure is closed but additions are made around it or outside of it

2 – figure not closed or incidental

Torrance Score – Elaboration • 1 • 2 • 3 • 4 • 5 • 6

(1 pt. is given for each idea/ feature that is added, after determining the basic portion of the response).

1= 0-5, 2= 6-16, 3= 13-19, 4= 20-26, 5= 27-33, 6= > 33

Journal of Psychological Inquiry
2021, Vol.25, No. 2, pp. 23-36

EXECUTIVE FUNCTION AND NOVEL WORD LEARNING IN ACTIVE AND INACTIVE BILINGUALS

ATHENA SZETO AND CASS FOURSHA-STEVENSON

MOUNT ROYAL UNIVERSITY

Abstract – The relationship between bilingualism and cognition has been studied extensively with a great deal of literature both in support of and challenging the idea of a bilingual advantage. However, the effect of language use on the bilingual advantage has been suspected to be a confounding variable. Previous studies have reported bilingual participants demonstrating faster reaction times (RTs) and greater accuracy than monolinguals in tasks that test inhibition, task switching, and novel word learning. If regular language use and switching does have an effect on cognition, we predicted that participants who regularly switched between languages would demonstrate faster RT and accuracy when compared with other participants. Participants were divided into three groups: monolinguals, inactive bilinguals (bilinguals who regularly use only one language), and active bilinguals (bilinguals who regularly use more than one language), and were compared on their task performance to determine whether there is an effect of active language use on accuracy and RT. Contrary to our predictions, the monolingual group demonstrated significantly faster RT than the active bilingual group on the Simon task. We found no significant differences between the groups' RT and accuracy in the other tasks. No evidence supporting a bilingual advantage as a result of language use was found in the Simon, ANT, card sort, or novel word learning tasks. Conversely, active language use can result in a disadvantage in inhibitory processing in young adult bilinguals. It is possible that active language use may exhaust cognitive resources, thereby increasing RT.

Keywords: bilingual advantage, bilingualism, executive function, novel word learning, language use, inhibition, task-switching.

Many researchers have studied the effects of bilingualism and have found evidence of a bilingual advantage in areas of executive functioning such as inhibition and task-switching (Bialystok & Martin, 2004; Bialystok et al., 2006; Bialystok, 2008; Bialystok et al., 2012; Colzatio et al., 2008; Costa et al., 2008; Costa et al., 2009; de Bruin et al., 2015b; Pelham & Abrams, 2014; Prior & MacWhinney, 2010; Soveri et al., 2011; Treccani et al., 2009; Vega & Fernandez, 2011; Vega-Mendoza et al., 2015; Verreyt et al., 2015). In addition to executive function advantages, a bilingual advantage in novel word learning has also been reported (Kan & Sadagopan, 2014; Kan et al., 2014; Kaushanskaya & Marian, 2009a; Kaushanskaya & Marian, 2009b; Kaushanskaya & Rechtzigel, 2012; Kaushanskaya et al., 2013; Nair et al., 2016). Advantages associated with bilingualism do not appear to stop there, as research has found an effect of bilingualism across a vast area of cognitive domains from conflict resolution, executive control, cognitive flexibility, and even in storing information in working memory (Bialystok, 2008).

However, a meta-analysis by de Bruin et al. (2015a) argued that despite the volume of research that supports the existence of a bilingual advantage, there is a publication bias that prevents studies that did not find a bilingual advantage from being added to the existing pool of literature. de Bruin et al. (2015b) challenged the idea of a bilingual advantage arguing that despite the vast body of literature that demonstrated a bilingual advantage, there were also many studies that did not find significant differences between monolingual and bilingual participants (see Incera & McLennan, 2018; Kousaie et al., 2014; Paap & Greenberg, 2013). In fact, finding a bilingual advantage may be more difficult than expected as bilingualism is an experience that is difficult to control. Some examples of possible confounding variables include social economic status (SES), immigration, acculturation, age of acquisition (AoA) of participants' second language

(L2), language proficiency, similarity between language pairs, immersion, etc. (de Bruin et al., 2015b; Soveri et al., 2011).

One specific variable that de Bruin et al. (2015b) identified was that within the group of bilingual participants, participants would vary on their use of languages. Other studies have also found that some bilingual participants were more likely to switch between languages during the day, sometimes even within a conversation, depending on the environment (Luk & Bialystok, 2013; Soveri et al., 2011). Previous research has indicated that performance on tasks changes based on how often a person switches between languages, as participants who practice switching between languages regularly are more balanced in their proficiency in both languages and have more practice juggling the cognitive load of multiple languages, leading to better performance in executive function tasks (Luk & Bialystok, 2013; Pelham & Abrams, 2014; Verreyt et al., 2015). In novel word learning, participants' language use is usually measured to ensure that bilinguals differ in language experience from monolinguals, but it appears that no research has examined possible effects of varying language use on word learning (Kan et al., 2014). While many studies aim for bilinguals who are more balanced in their language usage to compare against monolinguals, it can also be helpful to examine the variability within the bilingual group.

The current study seeks to determine whether there is a difference between monolinguals and bilinguals in cognitive functioning and novel word learning tasks by examining the role of regular language use on executive functioning and ability to learn new words in a short period of time.

Inhibition

One major area of research that has been studied extensively is the bilingual advantage in inhibition tasks (Bialystok et al., 2004; Bialystok et al., 2006; Bialystok, 2008; Costa et al., 2008; Costa et al., 2009; Pelham & Abrams, 2014; Verreyt et al., 2015). Inhibition is often defined as the executive control ability to focus on relevant information while ignoring competing information or responses in order to achieve goals (Gamboz et al., 2009; Soveri et al., 2011). There are many theories as to why bilinguals have an advantage over monolinguals in non-verbal inhibition tasks. One theory is that bilinguals have more practice with inhibition as they are inhibiting their other language when communicating with others (Bialystok, 2008; Costa et al., 2008).

Studies have found that bilinguals of different ages all tend to perform better in inhibition type tasks than monolinguals (Bialystok, 2008; Bialystok et al., 2004; Bialystok & Martin, 2004; Costa et al., 2008). Pelham and Abrams (2014) compared bilinguals who learned their second language (L2) early in life with participants who became bilingual during their teenage years. They found that these late bilinguals shared cognitive advantages similar to the participants who acquired their L2 during childhood, indicating that age of acquisition (AoA) or length of bilingualism does not affect the bilingual advantage, but rather regular use of more than one language (Pelham & Abrams, 2014). Tasks such as the Simon task (Bialystok, 2008; Bialystok et al., 2004) and the Attentional Network task (ANT) (Costa et al., 2008; Pelham & Abrams, 2014) are commonly used to test inhibition in bilingual and monolingual participants.

The Simon task is an inhibition task that involves a stimulus appearing on one side of the computer screen and participants are asked to indicate the color of the stimulus by pressing a key on either the right- or left-hand side of the keyboard. The reaction times (RT) for congruent trials (stimulus appears on the same side of the screen as the key) are compared with the RT from incongruent trials (stimulus and key are on opposing sides). Bilinguals generally show a smaller difference in RT between congruent and incongruent trials, leading to the conclusion that bilinguals have an advantage in inhibition (Bialystok, 2008; Bialystok et al., 2004; Verreyt et al., 2015).

The ANT task tests inhibition, but also tests other parts of the attention network including monitoring and alerting (Costa et al., 2008; Pelham & Abrams, 2014). For this study we focused on the inhibition aspect, treating it as a flanker task. Like the Simon task, participants are asked to indicate the direction of a central arrow while inhibiting the direction of the flanking arrows around it. Once again bilinguals have been reported to have faster RTs when compared with monolinguals, which researchers have concluded that bilinguals have advantages with inhibiting conflicting information and have more efficient attentional networks (Bialystok, 2008; Costa et al. 2008; Pelham & Abrams, 2014).

Attempts to replicate these advantages have not always been successful (Incera & McLennan, 2018; Kousaie et al., 2014). When de Bruin et al. (2015b) matched their bilingual and monolingual participants, they found no differences between group RT scores suggesting bilingual advantage. Paap and Greenberg (2013) tested several groups of participants using many tasks that had reported a bilingual advantage and they too found no bilingual advantage. In fact, on the Simon

task they found a bilingual disadvantage (Paap & Greenberg, 2013).

Task-Switching

Another area in which a bilingual advantage has been reported is task-switching (Hernández et al., 2013; Prior & MacWhinney, 2010; Soveri et al., 2011; Vega & Fernandez, 2011; Xie & Dong, 2017). The bilingual advantage in task-switching has been theorized to stem from bilinguals switching back and forth between their languages because the mechanisms that underlie language switching are thought to be similar to general task-switching (Hernández et al., 2013). Further research done by Hernández et al. (2013) with Spanish-English participants has demonstrated that bilinguals have lower restart costs by demonstrating faster RTs in trials that have repeated cues. These results seem to indicate the presence of a bilingual advantage.

Task-switching is measured by a variety of different paradigms, including a task like the Wisconsin Card Sort Task (WCST), which measures switching amongst other executive control functions (Gamboz et al., 2009; Miyake et al., 2000; Monchi et al., 2001; Vega & Fernandez, 2011; Xie & Dong, 2017). Participants are given a card and asked to sort it by one of its features (e.g. card with yellow circle will be placed in the pile with the color yellow). Part way through the task, the sorting criteria will change. Participants who were previously sorting based on color will have to sort according to another feature (e.g., shape). Accuracy between switching and non-switching conditions will be compared with the expectation that bilinguals will make fewer errors and adapt to new switching rules faster than monolinguals (Vega & Fernandez, 2011).

Similar to findings produced in inhibition studies, a fair share of studies have found no bilingual advantage in task switching (de Bruin et al., 2015b; Kousaie et al., 2014; Paap & Greenberg, 2013). Kousaie et al. (2014) found, contrary to previous findings, one of the monolingual groups in the study had outperformed the bilinguals on the WCST. Using various other task-switching paradigms, neither de Bruin et al. (2015b) nor Paap and Greenberg (2013) were able to replicate a bilingual advantage in task-switching.

Novel Word Learning

In addition to general cognition advantages, bilingual advantages have been reported in novel word learning tasks (Kan & Sadagopan, 2014; Kan et al., 2014; Kaushanskaya, 2012; Kaushanskaya & Marian, 2009a; Kaushanskaya & Marian, 2009b; Kaushanskaya & Rechtzigel, 2012; Nair et al., 2016). It has been hypothesized that bilinguals' increased experience with different language systems results in an easier time learning new words in a short period of time (Kan & Sadagopan, 2014; Kaushanskaya & Marian, 2009b). However due to the complexity of word learning and large variability of bilingual experiences, the exact nature of the advantage is still unknown (Kaushanskaya & Marian, 2009b; Nair et al., 2016).

The role of bilingual age of language acquisition and the effect of phonological familiarity (words that sound like they belong to a language that a person is familiar with) on novel word learning has been examined (Kaushanskaya & Rechtzigel, 2012; Kaushanskaya et al., 2013; Nair et al., 2016). These studies found that bilinguals who acquired their L2 later had more difficulty learning new words (Nair et al. 2016), and that bilinguals had an easier time accurately recalling novel words when they were phonologically familiar and associated with their native language (Kaushanskaya, et al., 2013). Kan et al. (2014) presented Spanish-English bilinguals with a novel word paired with a novel visual stimulus and tested their ability to identify the corresponding word or stimulus. Kan et al. (2014) reported that bilinguals recalled significantly more novel words after a short learning period and that there was an interaction between increased speech practice and more language exposure that led to better novel word recall.

Active and Inactive Bilingualism

Many studies have examined the bilingual experience, focusing on various variables from AoA, to language proficiency, and language environments (Soveri et al., 2011). However, de Bruin et al. (2015b) argued that when studying bilinguals, it is important to put a stronger emphasis on the participant's language use rather than just their knowledge of L2. While other studies have also examined the effect of language use and potential effects on bilingualism, de Bruin et al. (2015b) grouped participants into three groups: monolinguals, active bilinguals (bilinguals who regularly use more than one language), and inactive bilinguals (bilinguals who know a L2 but regularly use only one language). Participants were then matched to control for possible confounding variables (eg. immigration, socioeconomic status (SES), education, age, etc.) and were asked to complete inhibition and task-switching tasks that had previously reported a bilingual advantage (de Bruin et al., 2015b). The researchers found no significant differences between groups and concluded that there was no bilingual advantage present (de Bruin et al., 2015).

Other studies examining bilingualism argued that rather than dichotomizing bilingualism, treating bilingualism as a continuous variable may help control for possible confounding variables and individual differences between participants (Incera & McLennan,

2018; Luk & Bialystok, 2013; Soveri et al., 2011). Luk and Bialystok (2013) argued that since bilingualism is made up of many different dimensions, using a non-categorical approach best captures the experience.

Present Study

The current study aims to examine the effect of active bilingualism on executive functioning tasks measuring inhibition and task-switching while also looking at novel word learning in a university setting in Alberta, Canada. Although Canada is officially a bilingual country, most Albertans regularly only use English in professional and public settings. We are interested in examining whether or not the existence of a bilingual advantage can be found when participants from a wide range language backgrounds are divided into groups according to their language usage. Participants were asked to fill out a survey to gather basic information such as language experience, AoA, and how often they use more than one language per day. From here, participants will be divided into groups: monolingual, active bilingual, and inactive bilingual. Language usage will also be examined as a continuous variable to control for confounding variables. The current study focuses on tests that have reported bilingual advantages in the past (Simon task, ANT, and WCST). In addition, the role of active bilingualism in novel word learning for various bilinguals will also be examined by presenting the same words from the Kan et al. (2014) study.

If there is a bilingual advantage, then bilinguals are predicted to perform better on the executive functioning and novel word learning tasks than monolinguals. If there is an effect of active bilingualism, then active bilinguals will outperform inactive bilinguals. Specifically, better performance will be demonstrated by lower RTs on the inhibition tasks and greater accuracy on the WCST and novel word learning tasks. These findings could provide further insight into the potential advantages associated with bilingualism.

Method

Participants

Participants were recruited from the pool of Introductory Psychology students at Mount Royal University; students signed up for the study online and received 1% of credit towards their final grade as compensation. A total of 61 participants, 16 males and 45 females, from 31 different language backgrounds (Table 1), aged 18-43 years old (M = 23.41 years old, SD = 6.18) were recruited and divided into one of three groups according to their language use. One participant was excluded from the study because the questionnaire was completed incorrectly.

Table 1

Language background frequencies

Language	Number of Users	Language	Number of Users
English	61	Kapampangan	1
French	21	Korean	1
Spanish	13	Lithuanian	1
Italian	5	Mandarin	1
Japanese	4	Norwegian	1
Tagalog	4	Portuguese	1
German	3	Serbian	1
Cantonese	2	Shanghainese	1
Cree	2	Somali	1
Punjabi	2	Twi	1
Russian	2	Vietnamese	1
Stoney Nakoda	2		
Urdu	2		
Arabic	1		
Bengali	1		
Cambodian	1		
Danish	1		
Farsi	1		
Hebrew	1		
Hindi	1		

Participant Groups

Participants who indicated that they were exposed to an L2 less than 5% of the time were placed in the monolingual group, participants who reported being exposed to more than one language from 5-25% of the time were considered inactive bilinguals, and participants who spoke, read, listened, etc. to another language more than 25% of the time were categorized as active bilinguals. There were a total of 19 monolingual participants (Mage = 22.84 years old, SD = 5.44), 22 inactive bilinguals (Mage = 23.68 years old, SD = 5.22), and 20 active bilinguals (Mage = 23.65 years old, SD = 7.88). A one-way ANOVA was run ($F(2,58) = 0.11$, $p = .893$) and found no significant differences between language groups in terms of age.

While on the lower side in terms of participants per group, it is not unusual to have about 20 participants per group. Many studies are conducted with roughly 20-30 people in each language group (Bialystok et al., 2004; Bialystok et al., 2006; de Bruin et al., 2015b; Kan et al., 2014; Kan & Sadagopan, 2014; Kaushanskaya & Marian,

2009a; Kaushanskaya & Rechtzigel, 2012; Pelham & Abrams, 2014; Weber et al., 2015).

Language Switching and Word Substitution

Two separate one-way ANOVAs were conducted comparing the three groups on language switching and word substitution to check for language use. Language switching and word substitution were measured on a scale from rarely (1) to often (6). The groups differed significantly on how often they switched between languages per day ($F(2, 58) = 26.92$, $p < .05$). A Games-Howell post-hoc test was run to compare the three groups, with the active bilingual group ($M = 4.45$, $SD = 1.47$) switching between languages significantly more than the monolingual group ($M = 1.05$, $SD = 1.94$) where $p < .05$ and the inactive bilingual group ($M = 2.82$, $SD = 1.94$, $p = .010$). The inactive bilingual group switched between languages significantly more than the monolingual group at the $p < .05$ level.

The groups also differed significantly on word substitution ($F(2, 58) = 17.93$, $p < .05$). Another Games-Howell post-hoc was run to compare the groups. The monolingual group ($M = 1.37$, $SD = 0.83$) substituted words significantly less often than both the inactive bilingual ($M = 3.27$, $SD = 1.93$) and active bilingual ($M = 4.35$, $SD = 1.66$) groups at the $p < .05$ level; there was no difference between the inactive bilingual and active bilingual groups on word substitution ($p = .14$).

SES

To control for potential SES differences, information about participants' parent's education was gathered. Each participant was asked to indicate their parent's highest education level completed: High School Not Completed (1), High School Graduate (2), Undergrad Completed (3), and Graduate or Professional Degree Obtained (4). Two one-way ANOVAs were run to compare parents' education between the three language groups. There were no significant differences between groups for either father's ($F(2, 58) = 0.132$, $p = .877$) or mother's ($F(2, 58) = 1.34$, $p = .271$) education level. Parents of participants from all three language groups had completed similar levels of education.

Living Outside of Canada and L2 courses

Information about whether participants had lived in a country outside of Canada or had taken a second language course was also gathered, two one-way ANOVAs were run to compare the groups. Groups differed significantly in countries where they had lived ($F(2, 58) = 8.31$, $p < .05$), indicating that one group had significantly more participants that had lived in another country besides Canada. In a Games-Howell post hoc, the active bilingual group ($M = .65$, $SD = .49$) was significantly more likely to have lived in another country than the inactive bilingual ($M = .27$, $SD = .46$, $p < .05$) and the monolingual ($M = .11$, $SD = .32$, $p < .05$) groups. There were no significant differences between the inactive bilingual and monolingual groups ($p = .360$). There were no significant differences between groups when comparing participants' on whether they had taken an L2 course ($F(2, 58) = 0.347$, $p = .708$); indicating that regardless of language use, participants were equally likely to have taken an L2 course.

Materials and Design

The questionnaire measured the participants' use of language and gathered background information as described above. Responses were used to divide the participants into one of three groups: monolingual, inactive bilinguals, and active bilinguals. Following this, participants were asked to complete three executive functioning tasks (Simon, ANT, and WCST) and one novel word learning task.

This study used a computer with PEBL (Psychology Experiment Building Language) and DirectRT software installed to conduct executive functioning and novel word learning tasks. PEBL is a psychology test bank where the researchers used the ANT and WCST. DirectRT was used to measure RTs and accuracy for the Simon task and the novel word learning task. The collected data were analyzed using IBM SPSS version 24 software.

Simon Task

Participants began with the Simon task. The Simon task has participants watching a computer screen while red or green squares appear on either the left or right side of the screen. Participants are asked to press K (a key on the right-hand side of the keyboard) if the square is red and to press S (a key on the left) if the square is green, regardless of the squares' positions on the computer screen. There are a total of 32 trials, 16 congruent and 16 incongruent. When the position of the square on the computer screen is congruent to the position of the key pressed on the keyboard, then participants are expected to have faster RTs when compared to an incongruent trial where the position of the square and key are different. The Simon cost is calculated by comparing the RTs of congruent and incongruent trials. The Simon task is a common test of executive functioning used in bilingualism studies to measure inhibition with bilinguals having a smaller Simon cost than monolinguals, indicating an advantage in inhibition (Bialystok, 2008; Bialystok et al., 2004).

ANT

The next task is the ANT. Similar to the Simon task, the ANT also measures inhibition among other

executive functioning processes. During the ANT, a fixation cross appears on the screen. A star may or may not appear on the screen to indicate whether the stimuli will appear above or below the fixation cross. Afterwards, one arrow or a row of arrows will appear on the screen and the participant is to press the right or the left shift keys to indicate the direction of the central arrow. Like the Simon task, the ANT is made up of congruent and incongruent trials but also includes a neutral trial in which only one arrow is present on the screen. The difference between the neutral, congruent, and incongruent trials was used to compare the three different groups on their ability to pay attention to the central arrow, while ignoring the other arrows. On PEBL, the ANT consists of four blocks of trials. The first block was a practice trial with 24 trials. The other three blocks were experimental trials. There were a total of 144 experimental trials, with a final total of 168 trials in the default ANT parameter on PEBL. Previous studies have shown that bilinguals significantly outperform monolinguals in this task (Costa et al., 2008; Pelham & Abrams, 2014).

WCST

The final executive functioning task that the participants were asked to perform is the WCST. Here the participants' ability to switch between tasks was measured. On the computer screen a card appeared which could be sorted in several different ways (e.g. shape, color, number). The participant had to sort the image on one of its features and received feedback about whether their selection was correct. Participants had to figure out the sorting rule and the sorting criteria changes during the task. There were a total of 64 experimental trials on PEBL using the card sort parameter. Here the difference between switching and non-switching errors was used to determine the participant's proficiency in switching tasks. Switching tasks like the card sort task have been previously used in bilingual studies that have reported a bilingual advantage (Hernández et al., 2013; Prior & MacWhinney, 2010).

Novel Word Learning Task

Finally, participants performed a novel-word learning task. Here they were presented with a novel word (presented in writing and in sound) from Kan et al. (2014), paired with a novel visual stimulus created for this study (see Appendix for examples of novel stimuli) and were given fifteen seconds to become familiar with the stimulus and word. There were sixteen novel word and image pairings for the participant to memorize. Afterwards, participants were tested with trials in which they were presented with the stimulus and asked to pick its paired word. Participants were then presented with a stimulus and were asked to verbally produce its paired word. There were sixteen matching trials and six trials for recall. The accuracy on these trials was used to determine the participant's ability to learn new words in a short period of time. Novel word and novel image association tasks had reported a bilingual advantage in the past (Kan et al., 2014; Kan & Sadagopan, 2014).

Procedure

Upon arrival, participants were briefed by the researcher, signed consent forms, and were guided in filling out a questionnaire to assess their experience, use, and proficiency with languages. Participants were seated in front of a computer and received instructions on the tasks that they were asked to perform, beginning with the Simon task. They then completed the ANT followed by the WCST and finished with a novel word learning task. After completing the matching trials, participants were told to inform the researcher that they have reached the recall phase. The researcher then returned to the room and asked the participants to verbally recall the associated word when presented with its stimulus. The researcher informed participants to take their time and recall the word as accurately as possible before pressing the space bar to move on to the next word. The researcher recorded the participants' responses. Finally, participants received a debriefing form and were asked if they had any questions about the study; afterwards, the participants were free to leave the lab and were given class credit for their participation.

Results

Simon Task

The Simon task was analyzed using a 3x2 mixed model General Linear Model, with language group (monolingual, inactive bilingual, active bilingual) as the between-groups variable and trial type (congruent vs. incongruent) as the within-groups variable. RT and accuracy (number of correct responses) were analyzed separately.

RT regression analysis was carried out using a subtraction method by Fan et al. (2002) and Di Francesco et al. (2016) in which participants' congruent RT scores were subtracted from their incongruent scores. The same method was carried out with participants' accuracy scores, subtracting incongruent scores from the congruent.

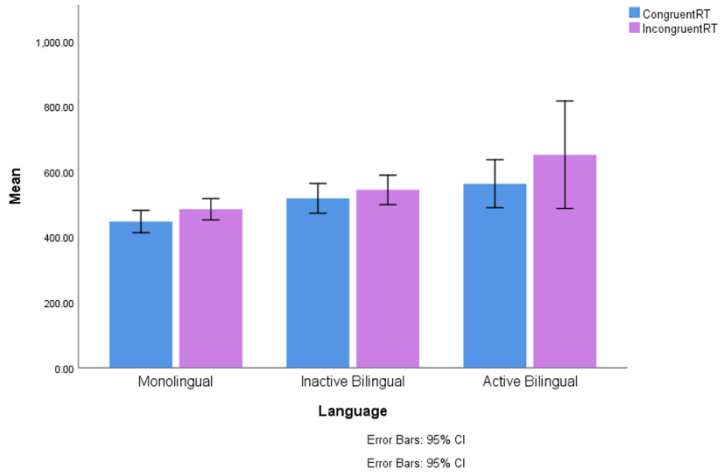
RT

There were significant differences between the three groups RTs in milliseconds ($F(2, 58) = 4.50$, $p < .05$,

ηp2 = .134) (see Tables 2 and 3); a Games-Howell test was run and the monolingual group (M = 468.91, SD = 60.35) produced significantly faster RTs than both the inactive (M = 533.40, SD = 97.82) and active bilingual (M = 613.38, SD = 22.34) groups when $p < .05$, there were no significant differences between the inactive and active bilingual groups ($p = .367$). The monolingual group demonstrated significantly faster RTs in the Simon task than both bilingual groups. There was a significant main effect of trial type ($F(1, 58) = 4.92$, $p < .05$, ηp2 = 0.78), with faster RTs in congruent trials (M = 511.34, SD = 123.62) than incongruent trials (M = 561.46, SD = 221.17); indicating that participants had faster RTs when the image on the screen appeared on the same side as the key they pressed. There was no significant interaction between language and trial condition ($F(2, 58) = 0.712$, $p = .495$), indicating that language use had no effect on performance across congruency (see Figure 1).

Figure 1

Means Comparing the RTs of the Three Language Groups from the Simon Task



Accuracy

There were no significant differences between groups accuracy scores (see Table 3). There was no significant interaction between language groups and trial type ($F(2, 58) = 1.46$, $p = .241$). There was a main effect of trial type ($F(1, 58) = 4.77$, $p < .05$, ηp2 = .076), participants produced more accurate responses on the congruent trials than incongruent trials (see Table 2).

Regression

The linear regression predicting RT scores from L1 exposure was insignificant (see Table 4). Therefore, participants L1 exposure does not predict their RT scores in the Simon task. Another linear regression was run to predict accuracy scores using L1 exposure and found a

Table 2

Means and standard deviations (in parenthesis) of scores from the Simon, ANT and WCST

	Monolinguals	Inactive Bilinguals	Active Bilinguals
Simon			
RT			
Congruent	447.70 (71.07)	518.81 (102.99)	563.60 (157.37)
Incongruent	485.40 (67.55)	544.74 (101.97)	652.11 (351.99)
Accuracy			
Congruent	.98 (.03)	.95 (.08)	.94 (.16)
Incongruent	.97 (.03)	.94 (.07)	.89 (.21)
ANT			
RT			
Incongruent	638.77 (112.54)	665.14 (132.16)	747.24 (281.42)
Neutral	543.60 (94.85)	574.95 (115.36)	648.47 (284.88)
Congruent	631.78 (97.98)	667.72 (159.94)	731.99 (287.10)
Accuracy			
Incongruent	.95 (.07)	.91 (.13)	.85 (.27)
Neutral	.97 (.06)	.97 (.06)	.90 (.25)
Congruent	.95 (.06)	.91 (.14)	.85 (.28)
WCST			
Preservative Errors	.15 (.09)	.12 (.05)	.17 (.09)
Non-Preservative Errors	.11 (.08)	.09 (.07)	.11 (.07)

Note. Accuracy and error scores were converted to error rates.

Table 3

Categorical results from the Simon, ANT, WCST, and Novel Word Learning Task

	F Scores	*p* Values	Effect Size	Power
Simon				
RT	4.50*	.015	.134	.748
Accuracy	1.44	.244	.048	.297
ANT				
RT	1.59	.212	.052	.324
Accuracy	1.52	.227	.050	.311
WCST				
Accuracy	1.56	.218	.051	.319
Novel Word Learning				
Matching	1.84	.075	.060	.369
Recall				
Correct	0.038	.963	.001	.055
Incorrect	0.411	.665	.014	.113
Partially Correct	1.073	.349	.036	.229

Note. * indicates the results listed above were significant at the $p < .05$ level.

Table 4

Regression results from the Simon, ANT, WCST, and Novel Word Learning Tasks

	Regression	*p* Values
Simon		
RT Incongruent-neutral	1.99	.164
Accuracy Incongruent-neutral	4.00*	.050
ANT		
RT		
Incongruent-neutral	0.02	.90
Incongruent-congruent	0.28	.60
Accuracy		
Neutral-incongruent	0.121	.729
Neutral-congruent	0.098	.756
WCST		
Errors	3.29	.075
Novel Word Learning		
Matching	0.675	.415
Recall		
Correct	0.003	.959
Incorrect	0.816	.370
Partially Correct	1.57	.216

Note. * indicates the results that were significant at the $p < .05$ level.

significant linear regression F(1, 59) = 4.00, p = .05, R2 = .063. Participants' accuracy scores decreased by -0.001 for each increase in L1 exposure. Indicating that the participants who frequently used L1 more often tended to have lower accuracy scores, supporting results from the dichotomized analysis. is suggesting a bilingual disadvantage (see Table 4).

ANT

The ANT was analyzed with a 3x3 mixed General Linear Model with language groups serving as the between-group variable and trial type (neutral, incongruent, congruent) as the within-groups variable. RT and accuracy (number of errors) were analyzed separately.

For the regression analyses, RTs and accuracy were analyzed using the same subtraction method from the Simon task; RT scores were generated by subtracting neutral and congruent scores from the incongruent scores, accuracy scores were created by subtracting congruent and incongruent scores from the neutral scores.

None of the ANT results were significant. See Table 3 for the categorical analyses and Table 4 for the regression results.

WCST

Accuracy

A 3x2 mixed General Linear Model was run to compare participants WCST scores. Language groups served as the between-groups variable and types of errors served as the within-groups variable, comparing the number of perseverative errors (errors that occurred when the sorting rule changed) with the number of non-perseverative errors (random errors).

The results were insignificant (see Table 3).

Regression

The linear regression performed to predict the effect of language usage on the percentage of perseverative errors was insignificant (see Table 4).

Novel Word Learning

Matching and recall conditions from the Novel Word Learning task were analyzed separately. A one-way ANOVA was run comparing the three language groups on number of correct answers in the matching condition.

Both the ANOVA and the regression in the matching condition were insignificant (see Tables 3 and 4).

Recall

Accuracy. For the recall condition, three separate one-way ANOVAs were run to compare groups on the number of correct, incorrect, and partially correct (participants were able to recall part of the word) novel words they were able to recall. (see Table 3).

Regression. The regression analysis was insignificant (see Table 4).

Discussion

The purpose of the study was to gain a better understanding of the effects of bilingual language use on cognitive function and word learning. Participants were asked to perform tasks that had previously reported a bilingual advantage, but we were unable to find evidence to support a bilingual advantage. We had hypothesized that a bilingual who regularly used more than one language would show a stronger bilingual advantage, but we found no evidence of an effect of active bilingualism either.

Inhibition

In the Simon task, we found a significant difference in the RTs between different language groups. However, our active bilingual group had the slowest RTs, which contradicts our hypothesis that there is an effect of

active bilingualism, indicating that increased language use negatively affected inhibition. Our results also contradict previous studies in which bilinguals were reported to demonstrate faster RTs on incongruent trials when compared to monolinguals (Bialystok, 2008; Bialystok et al., 2004; Verreyt et al., 2015), suggesting that monolinguals may have an advantage in the Simon task. Contrary to previous studies, our bilinguals did not outperform the monolingual group on the ANT either (Costa et al. 2008; Pelham & Abrams, 2014) possibly due to a number of variables that were not controlled for such as language proficiency, language pairs, acculturation, etc. While bilinguals have been reported to show an advantage in inhibition, this would not be the first study in which a bilingual advantage in inhibition was not found (de Bruin et al., 2015b; Incera & McLennan, 2018; Kousaie et al., 2014; Paap & Greenberg, 2013). Paap and Greenberg (2013) also found a bilingual disadvantage with inhibition tasks.

It has been previously hypothesized that bilinguals have an advantage in inhibition because they have practice with inhibiting another language while speaking (de Bruin et al., 2015b). In the case of our active bilinguals, since they regularly use more than one language, they are more likely to have practice with inhibition than the inactive bilingual or monolingual groups. Research has also suggested that the inhibition practiced by bilinguals inhibiting another language may not transfer as an advantage in inhibition tasks that do not involve languages (de Bruin et al., 2015b). We propose another explanation that active bilinguals who must regularly inhibit another language may use up more cognitive resources during inhibition which would slow them down during tasks, leading to slower RTs and more errors.

Task-Switching

There was no evidence of a bilingual advantage in task switching either as demonstrated on the WCST task. The active bilingual group, the group that we hypothesized would have the most practice with switching, did not make the fewest errors out of all the groups. Contrary to previous research, the bilingual group did not outperform the monolingual group (Hernández et al., 2013; Prior & MacWhinney, 2010; Vega & Fernandez, 2011; Xie & Dong, 2017). As mentioned above, the WCST measures several different executive functions (Gamboz et al., 2009), not just task switching, so perhaps using a different task, one that only measures task switching, would be more sensitive to finding differences between groups.

Research has also been split on the existence of a bilingual advantage on task-switching and that experience with switching between languages may not result in advantages in tasks designed to test task-switching (de Bruin et al., 2015b).

Novel Word Learning

No significant differences between groups were found in the novel word learning task either. The three groups did not differ in their ability to match or recall novel words when presented with their paired novel images (Kan & Sadagopan, 2014; Kan et al., 2014; Kaushanskaya & Marian, 2009b). While previous research has credited bilinguals with an advantage due to their experience with different language systems, bilinguals may have simply learned strategies that would help them learn new words in a short period of time (Kan & Sadagopan, 2014). Kan et al. (2014) found that besides language experience, strategies such as speech practice also helped with the retention of novel words.

The novel word learning task was also the last task that our participants were asked to complete. At that point, participants may have been experiencing fatigue which prevented them from performing at an optimal level. Floor effects have been shown in previous studies in which participants have been asked to verbally recall novel words, therefore it is possible that the task was too difficult to find any effects of bilingualism (Kan & Sadagopan, 2014; Kaushanskaya, 2012).

Previous research findings also indicate that bilinguals have more difficulty with picture-naming and verbal recall tasks, due to their experience with multiple language systems (Biaylstok, 2008). Using a different novel word learning task may lead to different results. Studies have suggested that when testing adult bilinguals, using tasks that are similar to classroom learning tasks may help facilitate novel word learning such as, matching novel words with translations or with pictures of familiar objects (Kaushanskaya et al., 2013; Nair et al., 2016).

Overall Findings

Taken together we had found no evidence of a bilingual advantage in any tasks (Simon, ANT, WCST, novel word learning), and in the case of the Simon task, we found a bilingual disadvantage. Previous research found bilingual advantages in areas of inhibition, task switching, and novel word learning (Bialystok et al., 2004; Bialystok, 2008; Costa et al., 2008; Costa et al., 2009; Hernández et al., 2013; Kan et al., 2014; Kan & Sadagopan, 2014; Kaushanskaya et al., 2013; Nair et al., 2016; Pelham & Abrams, 2014; Prior & MacWhinney, 2010; Soveri et al., 2011; Vega & Fernandez, 2011; Verreyt et al., 2015; Xie & Dong, 2017) that we were unable to replicate in this study, instead adding to the number of studies that found no advantage associated with bilingualism (de Bruin et al., 2015b; Incera & McLennan,

2018; Kousaie et al., 2014; Paap & Greenberg, 2013). As the bilingual experience is highly variable between individuals, there are a number of differences from participant to participant that make it difficult to come to a conclusion that generalizes to various populations.

We conducted our study using Introductory Psychology students, the majority of which are young adults (M = 23.41 years old, SD = 6.18). Studies have previously found that the effects of bilingualism are most evident in young children or older adult participants (de Bruin et al., 2015b), so we may find effects if we were to test a different age sample. It is also possible that a larger sample size may have led to different results, as the study was conducted with approximately 20 participants per group, which may not have been big enough to detect any effects. As shown in Table 3, the power and effect sizes for the majority of tasks were low. The combination of a small effect and low power could have led to an inability to detect any differences between participant groups.

Limitations

As mentioned in the introduction, there are several confounding variables that have been associated with bilingualism and we were unable to control for some such as language pairings, proficiency, and immigration.

Although our study broadly examines the effect of bilingualism across a wide variety of languages, it has been hypothesized that different language groups may exhibit different results, which is why some studies try to control for languages by testing specific groups of bilinguals (de Bruin et al., 2015b). With a wide variety of languages any effect that may exist could have been harder to detect. When comparing groups on whether they had lived outside of Canada, the active bilingual group had significantly more participants who previously lived in other countries. de Bruin et al. (2015b) listed immigration as a factor that should be controlled for in bilingualism studies as any number of factors from immigration could create confounds such as differences in SES, cultural influences, health, individual adaptation, language environment, cultural differences, and immersion (Costa et al., 2009; Kaushanskaya et al., 2013; Paap & Greenberg, 2013). Due to time and sample restrictions, we were unable to control for immigration or acculturation effects.

Another possible limitation is the lack of control for proficiency. Cummins (1976) suggested that inconsistencies in bilingualism literature are due to differing levels of language proficiency in bilingual participants. He hypothesized that for a bilingual participant to show a cognitive advantage, they had to have reached a certain threshold in proficiency in one language before learning another; if this threshold is not reached, then the participant will show disadvantages in cognitive functioning (Cummins, 1976). It is possible that inactive bilinguals who do not regularly use more than one language may have lost proficiency in the unused language and no longer reach the threshold to show cognitive advantages. For our study, we relied on self-reports about language proficiency, but it was not tested or controlled for. Bilinguals we happened to test may not have met the proficiency threshold necessary to gain advantages related to bilingualism. However, this is less likely an issue for our active bilingual group.

Finally, groups were determined based on self-reports, participants who ranked their L2 use from 5-25% were classified as inactive bilinguals. It is possible that it was too broad a range since there was great variability in language use within the group. It is possible that being exposed to another language at least 25% of the time was not enough to be considered an active bilingual, but it was used as the cut-off point due to the small sample size. Studies have also suggested using both self-reports and standardized tests to gain a more accurate picture of individual language experience (Luk & Bialystok, 2013; Pelham & Abrams, 2014). For the questionnaire, participants were asked to indicate their language exposure in all situations for one question. A better picture of participants language use and exposure could have possibly been captured by breaking the exposure down into multiple questions, for example separate questions about how many languages a participant uses in different environments (e.g. school, home, etc.) and separate questions about time spent reading, speaking, or listening, etc. (Marian et al., 2007). Again, this limitation is less likely to have an impact on our active bilingual group.

Future Research

de Bruin et al. (2015b) and Paap and Greenberg (2013) discussed how there are a wide variety of tasks that have been used to test various areas of executive function, but there is little correlation between the tasks. Possibly, one of the reasons for the discrepancies within the literature is because the different tasks may be measuring different aspects of cognitive functioning. According to Paap and Greenberg (2013), between various types of inhibition tasks, there was only a r = -.01 correlation, indicating that, while one task might find a bilingual advantage, it is highly likely that another task, designed to test the same aspect of executive function, will not. Paap and Greenberg (2013) suggested that it is more helpful to use multiple tests of one measure rather than just relying on one test to draw conclusions.

While we did not find an advantage of language use from the tasks that we used, we cannot rule out the

possibility that other tasks might. Executive function is an umbrella term for many different types of processes (de Bruin et al., 2015b), so a failure to find support for a bilingual advantage in certain areas does not entirely rule out the possibility of a bilingual advantage. As research on executive functioning continues, new and perhaps more reliable tests may be produced.

Likewise, future research can investigate other ways of examining language use. For our study, we relied on participants' estimates of their exposure to their second language (e.g. spoke, listened, etc.). Other ways of grouping participants can be explored in the future such as examining bilinguals' language switches in different contexts (e.g. work and home). Perhaps more research can examine treating bilingualism as a continuous variable rather than dichotomizing it.

Conclusion

Our study found no evidence supporting a bilingual advantage or an effect of active bilingualism. Interestingly, we found the opposite result in which bilinguals with more language exposure seemed to perform worse on inhibition tasks. On task switching and novel word learning monolinguals and both bilingual groups performed similarly, indicating no advantage of bilingualism. Further research on the effect of language use could help clarify some of the discrepancies in the research literature.

References

Bialystok, E. (2008). Bilingualism: The good, the bad, and the indifferent. Bilingualism: Language and Cognition, 12(1), 3-11. https://doi.org/10.1017/S1366728908003477

Bialystok, E., Craik, F. I., & Luk, G. (2012). Bilingualism: Consequences for mind and brain. Trends in Cognitive Sciences, 16(4), 240-250. https://doi.org/10.1016/j.tics.2012.03.001

Bialystok, E., Craik, F. I., Klein, R., & Viswanathan, M. (2004). Bilingualism, aging, and cognitive control: Evidence from the Simon task. Psychology and Aging, 19(2), 290-303. https://doi.org/10.1037/0882-7974.19.2.290

Bialystok, E., Craik, F. I., & Ruocco, A. C. (2006). Dual-modality monitoring in a classification task: The effects of bilingualism and ageing. Quarterly Journal of Experimental Psychology, 59(11), 1968-1983. https://doi.org/10.1080/17470210500482955

Bialystok, E., & Martin, M. M. (2004). Attention and inhibition in bilingual children: Evidence from the dimensional change card sort task. Developmental Science, 7(3), 325-339. https://doi.org/10.1111/j.1467-7687.2004.00351.x

Colzato, L. S., Bajo, M. T., van den Wildenberg, W., Paolieri, D., Nieuwenhuis, S., La Heij, W., & Hommel, B. (2008). How does bilingualism improve executive control? A comparison of active and reactive inhibition mechanisms. Journal of Experimental Psychology: Learning, Memory, and Cognition, 34(2), 302-312. https://doi.org/10.1037/0278-7393.34.2.302

Costa, A., Hernández, M., & Sebastián-Gallés, N. (2008). Bilingualism aids conflict resolution: Evidence from the ANT task. Cognition, 106(1), 59-86. https://doi.org/10.1016/j.cognition.2006.12.013

Costa, A., Hernández, M., Costa-Faidella, J., & Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: Now you see it, now you don't. Cognition, 113(2), 135-149. https://doi.org/10.1016/j.cognition.2009.08.001

Cummins, J. (1976). The influence of bilingualism on cognitive growth: A synthesis of research findings and explanatory hypotheses. Working Papers on Bilingualism, 9, 1-44.

de Bruin, A., Treccani, B., & Della Sala, S. (2015a). Cognitive advantage in bilingualism: An example of publication bias? Psychological Science, 26(1), 99-107. https://doi.org/10.1177/0956797614557866

de Bruin, A., Bak, T. H., & Della Sala, S. (2015b). Examining the effects of active versus inactive bilingualism on executive control in a carefully matched non-immigrant sample. Journal of Memory and Language, 85, 15-26. https://doi.org/10.1016/j.jml.2015.07.001

Di Francesco, S. A., Simione, L., López-Ramón, M. F., Belardinelli, M. O., Lupiáñez, J., & Raffone, A. (2016). Dispositional mindfulness facets predict the efficiency of attentional networks. Mindfulness, 8(1), 101-109. https://doi.org/10.1007/s12671-016-0634-5

Fan, J., McCandliss, B. D., Sommer, T., Raz, A., & Posner, M. I. (2002). Testing the efficiency and independence of attentional networks. Journal of Cognitive Neuroscience, 14(3), 340-347. https://doi.org/10.1162/089892902317361886

Gamboz, N., Borella, E., & Brandimonte, M. A. (2009). The role of switching, inhibition and working memory in older adults' performance in the

Wisconsin Card Sorting Test. Aging, Neuropsychology, and Cognition, 16(3), 260-284. https://doi.org/10.1080/13825580802573045

Hernández, M., Martin, C. D., Barceló, F., & Costa, A. (2013). Where is the bilingual advantage in task-switching? Journal of Memory and Language, 69(3), 257-276. https://doi.org/10.1016/j.jml.2013.06.004

Incera, S., & McLennan, C. T. (2018). Bilingualism and age are continuous variables that influence executive function. Aging, Neuropsychology, and Cognition, 25(3), 443-463, http://doi.org/10.1080/13825585.2017.1319902

Kan, P. F., & Sadagopan, N. (2014). Novel word retention in bilingual and monolingual speakers. Frontiers in Psychology, 5, 1-10. https://doi.org/10.3389/fpsyg.2014.01024

Kan, P. F., Sadagopan, N., Janich, L., & Andrade, M. (2014). Effects of speech practice on fast mapping in monolingual and bilingual speakers. Journal of Speech, Language, and Hearing Research, 57(3), 929-941. https://doi.org/10.1044/2013_JSLHR-L-13-0045

Kaushanskaya, M. (2012). Cognitive mechanisms of word learning in bilingual and monolingual adults: The role of phonological memory. Bilingualism: Language and Cognition, 15(3), 470-489. https://doi.org/10.1017/S1366728911000472

Kaushanskaya, M., & Marian, V. (2009a). Bilingualism reduces native-language interference during novel-word learning. Journal of Experimental Psychology: Learning, Memory, and Cognition, 35(3), 829-839. http://doi.org/10.1037/a0015275

Kaushanskaya, M., & Marian, V. (2009b). The bilingual advantage in novel word learning. Psychonomic Bulletin & Review, 16(4), 705-710. http://doi.org/10.3758/PBR.16.4.705

Kaushanskaya, M., & Rechtzigel, K. (2012). Concreteness effects in bilingual and monolingual word learning. Psychonomic Bulletin & Review, 19(5), 935-941. https://doi.org/10.3758/s13423-012-0271-5

Kaushanskaya, M., Yoo, J., & Van Hecke, S. (2013). Word learning in adults with second-language experience: Effects of phonological and referent familiarity. Journal of Speech, Language, and Hearing Research, 56(2), 667-678. https://doi.org/10.1044/1092-4388(2012/11-0084)

Kousaie, S., Sheppard, C., Lemieux, M., Monetta, L., & Taler, V. (2014). Executive function and bilingualism in young and older adults. Frontiers in Behavioral Neuroscience, 8, 250-262. https://doi.org/10.3389/fnbeh.2014.00250

Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. Journal of Cognitive Psychology, 25(5), 605-621. https://doi.org/10.1080/20445911.2013.795574

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. Journal of Speech, Language, and Hearing Research, 50(4), 940-967. https://doi.org/10.1044/1092-4388(2007/067)

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. Cognitive psychology, 41(1), 49-100. https://doi.org/10.1006/cogp.1999.0734

Monchi, O., Petrides, M., Petre, V., Worsley, K., & Dagher, A. (2001). Wisconsin Card Sorting revisited: distinct neural circuits participating in different stages of the task identified by event-related functional magnetic resonance imaging. Journal of Neuroscience, 21(19), 7733-7741. https://doi.org/10.1523/JNEUROSCI.21-19-07733.2001

Nair, V. K., Biedermann, B., & Nickels, L. (2016). Consequences of late bilingualism for novel word learning: Evidence from Tamil–English bilingual speakers. International Journal of Bilingualism, 20(4), 473-487. https://doi.org/10.1177/1367006914567005

Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. Cognitive Psychology, 66, 232–258. https://doi.org/10.1016/j.cogpsych.2012.12.002

Pelham, S. D., & Abrams, L. (2014). Cognitive advantages and disadvantages in early and late bilinguals. Journal of Experimental Psychology: Learning,

Memory, and Cognition, 40(2), 313-325. https://doi.org/10.1037/a0035224

Prior, A., & MacWhinney, B. (2010). A bilingual advantage in task switching. Bilingualism: Language and Cognition, 13(2), 253-262. https://doi.org/10.1017/S1366728909990526

Soveri, A., Rodriguez-Fornells, A., & Laine, M. (2011). Is there a relationship between language switching and executive functions in bilingualism? Introducing a within group analysis approach. Frontiers in Psychology, 2, 1-8. https://doi.org/10.3389/fpsyg.2011.00183

Treccani, B., Argyri, E., Sorace, A., & Della Sala, S. (2009). Spatial negative priming in bilingualism. Psychonomic Bulletin & Review, 16(2), 320-327. https://doi.org/10.3758/PBR.16.2.320

Vega, C., & Fernandez, M. (2011). Errors on the WCST correlate with language proficiency scores in Spanish-English bilingual children. Archives of clinical neuropsychology, 26(2), 158-164. https://doi.org/10.1093/arclin/acq097

Vega-Mendoza, M., West, H., Sorace, A., & Bak, T. H. (2015). The impact of late, non-balanced bilingualism on cognitive performance. Cognition, 137, 40-46. https://doi.org/10.1016/j.cognition.2014.12.008

Verreyt, N., Woumans, E., Vandelanotte, D., Szmalec, A., & Duyck, W. (2015). The influence of language-switching experience on the bilingual executive control advantage. Bilingualism: Language and Cognition, 19(1), 181-190. https://doi.org/10.1017/S1366728914000352

Weber, R. C., Johnson, A., Riccio, C. A., & Liew, J. (2015). Balanced bilingualism and executive functioning in children. Bilingualism: Language and Cognition, 19(2), 425-431. https://doi.org/10.1017/S1366728915000553,

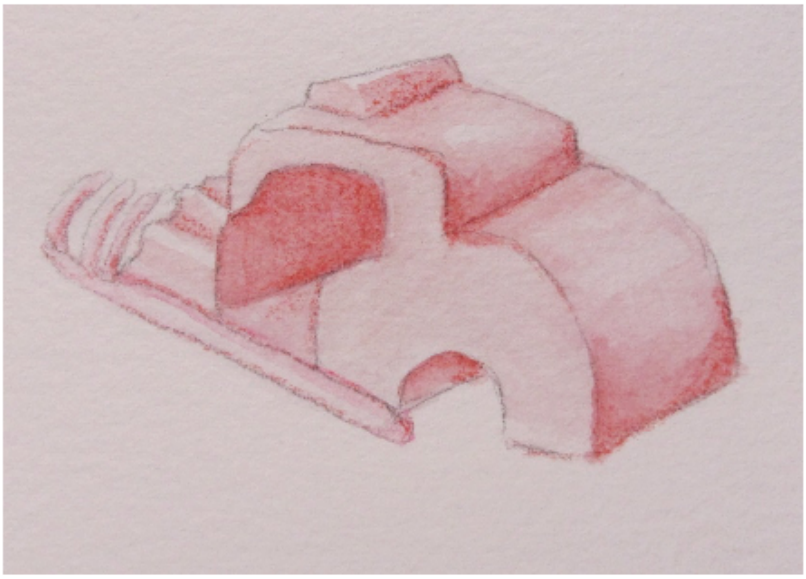
Xie, Z., & Dong, Y. (2017). Contributions of bilingualism and public speaking training to cognitive control differences among young adults. Bilingualism: Language and Cognition, 20(1), 55-68. https://doi.org/10.1017/S1366728915000474

Author Note

We would like to thank all of our participants and Amelia Vermeulen for creating the novel image stimuli. Parts of these findings were presented as a poster at the 2019 Scandinavian Association for Language and Cognition, Aarhus, Denmark; and at the 2019 International Symposium of Bilingualism, Edmonton, Alberta, Canada.

Appendix

Examples of Novel Visual Stimuli



Journal of Psychological Inquiry
2021, Vol.25, No. 2, pp. 37-45

Attachment, parental divorce, and religiosity: Exploring predictors of sociosexuality

Krista S. Zenk, Caitlin M. Makovicka, and Shannon E. Claxton

Morningside College

Abstract – A well-established relationship between romantic attachment and sociosexual orientation exists, and we aimed to examine factors that may predict this relationship. This study was cross-sectional and included 741 participants with a mean age of 22.26 (SD = 3.51). Out of the participants, 70% were female and 83.7% were White. The purpose of this research was to better understand how parental divorce and religiosity predict romantic and sociosexual orientation. Sociosexuality is defined as an individual's willingness to engage in casual sexual relationships; those with higher scores show a more unrestricted orientation. We found evidence of a weak, positive relationship between anxious and avoidant attachment and sociosexual orientation, such that those with higher attachment anxiety and attachment demonstrated more unrestricted sociosexual orientations. There was a moderate negative association between religiosity and sociosexual orientation, such that higher religiosity was associated with a more restricted sociosexual orientation. Divorce and sociosexual orientation were not related. Overall, romantic attachment (in particular avoidant attachment) and religiosity emerged as the strongest predictors of an unrestricted sociosexual orientation.

Keywords: sociosexual orientation, romantic attachment, parental divorce, religiosity

Researchers suggest that most young adults engage in at least one casual sexual relationship or experience (e.g., Claxton & van Dulmen, 2013). Sociosexual orientation, which is one's willingness to engage in sexual actions outside of a committed relationship may be a driving factor in participation in casual sexual relationships (Penke & Asendorpf, 2008). Researchers have found associations between sociosexual orientation and romantic attachment (e.g., Brase et al., 2014). Similarly, researchers suggest that parental divorce (e.g., Haddad et al., 2016) and religiosity (e.g., Simmons et al., 2009) may be important factors to consider in understanding sociosexual orientation because they are prominent in shaping the experiences of an individual such as expectations in romantic relationships. Although researchers suggest that there is a relationship between romantic attachment, parental divorce, religiosity, and sociosexual orientation, this relationship is not fully understood. The purpose of this paper was to extend research on sociosexaul orientation and to evaluate which factors best predict sociosexual orientation.

Sociosexual Orientation

Sociosexual orientation is an individual's willingness to engage in casual sexual relationships. The first measure of sociosexual orientation was uni-dimensional (see Gangestad & Simpson, 1990) and ranged from unrestricted to restricted. Individuals with a more restricted sociosexual orientation focus on long-term, committed relationships, whereas those with unrestricted sociosexual orientations are likely to participate in short-term, uncommitted relationships, such as one-night stands (Gangestad & Simpson, 1990). However, a more recent Revised Sociosexual Orientation Inventory (SOI) consists of multiple dimensions including prior behavior, opinion on casual sexual relationships, and desire for non-committed relationships (Penke & Asendorpf, 2008).

One's sociosexual orientation predicts their psychological and physical wellbeing. For example, having a stable relationship with a romantic partner promotes pro-relationship behaviors, such as long-term commitment and is associated with decreased infidelity

in romantic relationships (Rodrigues & Lopes, 2017). Another study concluded that people who are in long-term committed relationships tend to be happier compared to those who are single (Proulx & Snyder-Rivas, 2013). In addition, having unrestricted sexual relationships can be problematic for an individual. Some researchers suggest that engaging in casual sexual relationships is related to the psychological well-being of women, including higher psychological distress and increased alcohol or drug consumption (Bersamin et al., 2014; Dube et al., 2017). Given the potential downsides of having an unrestricted sociosexual orientation, it is important to understand the factors that may predict this type of orientation.

Romantic Attachment

One factor linked to sociosexuality is an individual's romantic attachment style, which is how an individual tends to interact and behave with romantic partners (Fraley & Shaver, 2000). Attachment styles originated from Bowlby's (1969, 1989) parental attachment theory. Bowlby suggested that relationships between infants and their caregivers affect their expectations for future relationships. Research supports the idea that attachment styles are broad and have relevance across the lifespan (Sutton, 2019). Romantic attachment style impacts the connection, quality, and longevity of the romantic and sexual relationships an individual has. These romantic attachments vary on the dimensions of attachment anxiety and attachment avoidance (e.g., Brennan et al., 1998), and these dimensions relate to sociosexuality.

Anxiously attached individuals desire emotional intimacy but fear losing their partners (Tracy et al., 2003). Individuals with this attachment style tend to have inconsistent relationship patterns including habits such as breaking up and getting back together, extreme jealousy, and obsessions over romantic partners (Collins et al., 1996). They are at risk for engaging in unwanted but consensual relationships (including casual sexual relationships) and risky sexual behavior because they rush into sexual behavior as a way to increase intimacy with a partner (Brase et al., 2014; Feeney et al., 1999; Gentzler & Kerns, 2004; Tracy et al., 2003). Overall, anxiously attached individuals may report permissive attitudes but actually crave committed relationships (Simon, 1997).

On the other hand, individuals who are avoidantly attached to their romantic partners tend to be colder and have a more self-focused approach to relationships (Schmitt & Jonason, 2015). Individuals with this attachment style tend to have dismissing and fearful relationship patterns, including spending less time grieving breakups, being emotionally withdrawn from their partners, and being distrustful of long-term relationships (Collins et al., 1996). This can lead to unrestricted sociosexual orientations because they prefer to engage in casual sex without emotional involvement (Brase et al., 2014; Gentzler & Kerns, 2004; Schmitt & Jonason, 2015; Simpson et al., 2004).

Lastly, individuals can also experience secure attachments within their romantic relationships, meaning that they are low on both attachment avoidance and attachment anxiety. Studies have found that individuals with secure attachments tend to be more invested within their relationships, supportive of their partners in times of need, and consistently seek mutually beneficial relationships (Collins et al., 1996). Those with secure attachments are less likely to express desire for casual sexual relationships because of the strong emotional connection and high levels of trust in the relationship (e.g., Simmons et al., 2009; Simon, 1997). In sum, romantic attachment anxiety and attachment avoidance relate to behaviors and habits in romantic relationships which in turn are related to sociosexual orientation.

The Role of Parental Divorce

It has been established that parental divorce can have negative effects on children because of the stress and trauma of a break-up (Sutton, 2019). Social learning theory provides one explanation of how parental divorce impacts later functioning in romantic relationships. This theory says that children learn through observation to better understand how the world works. Children who observed a divorce between their parents are less willing to be in a long-term committed relationship because their experience has shaped their views on relationships (Cui &, Finchman, 2010). In addition, some children from divorced families struggle in and have lower quality adult romantic relationships (Washington & Hans, 2013).

Additionally, parental divorce may be related to sociosexuality. One study found that children from divorced families are more sexually permissive than those whose parents are still married (Simmons et al., 2009). Similarly, Jónsson et al., (2000) concluded that adults whose parents were divorced engaged in sexual experiences earlier than their peers. However, the opposite has been found in other studies. Fraley and Heffernan (2013) suggested that children who witness their parents' divorce try to not make the same mistake as their parents, by taking relationships more seriously. This could create a more restricted sociosexual orientation because these children strive for more support and a more stable relationship. Overall, there is some evidence

that divorce may be related to sociosexual orientation, although the relationship has not always been found.

The Role of Religiosity

Religiosity is also a predictor of an individual's sociosexual orientation. It has been established by researchers that religiosity is significantly negatively correlated with sociosexuality (Haddad et al., 2016). That means if an individual has low levels of religiosity, he or she is more likely to engage in more unrestricted sociosexual behaviors, attitudes, and desires. Furthermore, researchers have also found that religiosity impedes the number of casual sexual partners an individual has (Barkan, 2006). In addition, Moon et al. (2019) showed that the sociosexual orientation of an individual is related to how the individual thinks and feels about his or her religion. This means that the mating strategies of some individuals tend to connect in one direction or the other to their attraction or repulsion towards their religion.

An aspect of religiosity that is associated with the individual's sociosexuality is the individual's religious service attendance. As individuals attend more religious services, they become more aware of the expectations that their religion has for engaging in casual sexual relationships, so individuals may tend to try to abstain from casual sexual relationships because of it (Rostosky et al., 2004). As shown, previous researchers have detected a relationship between religiosity and sociosexuality.

Control Variables

Many studies have noted that gender is related to sociosexual orientation. In particular, researchers suggest that men have more unrestricted sex when compared to women (e.g., Schmitt, & Jonason, 2015). Similarly, past researchers have shown that gender can have an impact on romantic attachment. Researchers have found that men tend to be higher in avoidant attachment styles and lower in anxiety attachment styles than women (Del Giudice, 2011). Not only have previous studies shown how gender plays a role in sociosexual orientation and romantic attachment, but previous researchers have shown that gender impacts the relationship between sociosexual orientation and romantic attachment. Therefore, within this study we controlled for the sex of our participants.

Also, researchers have shown that age may be associated with sociosexuality, in that younger adults report more unrestricted sociosexual orientations than older adults (Le Gall et al., 2002). Additionally, it is common for individuals to increase in their willingness to engage in casual sexual relationships when they transition from adolescence to emerging adulthood (Lyons et al., 2015). This is relevant because the sample used in the current study contains a larger age range than many previous studies.

Literature Advancement

Previous research found relationships between the variables of religion, parental divorce, and attachment, and sociosexulaity, but these factors have not been studied together systematically. Understanding how these factors predict sociosexuality when examined together provides information that can be used in a clinical setting to help identify individuals at risk of experiencing negative outcomes associated with unrestricted sociosexuality.

Another important aspect to note about previous research is that the data is usually from a limited and smaller sample size from mainly college-aged students (18-22). Thus, the current study aims to expand the evaluation of attachment, parental divorce, religiosity, and sociosexual orientation to non-college students and those in their mid-to-late 20s.

Hypotheses

Attachment

We hypothesized that there would be an association between romantic attachment and sociosexual orientation. Based on previous research (see Brase et al., 2014 for a summary) we expected that individuals who were anxiously attached to their romantic partners would have higher scores on the sociosexual orientation inventory (SOI) indicating a more unrestricted sociosexual orientation (H1). Similarly, we expected that those with higher levels of avoidant attachment would report higher SOI scores (H2).

Divorce

We also hypothesized that parental divorce would have an association with a higher SOI score (H3).

Religiosity

In addition, we expected that individuals who had higher religiosity scores would have lower SOI scores than those who had lower religiosity scores (H4) based on research done by Barkan (2006). Additionally we expect that in a multiple linear regression, religion, divorce, and attachment will all predict SOI (H5).

Method

Participants

Prior to conducting this study, we received approval from the institution's institutional review board. A subsample of participants from a larger online study were utilized for the current study. To be included in the

study, participants were limited to individuals between the ages of 18-39 not currently in a romantic relationship. Participants also needed to have complete data on all major study variables (attachment, parental divorce, religiosity, and parental divorce) to be included in the sample, leaving the final sample of 741 individuals. Participants were recruited through Amazon's Mechanical Turk (MTurk) online labor market (*N* = 306) and a small, Midwestern college (*N* = 435). Amazon's MTurk is a desirable method of participant recruitment because it offers a more diverse population than the samples found on college campuses, which helps to make the participants more representative of the entire population (Buhrmester et al., 2018).

Of the 741 participants, 222 men and 519 women were included in the study. The average age was 22.26 (*SD* = 3.51). The majority of participants identified as White (83.7%) followed by African American/Black (7.0%), Asian American/Pacific Islander (4.0%), Native American/Alaskan Native (1.5%), Hispanic/Latino(a) (4.6%), and Biracial (1.8%). Out of all the participants, 86.4% identified as heterosexual, 2.4% as homosexual, and 4.7% as bisexual. Of the participants, 64.2% attended or finished college or some type of higher education, 30.2% graduated high school, 1.1% attended a trade school, and 0.8% did not graduate high school. 60.6% were currently full time students, 2.7% were currently part time students, and 36.2% were not currently in school.

The participants recruited from the Amazon MTurk online labor market were paid $2.00 for completing the survey. In addition, those participants recruited from the small, Midwestern college were compensated through receiving research credits.

Measures

Sociosexual Orientation

To assess sociosexual orientation, participants were asked to fill out Penke and Asendorpf's (2008) 9-item Sociosexual Orientation Inventory- Revised (SOI-R) scale. The SOI-R includes three factors of sociosexual orientation: previous behavior (number of casual sexual intercourses and changing partners, e.g., "With how many different partners have you had sex within the past 12 months?"), attitudes (towards uncommitted sexual relationships, e.g., "Sex without love is OK."), and desire (for sexual relations with people with whom an individual has no romantic connection, e.g., "I do not want to have sex with a person until I am sure that we will have a long-term, serious relationship."). Each subscale consists of 3-items on a 9 point scale. Questions related to behavior have anchors ranging from 0 to 20 or more. Items related to attitude have anchors from *strongly disagree* to *strongly agree* and items related to desire have anchors from *never* to *at least once a day*. Items were combined to create a mean for each subscale as well as an overall SOI score (an overall average of all 9 items) ranging from 1 to 9. Higher scores indicate a more unrestricted sociosexual orientation. The overall scale had a Cronbach's alpha of .872.

Romantic Attachment

To assess romantic attachment, participants completed the Experiences in Close Relationships-Revised (ECR-R; Fraley et al., 2000), a self-report questionnaire about attachment styles which measures two dimensions: attachment anxiety and attachment avoidance. The ECR-R is a revised version of Brennan et al. (1998) Experiences in Close Relationships questionnaire and consists of 36 items using a 5-point likert scale response, with the scale ranging from 1 (strongly disagree) to 5 (strongly agree). Within the assessment, 18 items measured attachment-related anxiety (e.g., "I'm afraid I will lose my partner's love") and 18 items measured attachment-related avoidance (e.g., "I prefer not to show a partner how I feel deep down"). Several items were reverse coded, and the items were re-coded so that higher scores indicate higher attachment-related anxiety and attachment-related avoidance, respectively. The subscales for attachment anxiety (Cronbach's alpha = .936) and attachment avoidance (Cronbach's alpha = .941) indicated good internal reliability.

Parental Divorce

Parental divorce was measured by a multiple choice question ("What is the marital status of your biological parents?"). Participants had the option to answer with *never married*, *engaged*, *married*, *separated, divorced*, or *widowed*. The scores were separated into two categories of "married/engaged" or "separated/divorced." The participants who answered *widowed* or *never married* were excluded because those experiences are inherently different from the construct of interest (divorce).

Religiosity

In order to assess religiosity, participants were asked to fill out a 2-item measure established by Burdette et al. (2009). The measure included questions to assess an individual's frequency of church attendance (i.e., religious behavior) and his or her subjective religiousness. To assess the individual's frequency of church attendance the participants were asked: "How often do you attend religious services?" The participants' responses could range from (1) never or almost never to (4) almost every week. To assess an individual's

subjective religiousness, the participants were asked to indicate their response to: "How religious do you consider yourself to be?" on a response scale of (1) not at all religious to (4) very religious. These two items were combined into a single measure of religiosity (Cronbach's alpha = .801).

Gender

To assess the participant's gender, the participants were asked the following item: "What is your gender?" Response options were *Male*, *Female*, or *I Prefer Not to Answer*.

Age

Age was assessed by asking participants for their birthdate. Using the date of the survey and the birth date, the participants' age in years was calculated.

Results

For the overall sample, the average SOI score was 3.92 (*SD* = 1.69), which is slightly below the averages reported in Penke, 2013. The average attachment avoidance and attachment anxiety scores were 2.66 (*SD* = 0.85) and 2.80 (*SD* = 0.88) respectively. The average score for religiosity was 1.88 (*SD* = .87). For the divorce measure, 62.6% of participants indicated that their biological parents were married/engaged, and 37.4% indicated that their parents were divorced/separated. Age was significantly associated with SOI, $r = .25$, $p < .001$, although interestingly the association was positive, suggesting that in this sample of young adults, older individuals reported a more unrestricted sociosexual orientation. Men also reported statistically significantly higher SOI scores than women, indicating more unrestricted sociosexual orientations (see Table 1 for comparisons across men and women for major study variables).

Association between attachment and SOI

In order to evaluate H1 and H2 (that there would be a positive association between avoidant attachment and SOI and anxious attachment and SOI), we ran a series of correlations. These hypotheses were confirmed. There was a statistically significant, weak, correlation between anxious attachment and SOI such that those with higher attachment anxiety reported more unrestricted sociosexual orientations, $r = .09$, $p = .017$. Similarly, there was a statistically significant weak association between attachment avoidance and SOI such that those with higher attachment avoidance reported more unrestricted sociosexual orientations, $r = .15$, $p < .001$.

Parental Divorce

In order to evaluate the differences in SOI based on parental divorce (H3), we ran an independent samples t-test. Our hypothesis was not supported; there was no evidence of a difference in SOI based on parental divorce status, $t(741) = .73$, $p = .466$, $d = .06$.

Religiosity

In order to evaluate whether there was a significant relationship between religiosity and SOI (H4), we ran a correlation. Our hypothesis was confirmed. There was a statistically significant moderate association between religiosity and SOI, $r = -.26$, $p < .001$, such that higher religiosity was associated with a more restricted sociosexual orientation.

Table 1
Independent Samples T-Test Comparing Attachment Avoidance, Attachment Anxiety, and SOI between Men and Women, N = 741

	Men			Women				
	M	*SD*	*SE*	*M*	*SD*	*SE*	*t*	*d*
Attachment Avoidance	2.73	0.81	0.05	2.63	0.86	0.04	1.47	0.12
Attachment Anxietv	2.81	0.92	0.06	2.79	.87	0.04	.34	0.03
Sociosexual Orientation Inventory (SOI)	5.00	1.53	0.10	3.46	1.53	0.07	12.59***	1.01

*** $p < .001$

Note. M=Mean. *SD*=Standard Deviation. *SE*=Standard Error.

Linear Regression

In order to evaluate whether religiosity, attachment, and divorce uniquely predict SOI, we ran a simultaneous multiple linear regression with sociosexual orientation as the outcome and with religiosity, attachment avoidance, attachment anxiety, divorce as predictors and age and gender as control variables. The linear regression was statistically significant and explained a moderate amount of the variability in SOI, $F(6, 734) = 42.23$, $p < .001$, $R^2 = .26$. Attachment avoidance ($\beta = 0.09$, $p = .014$) and religiosity ($\beta = -0.22$, $p < .001$) were significant predictors of SOI in the overall model, but attachment anxiety and divorce were not (see Table 2). Additionally, gender and age were associated with SOI such that men and older individuals reported higher SOI scores.

Table 2

Regression Results, N = 741

Predictor	Estimate (B)	SE	p	Stand. Estimate (β)
Intercept	3.63	0.45	< .001	
Religiosity	-0.42	0.06	< .001	-0.22
Attachment Avoidance	0.17	0.07	0.014	0.09
Attachment Anxiety	0.10	0.07	0.139	0.05
Age	0.06	0.02	< .001	0.12
Divorce[a]	-0.04	0.11	0.734	-0.02
Sex[b]	-1.34	0.12	< .001	-0.79

Note. [a] reference group is married. [b] reference group is males

Discussion

Attachment and SOI

The results confirmed our prediction of a positive relationship between both anxious attachment and avoidant attachment and SOI at the bivariate level, although it is important to note that these associations were small. Research in this field has shown evidence of this same relationship. Anxiously attached individuals are often clingy in relationships and may engage in unwanted casual sexual relationships in order to obtain intimacy with a partner and those who are avoidantly attached may fear intimacy leading to more casual sexual relationships and sexual permissiveness (e.g., Claxton & van Dulmen, 2013; Schmitt & Jonason, 2015; Simpson et al., 2004). However, when anxious and avoidant attachment were included in the overall model, only avoidant attachment predicted SOI, suggesting that avoidant attachment is a stronger predictor of SOI than anxious attachment.

Divorce and SOI

SOI and parental divorce were not related in either bivariate examinations or in the overall linear regression. An individual whose parents were divorced was not more likely than those whose parents were still married to have a higher SOI score. These findings contradict what other research in this field has found between divorce and sociosexual orientation. Barber (1998) and Simons et al. (2013) suggested that parental divorce was a predictor of more unrestricted sociosexual orientation and casual sex. One explanation for why we did not find this relationship is because parental divorce has a wide variety of negative and positive effects on individuals that our study may not capture. For example, researchers suggest that divorce has more of an impact at later ages than earlier ages (Newcomer & Urdy, 1987). Specifically, if a parental divorce happens before age 10, then the relationship between parental divorce and unrestricted sexual behavior is not significant (Newcomer & Urdy, 1987). Another explanation for why we did not find significant results could be because after a parental divorce, a child's life could have improved if it was a high-conflict home prior. Our study did not consider the age at which individuals were when their parents divorced so we were unable to account for these effects of age which could have weakened the relationship between divorce and SOI.

Religiosity and Attachment and SOI

Religiosity did show a moderate, negative relationship with SOI, both at the bivariate level and in the linear regression, which means that individuals who indicated more religious behavior and subject religiosity had more restricted sociosexual orientations than individuals who indicated low religious behavior and subjective religiosity. This confirms previous research (Haddad et al., 2016; Rostosky et al., 2004).

Strength of Predictors

When examining attachment anxiety, attachment avoidance, religiosity, divorce, sex, and age as predictors of SOI, attachment avoidance, religiosity, sex, and age were the only significant predictors in the model. This suggests that attachment avoidance and religiosity serve as better predictors of SOI than attachment anxiety and divorce, and that religiosity is a stronger predictor of SOI than avoidant attachment. Given the small effect sizes for the other variables, overall, this study suggests that religiosity is an important predictor of SOI in emerging and young adults, whereas attachment style and divorce may not be as important for understanding SOI.

Limitations

One limitation of this study is that when measuring parental divorce, participants were asked the marital status of their biological parents. For participants who were adopted, it would be more informative to ask about the marital status of their adoptive parents, rather than their biological parents, especially if the adoption occurred at a young age. Although it is unlikely that this affected a large proportion of the individuals (as adopted children make up only about 2% of the population in the United States; Adoption Network, 2020), it could have affected some participants. In addition, we were unable to evaluate additional information about the divorce (e.g., when it occurred, the relationship between the parents) which limited our ability to examine aspects of the divorce that may have moderated the association with SOI.

Furthermore, a limitation to this study was the cross-sectional design. Because the data was collected at one time and not longitudinally over a period of time, we are unable to establish temporal precedence. For example, it is unclear whether attachment or SOI comes first in the relationship. This is in fact an issue with much of the research because there is no research that can claim a causal relationship between SOI and the predictors evaluated.

Finally, it is important to note that many of the associations found in this study demonstrated small effect sizes, especially those related to attachment avoidance and attachment anxiety, suggesting that there are other constructs that may be better predictors of sociosexual orientation. The findings related to associations between attachment and sociosexual orientation should be interpreted with caution given the small effects.

Future Directions

The limitations of our study help identify possible future directions for future researchers. We found a relationship between romantic attachment and religiosity and sociosexual orientation in a cross-sectional design, but a longitudinal study could improve research about how these variables are connected across time.

Furthermore, previous researchers have suggested that the age at which children experience parental divorce can have an effect on the relationship between divorce and attachment (Newcomer & Urdy, 1987), so including information on when parental divorce occurs in the parental divorce measure could allow for an examination of this factor.

Future research could further evaluate how romantic attachment, religiosity, and sociosexual orientation connect specifically to wellbeing. Research has suggested that engaging in casual relationships (i.e., having a more unrestricted sociosexual orientation) may be related to lower well being (Bersamin et al., 2014; Dube et al., 2017). However, future research could evaluate the specific roles of attachment and religiosity in understanding this connection.

Finally, previous researchers have suggested that the type of religion an individual participates in can be influential on the individual's sociosexual orientation (Lefkowitz et al., 2004). Future research may include an additional question within the religiosity scale in order to determine which religion an individual identifies with, because it may help to explain which religions have the greatest influence on the willingness people have in engaging in casual sexual relationships.

Implications

There are several ways in which our results connect with society and the field of study. The established connection between romantic attachment and sociosexual orientation helps us understand how an individual's attachment within romantic relationships (particularly avoidant attachment) may be associated with the individual's sexual experiences, although the association in the current study only demonstrated a weak association. Additionally, this study provides some confirmation of previous researchers who argued that divorce may not have a long lasting association with sociosexuality. Furthermore, this study shows that higher levels of religiosity are associated with a more restricted sociosexual orientation, suggesting that religiosity may help predict engagement in casual sexual relationships and experiences. Together, this provides us with a clearer understanding of the predictors of sociosexuality.

References

Adoption Network. (2020). *US adoption statistics*. Retrieved July 1, 2020, from https://adoptionnetwork.com/adoption-statistics

Barber, N. (1998). Sex differences in disposition towards kin, security of adult attachment, and sociosexuality as a function of parental divorce. *Evolution and Human Behavior*, *19*(2), 125-132. https://doi.org/10.1016/S1090-5138(98)00004-X

Barkan, S. E., (2006). Religiosity and premarital sex in adulthood. *Journal for the Scientific Study of Religion, 45*(3), 407-417. https://doi.org10.1111/j.1468-5906.2006.00315.x

Bersamin, M. M., Zamboanga, B. L., Schwartz, S. J., Donnellan, M. B., Hudson, M., Weisskirch, R. S., . . . & Caraway, S. J. (2014). Risky business: Is there an association between casual sex and mental health among emerging adults? *Journal of Sex Research, 51*(1), 43-51. https://doi.org/10.1080/00224499.2013.772088

Bowlby, J. (1969). *Attachment: Attachment and loss*. Basic Books.

Bowlby, J. (1989). *The role of attachment in personality development and psychopathology*. In S. I. Greenspan & G. H. Pollock (Eds.), *The course of life, Vol. 1. Infancy* (p. 229-270). International Universities Press, Inc.

Brase, G. L., Adair, L., & Monk, K. (2014). Explaining sex differences in reactions to relationship infidelities: Comparisons of the roles of sex, gender, beliefs, attachment, and sociosexual orientation. *Evolutionary Psychology, 12*(1), 73-96. https://doi.org/10.1177/147470491401200106

Brennan, K. A., Clark, C. L., & Shaver, P. R. (1998). Self-report measurement of adult attachment: An integrative overview. In J. A. Simpson & W. S. Rholes (Eds.), *Attachment theory and close relationships* (pp. 46-76). Guilford Press.

Buhrmester, M. D., Talaifar, S., & Gosling, S. D. (2018). An evaluation of Amazon's Mechanical Turk, its rapid rise, and its effective use. *Perspectives on Psychological, 13*(2), 149-154. https://doi.org/10.1177/1745691617706516

Burdette, A. M., Ellison, C. G., Hill, T. D., & Glenn, N. D. (2009). "Hooking up" at college: Does religion make a difference?. *Journal for the Scientific Study of Religion, 48*(3), 535-551. https://doi.org/10.1111/j.1468-5906.2009.01464.x

Claxton, S. E., & van Dulmen, M. H. (2013). Casual sexual relationships and experiences in emerging adulthood. *Emerging Adulthood, 1*(2), 138-150. https://doi.org/10.1177/2167696813487181

Collins, N., Clark, C. L., & Shaver, P. R. (1996). Attachment styles and internal working models of self and relationship partners. In G. J. O. Fletcher & J. Fitness (Eds.), *Knowledge structures in close relationships: A social psychological approach* (pp. 25- 62). Lawrence Erlbaum Associates.

Cui, M., & Fincham, F. D. (2010). The differential effects of parental divorce and marital conflict on young adult romantic relationships. *Personal Relationships*, *17*(3), 331-343. https://doi.org/10.1111/j.1475-6811.2010.01279.x

Del Giudice, M. (2011). Sex differences in romantic attachment: A meta-analysis. *Personality and Social Psychology Bulletin, 37*(2), 193-214. https://doi.org/10.1177/0146167210392789

Dube, S., Lavoie, F., Blais, M., & Hebert, M. (2017). Consequences of casual sex relationships and experiences on adolescents' psychological well-being: A prospective study. *The Journal of Sex Research, 54*(8), 1006-1017. https://doi.org/10.1080/00224499.2016.1255874

Feeney, J. A., Kelly, L., Gallois, C., Peterson, C., & Terry, D. J. (1999). Attachment style, assertive communication, and safer-sex behavior. *Journal of Applied Social Psychology, 29*(9), 1964-1983. https://doi.org/10.1111/j.1559-1816.1999.tb00159.x

Fraley, R. C., & Heffernan, M. E. (2013). Attachment and parental divorce: A test of the diffusion and sensitive period hypotheses. *Personality and Social Psychology Bulletin*, *39*(9), 1199-1213. https://doi.org/10.1177/0146167213491503

Fraley, R. C., & Shaver, P. R. (2000). Adult romantic attachment: Theoretical developments, emerging controversies, and unanswered questions. *Review of General Psychology, 42*(2), 132-154. https://doi.org/10.1037/1089-2680.4.2.132

Fraley, R. C., Waller, N. G., & Brennan, K. A. (2000). An item response theory analysis of self-report measures of adult attachment. *Journal of Personality and Social Psychology, 78*(2), 350-365. https://doi.org/10.1037/0022-3514.78.2.350

Gangestad, S. W., & Simpson, J. A. (1990). Toward an evolutionary history of female sociosexual variation. *Journal of Personality*, *58*(1), 69-96. https://doi.org/10.1111/j.1467-6494.1990.tb00908.x

Gentzler, A. L., & Kerns, K. A. (2004). Associations between insecure attachment and sexual experiences. *Personal Relationships, 11*(2), 249-265. https://doi.org/10.1111/j.1475-6811.2004.00081.x

Haddad, B., Angman, M., Archer, T., & Garcia, D. (2016). Dark triad, sociosexual orientation and religious affiliation: An association and moderation study. *Clinical and Experimental Psychology, 2*(2), 124- 128. https://doi.org/10.4172/2471-2701.1000124

Jónsson, F. H., Njarðvik, U., Ólafsdóttir, G., & Grétarsson, S. (2000). Parental divorce: long-term effects on mental health, family relations and adult sexual behavior. *Scandinavian Journal of Psychology*, *41*(2), 101-105. https://doi.org/10.1111/1467-9450.00177

Le Gall, A., Mullet, E., & Shafighi, S. R. (2002). Age, religious beliefs, and sexual attitudes. *Journal of Sex Research*, *39*(3), 207-216. https://doi.org/10.1080/00224490209552143

Lefkowitz, E. S., Gillen, M. M., Shearer, C. L., & Boone, T. L. (2004). Religiosity, sexual behaviors, and sexual attitudes during emerging adulthood. *Journal of Sex Research, 41*(2), 150-159. https://doi.org/10.1080/00224490409552223

Lyons, H. A., Manning, W. D., Longmore, M. A., & Giordano, P. C. (2015). Gender and casual sexual activity from adolescence to emerging adulthood: Social and life course correlates. *Journal of Sex Research, 52*(5), 543-557. https://doi.org/10.1080/00224499.2014.906032

Moon, J. W., Krems, J. A., Cohen, A. B., & Kenrick, D. T. (2019). Is nothing sacred? Religion, sex, and reproductive strategies. *Current Directions in Psychological Science, 28*(4), 361-365. https://doi.org/10.1177/0963721419838242

Newcomer, S., & Udry, J. R. (1987). Parental marital status effects on adolescent sexual behavior. *Journal of Marriage and the Family*, *49*(2) 235-240. https://doi.org/10.2307/352296

Penke (2013). The revised Sociosexual Orientation Inventory. In T. D. Fisher, C. M. Davis, W. L. Yarber, & S. L. Davis (Eds.): *Handbook of sexuality-related measures* (3rd Ed.). Taylor & Francis.

Penke, L., & Asendorpf, J. B. (2008). Beyond global sociosexual orientations: a more differentiated look at sociosexuality and its effects on courtship and romantic relationships. *Journal of Personality and Social Psychology*, *95*(5), 1113-1135. https://doi.org/10.1037/0022-3514.95.5.1113

Proulx, C. M., & Snyder-Rivas, L. A. (2013). The longitudinal associations between marital happiness, problems, and self-rated health. *Journal of Family Psychology*, *27*(2), 194-202. https://doi.org/10.1037/a0031877

Rodrigues, D., & Lopes, D. (2017). Sociosexuality, commitment, and sexual desire for an attractive person. *Archives of Sexual Behavior, 46*(3), 775-788. https://doi.org/10.1007/s10508-016-0814-3

Rostosky, S. S., Wilcox, B. L., Wright, M. L. C., & Randall, B. A. (2004). The impact of religiosity on adolescent sexual behavior: A review of the evidence. *Journal of Adolescent Research, 19*(6), 677-697. https://doi.org/10.1177/0743558403260019

Schmitt, D. P., & Jonason, P. K. (2015). Attachment and sexual permissiveness: Exploring differential associations across sexes, cultures, and facets of short-term mating. *Journal of Cross-Cultural Psychology*, *46*(1), 119-133. https://doi.org/10.1177/0022022114551052

Simmons, B. L., Gooty, J., Nelson, D. L., & Little, L. M. (2009). Secure attachment: Implications for hope, trust, burnout, and performance. *Journal of Organizational Behavior, 30*(2), 233-247. https://doi.org/10.1002/job.585

Simon, E. P. (1997). *Adult attachment style and sociosexuality [Doctoral dissertation] Dissertation Abstracts International.* https://psycnet.apa.org/record/1997-95006-038

Simons, L. G., Burt, C. H., & Tambling, R. B. (2013). Identifying mediators of the influence of family factors on risky sexual behavior. *Journal of Child and Family Studies*, *22*(4), 460-470. https://doi.org/10.1007/s10826-012-9598-9

Simpson, J. A., Wilson, C. L., & Winterheld, H. A. (2004). Sociosexuality and romantic relationships. *The handbook of sexuality in close relationships* (97-122). Psychology Press.

Sutton, T. E. (2019). Review of attachment theory: Familial predictors, continuity and change, and intrapersonal and relational outcomes. *Marriage & Family Review*, *55*(1), 1-22. https://doi.org/10.1080/01494929.2018.1458001

Tracy, J. L., Shaver, P. R., Albino, A. W., & Cooper, M. L. (2003). Attachment styles and adolescent sexuality. In P. Florsheim (Ed.), *Adolescent romance and sexual behavior: Theory, research, and practical implications* (p. 137-159). Lawrence Erlbaum Associates.

Vrangalova, Z., & Ong, A. D. (2014). Who benefits from casual sex? The moderating role of sociosexuality. *Social Psychological and Personality Science*, *5*(8), 883-891. https://doi.org/10.1177/1948550614537308

Washington, K. N., & Hans, J. D. (2013). Romantic attachment among young adults: The effects of parental divorce and residential instability. *Journal of Divorce & Remarriage*, *54*(2), 95-111. https://doi.org/10.1080/10502556.2012.752684

Author Note

Correspondence concerning this article should be addressed to Shannon E. Claxton, Morningside College, 1501 Morningside Ave, Sioux City, IA 51106. Email: claxtons@morningside.edu

Journal of Psychological Inquiry
2021, Vol.25, No. 2, pp. 46-54

FACTORS AFFECTING RESPONSE PERSISTENCE IN COMPUTER GAMING

MCKENZIE B. WALLACE AND JAMES V. LUPO

CREIGHTON UNIVERSITY

Abstract – As the popularity of video gaming steadily increases, so too does the prevalence of problematic gaming behavior and associated negative health outcomes. Schedules of reinforcement are frequently utilized in video game design to promote continued play and are considered a major factor contributing to excess gaming. Even so, there is a lack of empirical data examining the specific effect of particular video game structural characteristics, including reinforcement schedules, on the development of persistence in problem video game behaviors. In the present study, a computer program was written to create a "spaceship shooter game" in which specific game variables, including reinforcement schedule and visual graphic elements, could be manipulated. Consenting undergraduate students (N = 24) were randomly assigned to four groups, playing the game on either a continuous or a variable schedule of reinforcement in the presence or absence of graphic elements. Following 20 reinforced trials, mean responses in a five-minute extinction phase were compared. A significant effect of graphic elements was observed (p =.04), with presence of graphics promoting resistance to extinction. Though the main effect of reinforcement schedule only approached significance, variable reinforcement schedules appeared to promote persistence of play, a phenomenon widely supported by the literature. Our data provide additional empirical evidence for the relationship between response persistence in video game play and specific structural features of the game, including reinforcement schedules and sensory features.

Keywords: reinforcement schedules, extinction, partial reinforcement, operant conditioning, video games, computer games, gaming addiction

As one of the fastest growing entertainment industries in the United States, the video game industry alone added almost $12 billion to the country's economy in 2016 (Entertainment Software Association, 2017). According to a 2017 report by the Entertainment Software Association, the annual consumer spending on video game content is rapidly increasing, from $17.5 billion in 2010 to $24.5 billion in 2017. As the popularity of video games increases, however, so do the rates of problematic gaming behavior. While there is significant debate over the classification of these problematic behaviors as "addictive" (Yellowlees & Marks, 2007), "internet gaming disorder" was recently included in the appendix of the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 2013) and "gaming disorder" was added to the World Health Organization's 11th revision of its International Classification of Diseases (World Health Organization, 2018). Concern with video gaming overuse stems from an increasing body of evidence outlining associated negative consequences. Psychosocial impacts of excess gaming behaviors include increased stress (Batthyány et al., 2009; Kuss, 2013), attention deficits (Batthyány et al., 2009; P. Chan & Rabinowitz, 2006), poorer social skills (Zamani et al., 2010), loneliness (Lemmens et al., 2011), and decreased academic achievement (Jeong & Kim, 2011; Skoric et al., 2009). Excessive gaming behaviors are often associated with a reduction in time spent engaging in other hobbies, socializing, spending time with family, working, and sleeping (Batthyány et al., 2009; Griffiths et al., 2004; Peng & Liu, 2010; Yee, 2006). Furthermore, significant health consequences have been reported and range in severity from various musculoskeletal injuries including tenosynovitis and wrist pain (Griffiths et al., 2012) to sleep disruptions (Dworak et al., 2007) and epileptic seizures (Chuang, 2006).

Given the continuously increasing popularity of video games in the general population and the resultant elevation in prevalence of problematic gaming behavior, factors contributing to the initiation and maintenance of these behavioral patterns are of particular interest. It has been suggested, for example, that the underlying neural processes associated with problematic gaming share

similarities with other behavioral addictions and substance addictions and that genetic predispositions may increase susceptibility to excess gaming (Griffiths et al., 2012). Increasingly, research has also examined the role of intrinsic game features that promote play as, much like gambling addiction, individual consumer factors coupled with these structural features motivate the gamer to continue playing despite negative consequences.

The influence of structural gaming features can be difficult to study empirically (King et al., 2010a) and early attempts to systematically classify the different structural features of games focused on gaming features that appealed to players (Wood et al., 2004). Because this early attempt at classification focused on reportedly pleasurable aspects of the gaming experience, it likely overlooked important game features that lead to play persistence (King et al., 2010a) As such, additional structural characteristics of games have since been identified that may contribute to problem game playing, including social features, reward and punishment features, and presentation features (e.g., visual and auditory features of the game) among others (King et al., 2010a).

Of particular interest to the current investigation are the role of so-called "reward and punishment features" and "presentation features" in continued gaming. Much like the role of operant conditioning in gambling behaviors, specific schedules of reinforcement are utilized in game design and appear to promote consistent responding (Ferster & Skinner, 1957; Hopson, 2001; Yee, 2006). For example, reinforcement of behavior on a variable ratio schedule produces the most consistent pattern of responding, is the least susceptible to extinction, and is also one of the most commonly utilized reinforcement schedules in video game design (Chumbley & Griffiths, 2006; Hopson, 2001; King et al., 2010a). In addition to variable ratio reinforcement schedules, gamers are commonly reinforced on concurrent, fixed ratio schedules (Hopson, 2001; Yee, 2006) that result in the typical pattern of rapid responding with minimal post-reinforcement pause (King et al., 2010b; King & Delfabbro, 2009).

The view that operant conditioning and specific reinforcement schedules play a significant role in problem gaming behavior is not new. In 1983, for example, Loftus and Loftus asserted that game designers integrated specific reinforcement schedules with the intent of making them more addictive (Loftus and Loftus, 1983). Today, operant conditioning theory remains the dominant framework used by researchers to explain the intrinsic game features that promote problematic behavior and reinforcement schedules are the most popular intrinsic gaming feature used in explanation of problematic game play (King et al., 2010b). Although specific reinforcement schedules are commonly used in video game design to promote play (Hopson, 2001; Klemm & Pieters, 2017; Yee, 2006) and it is well-understood that these reinforcement schedules result in consistent response patterns outside of the video gaming environment (Chumbley & Griffiths, 2006; Ferster & Skinner, 1957; Hopson, 2001; Yee, 2006), there is a lack of published experimental data examining the specific effect of particular video game structural characteristics, including reinforcement schedules, on the development of persistence in problem video game behaviors (King et al., 2010b). Much of the current literature relies on reportable measures of gaming behavior, including consumer surveys (King et al., 2011) and interviews (King & Delfabbro, 2009) as well as affective measures of excitement and frustration (Chumbley & Griffiths, 2006), and analogies to the literature on gambling behaviors (King et al., 2010b). Direct, empirical measures of persistence in extinction are rare; the paucity of experimental data is a result of significant challenges associated with the empirical study of gaming behavior including the challenge of creating an "ecologically valid" gaming experience within an experimental setting (Wood et al., 2004), the variable nature of rewards in video games compared to slot machines, and pragmatic reasons such as a lack of funding (King et al., 2010b). More empirical research is needed to better understand features of video games and online gaming environments that contribute to addictive gaming behavior (King et al., 2019) as there are important implications both for ethical game design and for management and treatment of problematic gaming (Klemm & Pieters, 2017).

The current investigation seeks to provide necessary empirical evidence for the relationship between response persistence in video game play and specific structural features of the game, including reinforcement schedules and sensory features. A computer program was designed in which schedules of reinforcement and sensory features could be specifically manipulated in a gaming environment similar to basic video games. Although many studies have used similar computer programs, video games, or virtual environments to investigate basic learning in humans (Catania, 2005; Fox & Oakes, 1984; Greville et al., 2013; Raia et al., 2000), few have applied them to the investigation of persistence in video game play specifically. Using extinction time as a measure of persistence, we varied the reinforcement schedule and presence of sensory feedback in a spaceship shooter game and measured average response rate in extinction. As a well-established principal in operant

conditioning theory, we expected variable reinforcement schedules to display the greatest resistance to extinction in the virtual gaming environment, indicated by higher average rates of responding during extinction. Additionally, we expected the presence of sensory feedback following nonreinforced trials of "correct play" to similarly promote persistence during the extinction phase.

Method

Participants

Participants included 24 volunteers from Psychology courses who were required to participate in a research study of their choosing or complete an alternative assignment for course credit. Participants' ages ranged from 18 to 22 and included an equal distribution of males and females.

Procedure

The study was conducted in a small research room (approximately 3m x 3m) with minimal extraneous noise and standard lighting. Participants were seated at a desk, one meter from an iMac computer (53cm screen) with a wireless mouse and keyboard. A standard testing session involved a series of organized steps that are required by research ethics procedures approved by the Institutional Review Board. Participants were escorted into the testing room and introduced to the requirements of the project. Upon entrance, a welcome screen was displayed on the computer which read "Please wait for further instruction." The participant sat facing the computer while the researcher observed from the back of the room for the duration of the study. Researchers introduced the game and explained the general task; "Welcome to our study. During this session, you will be asked to control a ship used to aim and fire missiles in order to hit a moving flying saucer using the wireless mouse." A brief demonstration of the testing procedures presented a view of the experimental conditions and provided greater detail about the requirements of the experiment. Participants were instructed to select the "Demo" option in the bottom of the screen and could move the ship and fire missiles until they were ready to continue. Upon confirmation that the subject was finished experiencing the demo, they were returned to the home screen. After informed consent procedures, the test was presented, and the data were collected and stored. At the conclusion of the session, participants were debriefed, and their questions were answered.

Design

Game play in the study consisted of a training phase with 20 reinforced trials followed by a five-minute extinction phase for all participants. Participants were divided into a total of four groups in a 2 x 2 Reinforcement Schedule by Graphics Presence factorial design. In other words, reinforcement schedule and the presence or absence of graphic elements were varied to create four unique combinations. Two of the groups received continuous reinforcement (CR), receiving reinforcement in the form of saucer destruction after every successful hit, while the other two groups received partial reinforcement (PR) on a progressive variable ratio schedule, presented in detail below. Of the two groups exposed to each reinforcement schedule, one group received feedback on nonreinforced hits in the form of graphic elements (i.e., the appearance of "shield factors" ("+SF") around the ship when a hit landed but the ship did not explode) while the other received no feedback in the form of shield factors ("-SF") from graphic elements within the game (Table 1). Regardless of reinforcement schedule or graphic elements, all participants received 20 reinforced trials during the training phase and subsequently entered a five-minute extinction phase during which no reinforcement occurred. The average response rate during extinction as well as the total number of responses during each of five consecutive one-minute extinction blocks were measured to determine the effect of reinforcement schedule and graphic elements on play persistence and response pattern during extinction.

		Continuous Reinforcement (CR)	**Partial Reinforcement (PR)**
Graphic Elements	**"Shield Factors" present (+SF)**	Continuous Reinforcement/ Shield Factor Group ("CR/+SF")	Partial Reinforcement/ Shield Factor Group ("PR/+SF")
	No "Shield Factors" (-SF)	Continuous Reinforcement/ No Shield Factor Group ("CR/-SF")	Partial Reinforcement/No Shield Factor Group ("PR/-SF")

Table 1. A 2x2 Reinforcement Schedule by Graphic Elements Factorial Design

CR/+SF group

During the training phase, participants were required to shoot the missile and successfully hit the saucer for a total of 20 hits. Every shot fired that contacted the flying saucer was considered a "hit" and was reinforced via saucer destruction, denoted by a combination of particle emission from the flying saucer

and a visual/auditory explosion lasting approximately 1.5 seconds and occurring at the moment the missile hit the saucer. Participants received 20 reinforcements, after which they entered a five-minute extinction phase. During extinction, reinforcement was characteristically absent; the saucer did not explode when hit by a missile. For this group, however, each nonreinforced hit included graphic elements termed "shield factors" (SF). Though the saucer never exploded during extinction, a series of green concentric circles emanated outwards from the saucer when it was hit by the missile, indicating the presence of its "shields". Notably, participants in this group did not experience shields until the extinction phase as they were continually reinforced during the training phase.

CR/-SF group

This group was also reinforced on a continuous schedule for a 20-reinforcement training phase followed by a five-minute extinction phase. The training phase for this group was identical to that of the CR/+SF Group. During the extinction phase, however, each nonreinforced hit lacked the "shield factor" graphic elements. No shields or other visual elements appeared when the missile hit the saucer which, characteristically, failed to explode. Though the shields were absent from nonreinforced trials throughout training and extinction phases for the -SF groups, a very small, unintended movement of the saucer occurred upon contact with the missile (see discussion).

PR/+SF group

During training phase for this group, participants were required to shoot the missile and hit the saucer on a progressive schedule of partial variable reinforcement (VR). These participants received three different stages of reinforcement, with each successive stage requiring more successful hits prior to reinforcement (saucer explosion). The first stage involved five reinforcements on a VR2 ± 1 schedule. Generally, "VR2 ± 1" refers to a variable ratio schedule with an average of two responses required for reinforcement but varying between one and three responses required for each trial. In the context of the current study, the number of successful hits required to destroy the saucer varied from one to three between trials, for an overall stage one average of two hits required to destroy the saucer. After five successfully destroyed saucers, participants entered the second stage of partial reinforcement, characterized by an additional five reinforcements on a VR2 + 5 schedule in which an average of two hits was again required to destroy the saucer for the entire stage, but with the number of hits required for each individual trial varying from one to seven hits. The final stage of partial reinforcement required participants to successfully destroy 10 saucers on a VR3 + 5 schedule, with an average of three hits required to destroy each saucer and with the required number of hits for individual trials varying from one to eight. As with the continuous reinforcement groups, saucer destruction served as reinforcement during each training stage and involved particle emission and an explosion. On nonreinforced trials during the training phase as well as throughout the extinction phase, shields appeared around the saucer upon missile contact ("+SF"). During extinction, participants in this group were presented with the graphic shield element when the missile contacted the saucer, though no explosion occurred.

PR/-SF group

The progressive partial reinforcement schedule for this group was identical to that for the PR/+SF group, however nonreinforced trials were not met with the appearance of shield factor. Instead, the missile passed through the saucer with no effect. During the five-minute extinction period, a "hit" was similarly ineffective: neither an explosion nor shields were presented.

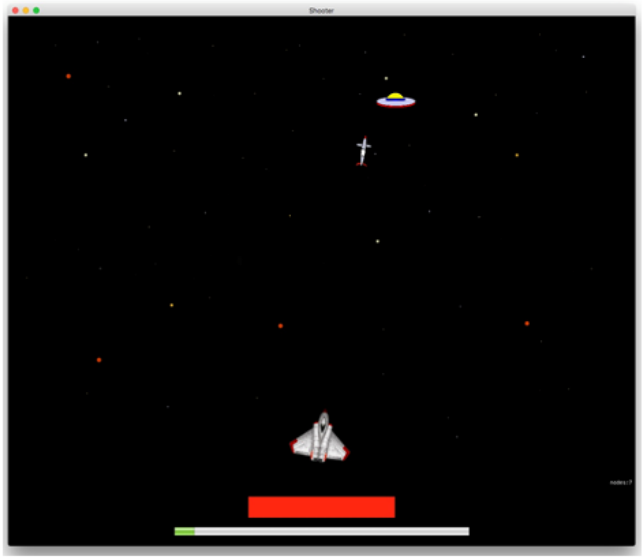


Figure 1. Shooter test environment display area.

Apparatus

The ***Shooter*** test environment is shown in Figure 1. The participants positioned the cursor in the red "fire control" area in the bottom center of the display. By moving the cursor back and forth in this area the defender ship could be rotated to aim the missile. Clicking the mouse button launched the missile. Only one missile was allowed in space at a time. This feature was

built into the test to avoid a "machine gun" approach to hitting the target and making the task a trivial one. The progress bar across the bottom of the screen advanced with each destroyed saucer. When the missile hit the saucer on a reinforced trial there was a visual explosion and audio feedback indicating saucer destruction.

Data Analysis

The total number of responses during extinction was measured and used to calculate the average response rate during extinction, as all participants underwent a five-minute extinction period. A three-way ANOVA with reinforcement schedule and sensory feedback as between-participants factors and time block in extinction as a within-subject factor was used to compare response rates. Post-hoc analyses for the significant main effect of time block indicated by ANOVA were done via Newman-Keuls pairwise comparisons.

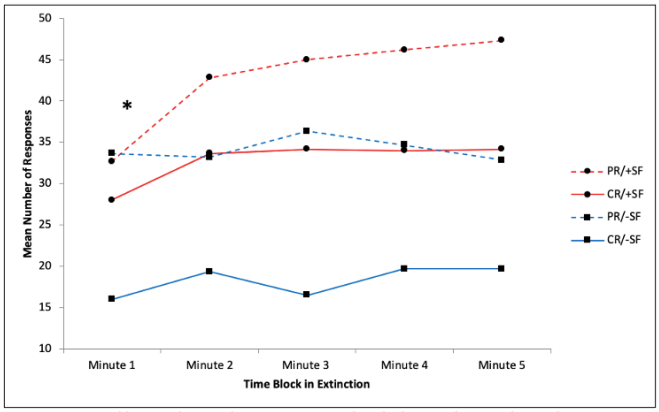


Figure 2. Effect of reinforcement schedule and graphic elements on mean responses during five consecutive one-minute time blocks in extinction.

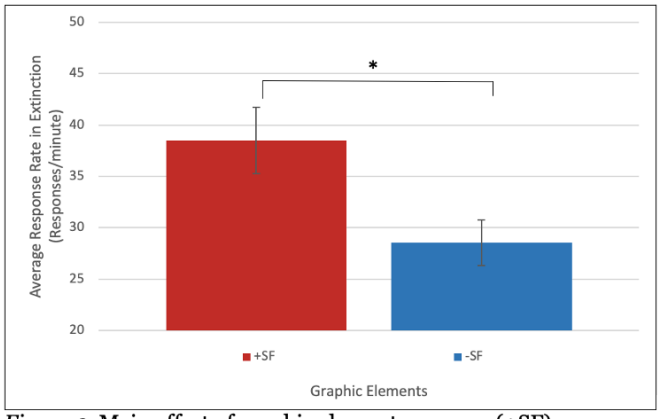


Figure 3. Main effect of graphic element presence (+SF) on average response rate in extinction.

Results

The three factor interaction between sensory feedback, reinforcement schedule, and time block approached significance, [F(4, 76) = 2.28, p = .07, see Figure 2], but only achieved a power of 0.37 to detect a medium effect size (f = 0.5). While no follow-up analyses were performed for this non-significant effect, examination of Figure 2 suggests that while the responses for groups CR/+SF, CR/-SF, and PR/-SF remained steady throughout extinction, the number of responses for group PR/+SF showed an increase over the five minutes of extinction testing, skewing the mean for the overall group.

Three-way ANOVA indicated a significant main effect of sensory feedback, which significantly increased average response rate during extinction [F(1,19) = 4.83, p = .04; see Figures 3 and 5]. Participants who experienced audiovisual feedback in the form of shield factors during nonreinforced hits fired an average of 38.78 missiles per minute during extinction (SD = 11.15) while participants without shield factors exhibited an average of 28.55 responses per minute during extinction (SD = 8.36). Given the effect size (f = 0.64), this analysis achieved an 86% power to correctly reject the null hypothesis.

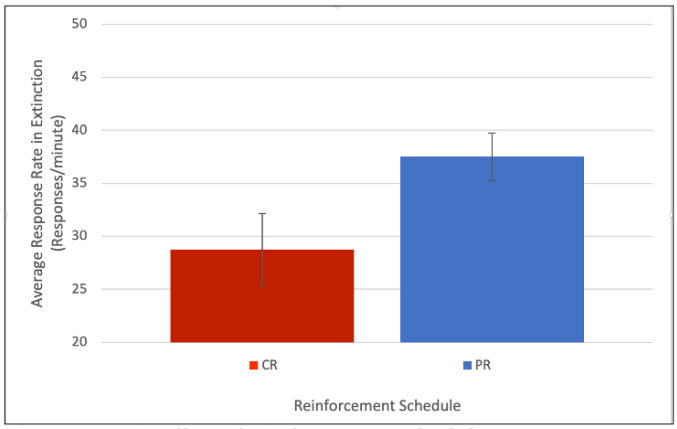


Figure 4. Main effect of reinforcement schedule on average response rate in extinction.

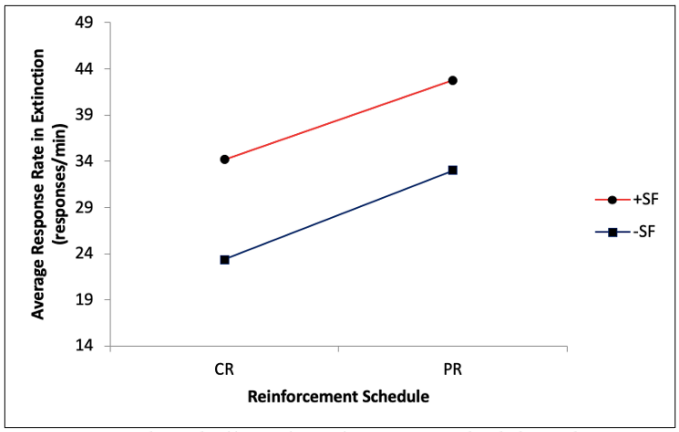


Figure 5. Combined effect of reinforcement schedule and presence of graphic elements on average response rate in extinction.

Participants who received the progressive partial reinforcement schedule during training tended to respond at a higher rate during extinction than continuously reinforced participants, although this effect only approached significance [$F(1, 19) = 3.87$, $p = .06$, see Figures 4 and 5]. Participants who had been reinforced on a partial schedule during training responded an average of 38.40 times per minute during extinction (SD = 7.66) while participants who were reinforced continuously during training responded an average of 28.8 times per minute in extinction (SD = 11.72). The current analysis achieved a 79% power to correctly reject the null hypothesis (f = 0.591).

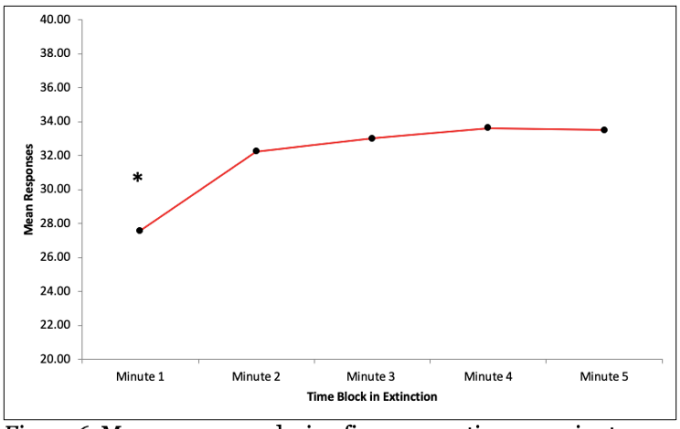


Figure 6. Mean responses during five consecutive one-minute time blocks in extinction for all participants.

There was also a significant main effect of time block in extinction on average response rate [$F(4, 76) = 4.37$, $p < .01$, see Figure 6]. Follow-up analyses using Newman-Keuls pairwise comparisons indicated that the mean number of responses in the first minute of extinction (M = 29.21, SD = 8.11) was significantly lower than that for all remaining time blocks (all significant p's < .05), none of which were significantly different from one another. The means (with standard deviations in parentheses) for minutes two-five were 33.79 (9.69), 33 (11.95), 33.63 (10.85), and 33.5 (11.31), respectively.

Discussion

Sensory feedback during nonreinforced trials significantly increased response rate in extinction as expected. For participants receiving sensory feedback, it is possible that the appearance of shields after a hit increased persistence in extinction by acting as a substitute reinforcer. Though the saucer did not blow up and participants were therefore not "reinforced" on trials with the shields, the presence of the shield may have acknowledged a successful hit and consequently reinforced a relatively high rate of responding. These results support the idea that even in the absence of direct reinforcers, visual aspects of video games such as the presence of detailed graphics contribute to gaming persistence. The Entertainment Software Association (2017), for example, reports that 67% of gamers in the United States rate "Quality of the Graphics" as the most significant factor influencing decisions to purchase video games. Graphic elements in games have also been postulated to contribute to problem gaming behavior (King et al., 2010a). The current findings provide empirical evidence for the relationship between graphic elements and persistent gaming behavior. Future research should investigate the degree to which the quality of graphics differentially reinforces game play and should further explore the mechanisms behind this relationship.

It is worth mentioning that the effect of sensory feedback may be more significant than was found in the current study. As briefly noted in the Methods section, the groups that did not receive sensory feedback in the form of shields (the "-SF" groups) still observed a small, consistent movement of the saucer when contacted by the missile on non-reinforced trials. This effect was first observed by the researchers during data collection and was unintentional. Even a minute movement of the saucer coinciding with missile contact is problematic on nonreinforced trials because, much like the "shields" effect, the saucer movement acknowledges missile contact and could act as a reinforcement substitute, albeit smaller in magnitude. Though our data still yielded significant results, correcting this small error in the program may produce a more robust effect and should be considered in future data collection.

The effect of reinforcement schedule on persistence in extinction only approached significance, a result that was inconsistent with our original hypothesis as well as the literature (Chumbley & Griffiths, 2006; Ferster & Skinner, 1957; King et al., 2010a, 2010b; King & Delfabbro, 2009). While our data did not display a statistically significant relationship between reinforcement schedule and response rate in extinction, we suspect that further data collection would yield significant results for a few reasons. First, the partial reinforcement extinction effect is a well-established learning phenomenon (Chumbley & Griffiths, 2006; Ferster & Skinner, 1957; King et al., 2010a, 2010b; King & Delfabbro, 2009). In addition, the prevalence of partial schedules of reinforcement in popular video games has long been thought to contribute to persistence of play in gamers (Hopson, 2001; King & Delfabbro, 2009; Yee, 2006). As such, it is reasonable to expect a relationship between partial reinforcement schedules and increased persistence in extinction in a virtual gaming

environment. The lack of significant results observed in the current study is likely due to an insufficient sample size to achieve appropriate power. Due to time constraints and lack of volunteers, the sample size of the current study was relatively small and achieved only a 15% power to detect a small effect (f = 0.2). For these reasons, we hypothesize that future data collection to increase the sample size would produce a significant main effect of reinforcement schedule on response rate in extinction.

While a significant main effect of time block in extinction may not have been surprising generally, the results of the follow-up analyses were somewhat unexpected. Typically, regardless of reinforcement schedules or visual elements, response rates decline in extinction given sufficient time. In fact, rate decrement is a characteristic feature of the extinction of a response (Ferster & Skinner, 1957). However, our data indicate that response rates for the first minute of extinction were significantly lower than the rates in subsequent one-minute time blocks across all groups. One might speculate, therefore, that a period of five minutes may not have been sufficient time to truly elicit extinction behavior. Increased response rate appeared to be most salient for the PR/+SF group (Figure 2); although the three-factor interaction of sensory feedback, reinforcement schedule, and time block was insignificant, this group received both partial reinforcement and shield factors, both tending to promote resistance to extinction. Increasing response rates in the PR/+SF group possibly skewed the mean response rates for the whole group in later time blocks and further supports the idea that the five-minute extinction phase was not long enough to truly extinguish responding particularly in participants receiving partial reinforcement and sensory feedback. Given that the response rate failed to decline significantly as is characteristic of extinction, future research should extend the extinction phase.

Finally, generalizability of this study is limited by the experimental environment in which it was conducted. As previously mentioned, one of the major barriers to experimental examination of persistent gaming behaviors is the challenge of creating an experimental environment similar to the true gaming environment (Wood et al., 2004). Although the current study aimed to create a valid gaming environment, multiple factors unique to an experimental setting likely influenced participant behavior throughout the experiment. For example, perceived social pressures experienced by the participants during the session may have influenced responding during extinction. The researchers conducting the experiments remained in the testing room throughout the session. Though seated quietly in the back of the room to observe gaming behaviors, researchers reported receiving inquiries as participants entered the extinction phase, including “Is this what I’m supposed to be doing?” and “Is the game broken?” The researchers provided as little feedback as possible, however participants may have felt as if they were expected to continue firing missiles at the saucers because they were being observed or because they felt compelled to complete the experimental session (although they were ensured prior to study initiation that completion was not necessary). Future research in experimental settings should exclude the researcher from the testing room during the session and ensure participants that their behavior is not being recorded. Furthermore, the same gaming program could reasonably be administered in a remote environment that would replicate participants’ preferred, real-world gaming environment. Further investigations should explore the possibility of remote administration of the shooter game to improve external validity.

The current study adds to the body of empiric data demonstrating that features intrinsic to the video game environment promote persistence of game play. These findings have important implications for the understanding and management of problem gaming behaviors and can help to inform ethical video game design and prevention of problematic play. Furthermore, methods used in this study can inform future investigations and help to address previous difficulties in studying video game behaviors experimentally (Wood et al., 2004) via remote administration emphasizing a realistic, externally valid gaming environment.

Outside of the context of video gaming, the current study emphasizes the more general role of reinforcement schedules and sensory feedback on preferred human activities. These results reinforce the idea that features inherent to the activities chosen motivate human behavior at least to some degree. This information is widely applicable for understanding and supporting current theories of human behavioral persistence despite lack of reinforcement, including gambling addiction and other compulsive behaviors.

References

Abrahamson, D. J., Hornyak, L. M., & Rehm, L. P. (2010). Understanding depression and effective treatment. *American Psychological Association*. Retrieved from http://www.apa.org/helpcenter/understanding-depression.aspx

Addis, M. E., & Mahalik, J. R. (2003). Men, masculinity, and the contexts of help seeking. *American Psychologist, 58*, 5-14. doi: 10.1037/0003-066X.58.1.5

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). https://doi.org/10.1176/appi.books.9780890425596

Batthyány, D., Müller, K., Benker, F., & Wölfling, K. (2009). Computer game playing: Clinical characteristics of dependence and abuse among adolescents. Wiener klinische Wochenschrift, 121(15–16), 502–509. https://doi.org/10.1007/s00508-009-1198-3

Catania, A. (2005). The operant reserve: A computer simulation in (accelerated) real time. Behavioural Processes, 69(2), 257–278. https://doi.org/10.1016/j.beproc.2005.02.009

Chan, P. A., & Rabinowitz, T. (2006). A cross-sectional analysis of video games and attention deficit hyperactivity disorder symptoms in adolescents. Annals of General Psychiatry, 5, Article 16. https://doi.org/10.1186/1744-859X-5-16

Chuang, Y.-C. (2006). Massively multiplayer online role-playing game-induced seizures: A neglected health problem in Internet addiction. Cyberpsychology & Behavior: The Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society, 9(4), 451–456. https://doi.org/10.1089/cpb.2006.9.451

Chumbley, J., & Griffiths, M. (2006). Affect and the computer game player: The effect of gender, personality, and game reinforcement structure on affective responses to computer game-play. Cyberpsychology & Behavior : The Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society, 9(3), 308–316. https://doi.org/10.1089/cpb.2006.9.308

Dworak, M., Schierl, T., Bruns, T., & Strüder, H. K. (2007). Impact of singular excessive computer game and television exposure on sleep patterns and memory performance of school-aged children. Pediatrics, 120(5), 978–985. https://doi.org/10.1542/peds.2007-0476

Entertainment Software Association. (2017). Essential Facts about the Computer and Video Game Industry. 2017 sales, demographic, and usage data. Retrieved from https://www.theesa.com/wp-content/uploads/2019/03/ESA_EssentialFacts_2017.pdf

Ferster, C., & Skinner, B. F. (1957). Schedules of Reinforcement. Appleton-Century-Crofts. https://doi.org/10.1037/10627-000

Fox, P. E., & Oakes, W. F. (1984). Learned helplessness: Noncontingent reinforcement in video game performance produces a decrement in performance on a lexical decision task. Bulletin of the Psychonomic Society, 22(2), 113–116. https://doi.org/10.3758/BF03333778

Greville, W. J., Newton, P. M., Roche, B., & Dymond, S. (2013). Conditioned suppression in a virtual environment. Computers in Human Behavior, 29(3), 552–558. https://doi.org/10.1016/j.chb.2012.11.016

Griffiths, M. D., Davies, M. N. O., & Chappell, D. (2004). Demographic factors and playing variables in online computer gaming. Cyberpsychology & Behavior: The Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society, 7(4), 479–487. https://doi.org/10.1089/cpb.2004.7.479

Griffiths, M. D., Kuss, D. J., & King, D. (2012). Video game addiction: Past, present and future. Current Psychiatry Reviews, 8(4), 308–318. https://doi.org/10.2174/157340012803520414

Hopson, J. (2001, April 27). Behavioral Game Design. Gamasutra: The Art & Business of Making Games. https://www.gamasutra.com/view/feature/131494/behavioral_game_design.php

Jeong, E. J., & Kim, D. H. (2011). Social activities, self-efficacy, game attitudes, and game addiction. Cyberpsychology, Behavior and Social Networking, 14(4), 213–221. https://doi.org/10.1089/cyber.2009.0289

King, D., & Delfabbro, P. (2009). Understanding and assisting excessive players of video games: A community psychology perspective. The Australian Community Psychologist, 21(1), 62-74.

King, D., Delfabbro, P., & Griffiths, M. (2010a). Video game structural characteristics: A new psychological taxonomy. International Journal of Mental Health and Addiction, 8(1), 90–106. https://doi.org/10.1007/s11469-009-9206-4

King, D., Delfabbro, P., & Griffiths, M. (2010b). The role of structural characteristics in problem video game playing: A review. Cyberpsychology: Journal of Psychosocial Research on Cyberspace, 4(1), Article 6. Retrieved from https://cyberpsychology.eu/article/view/4229/3272

King, D., Delfabbro, P., & Griffiths, M. (2011). The role of structural characteristics in problematic video game play: An empirical study. Int J Ment Health Addiction, 9(1), 320–333. https://doi.org/10.1007/s11469-010-9289-y

King, D., Koster, E., & Billieux, J. (2019). Study what makes games addictive. Nature, 573(7774), 346. https://doi.org/10.1038/d41586-019-02776-1

Klemm, C., & Pieters, W. (2017). Game mechanics and technological mediation: An ethical perspective on the effects of MMORPG's. Ethics and Information Technology, 19(2), 81–93. https://doi.org/10.1007/s10676-017-9416-6

Kuss, D. J. (2013). Internet gaming addiction: Current perspectives. Psychology Research and Behavior

Management, 6(1), 125–137. https://doi.org/10.2147/PRBM.S39476
Lemmens, J. S., Valkenburg, P. M., & Peter, J. (2011). Psychosocial causes and consequences of pathological gaming. Computers in Human Behavior, 27(1), 144–152. https://doi.org/10.1016/j.chb.2010.07.015
Loftus, G. R., Loftus, D., & Loftus, E. F. (1983). Mind At Play. Basic Books.
Peng, W., & Liu, M. (2010). Online gaming dependency: A preliminary study in China. Cyberpsychology, Behavior and Social Networking, 13(3), 329–333. https://doi.org/10.1089/cyber.2009.0082
Raia, C. P., Shillingford, S. W., Miller, H. L., Jr, & Baier, P. S. (2000). Interaction of procedural factors in human performance on yoked schedules. Journal of the Experimental Analysis of Behavior, 74(3), 265–281. https://doi.org/10.1901/jeab.2000.74-265
Skoric, M. M., Teo, L. L. C., & Neo, R. L. (2009). Children and video games: Addiction, engagement, and scholastic achievement. Cyberpsychology & Behavior: The Impact of the Internet, Multimedia and Virtual Reality on Behavior and Society, 12(5), 567–572. https://doi.org/10.1089/cpb.2009.0079
Wood, R. T. A., Griffiths, M., Chappell, D., & Davies, M. N. O. (2004). The structural characteristics of video games: A psycho-structural analysis. CyberPsychology & Behavior, 7(1), 1–10. https://doi.org/10.1089/109493104322820057
World Health Organization. (2019). International statistical classification of diseases and related health problems (11th ed.). https://icd.who.int/
Yee, N. (2006). The Psychology of MMORPGs: Emotional Investment, Motivations, Relationship Formation, and Problematic Usage. In R. Schroeder & A. Axelsson (Eds.), Avatars at Work and Play: Collaboration and Interaction in Shared Virtual Environments (pp. 187-207). London: Springer-Verlag.
Yellowlees, P. M., & Marks, S. (2007). Problematic Internet use or Internet addiction? Computers in Human Behavior, 23(3), 1447–1453. https://doi.org/10.1016/j.chb.2005.05.004
Zamani, E., Kheradmand, A., Cheshmi, M., Abedi, A., & Hedayati, N. (2010). Comparing the social skills of students addicted to computer games with normal students. Addiction & Health, 2(3–4), 59–65.

Author Note
James V. Lupo is now deceased.
McKenzie B. Wallace is now at the University of Missouri School of Medicine, Columbia, MO.
Supervision of the writing and revision process for this manuscript was done by Dr. Matthew T. Huss on behalf of Dr. James Lupo. We have no known conflict of interest to disclose.

Correspondence concerning this article should be addressed to McKenzie B. Wallace, University of Missouri – Columbia School of Medicine, 1 Hospital Dr, Columbia, MO 65212, United States. Email: mbwgbm@health.missouri.edu

Journal of Psychological Inquiry
2021, Vol.25, No. 2, pp. 55-60

Special Feature
Evaluating Controversial Issues

NEGATIVE AND POSITIVE IMPLICATIONS OF CO-PARENTING AFTER DIVORCE

TAYLOR D. GIPE, RACHEL G. KUNKEL, AND SARA E. BRADY

CONCORDIA UNIVERSITY

Abstract – Although previous research has extensively explored the negative aspects divorce has on children's well-being, few studies have examined the psychological influences of co-parenting for both children and parents. After a divorce, co-parenting is the reciprocal process in which both parents influence their children's education, childrearing, and life planning (Diogo et al., 2016). This review paper discusses the controversies behind the positive and negative implications of co-parenting for children and parents. For the negative implications, issues arise regarding violence between co-parents, mental health of children, and court disputes. For the positive implications, co-parenting can facilitate forgiveness and cooperation, promote children's psychological well-being, and diminish conflict exposure for children.

In 2019, 45 states reported to the National Vital Statistics System a total of 782,038 divorces that occurred within the United States (Centers for Disease Control and Prevention [CDC], 2020). According to the CDC, for every 1,000 people, there are 6.1 marriages and 2.7 divorces in the United States. While divorce is quite common, the question is whether or not to co-parent if children are involved in the family. There are several negative and positive viewpoints on whether co-parenting is beneficial to the child and the parents involved. According to Diogo et al. (2016), co-parenting is defined by "the reciprocal and conjoint involvement of both parents in education, childrearing, and planning of children's life decisions" (p. 717). This paper will discuss the positive and negative implications of co-parenting after divorce. The document will first examine the negative implications that can come from co-parenting which involve the continued violence or discord that can be present after a divorce between parents (Hardesty et al., 2016; Hardesty et al., 2017). In addition, negative implications for children will also be discussed (O'Hara et al., 2019; Parry et al., 2020; Wolchik et al., 2013), as well as issues regarding custody (Donner, 2006; Sbarra & Emery, 2005). This paper will then discuss the positive influence that co-parenting can have on the children and adults involved. The aspect of how forgiveness and cooperation between co-parents positively influences the child will be discussed (Yárnoz-Yaben et al., 2016). While there are negative mental health aspects that are discussed, there are also plenty of positive psychological and behavioral implications (Diogo et al., 2016; Feinberg et al., 2010). Finally, the idea of lower levels of conflict due to cooperative co-parenting will also be examined (Diogo et al., 2016).

Negative Implications of Co-Parenting

Premise

When two people get married, they seldom go into the marriage with the intent to get divorced. Divorce can be very draining on a person physically, mentally, and financially (Hardesty et al., 2016; Hardesty et al., 2017; Murray et al., 2016; Wolchik et al., 2013). Now add a child into the situation and it can become even more complicated. Families try to keep that aspect of normalcy by engaging in co-parenting. Many unforeseen challenges can come from co-parenting that may negatively affect the child and the relationship between the co-parents. There are reasons why co-parenting can negatively influence everyone involved including continued violence and disputes among the co-parents (Hardesty et al., 2016), negative mental health and behavior implications on the child (Wolchik et al., 2013), and ongoing custody disputes that can put strain on all involved (Donner, 2006).

Violence Between Co-Parents

Divorces occur for several reasons ranging from infidelity to different abuse that can occur. Marital violence is a big issue that often goes without being talked about due to those enduring the violence being fearful of judgment from friends, family, and society (Murray et al., 2016). After a couple has divorced, however, there can still be discord and violence that occurs in co-parenting. This not only affects the parents, but also the children. Intimate partner violence (IPV) is something that is still common even after a divorce, especially in divorces that have high amount of conflict already (Hardesty et al., 2016; Hardesty et al., 2017). Those who experienced IPV during the marriage also reported continued harassment even after the divorce when trying to co-parent (Murray et al., 2016). Often, nonphysical abuse, such as harassment, emotional abuse, and verbal abuse, was what one parent would endure after divorce due to a lack of physical access because the co-parents were no longer in physical proximity of each other (Hardesty et al., 2017). Specifically, mothers would report that there were higher levels of co-parenting conflict, higher levels of harassment after a divorce while co-parenting, lower levels of co-parenting support, and lower levels of communication about child rearing with the opposite parent (Hardesty et al., 2016; Hardesty et al., 2017). There is often conflict and discord among the parents. Sometimes this conflict and discord manifests in the child's surrounding, poor cooperation between parents, disengagement over childrearing issues, lack of involvement on parenting the child, and increased child-parent conflict (Bornovalova et al., 2013; Parry et al., 2020). When parents of a child engage in some sort of discord or violence, this often has a lasting influence on a child's mental health and behavior. Children who are exposed to violence between co-parents are also at an increased risk of emotion dysregulation, externalizing and aggressive behavior, and internalizing behavior (Murray et al., 2016). Externalizing behavior can be classified as problems manifesting as negative outward behavior and can include delinquency, temper tantrums, physical aggression, etc. Internalizing behavior can be classified as problems manifesting as negative psychological and emotional states and can include depression, anxiety, etc. Along with these increased risks for negative behavior, the child can also perceive discord between parents as the parents being unable to communicate and support the child in a way that fulfills the child's basic need of belonging and security (Parry et al., 2020).

Mental Health Implications on the Child

People commonly think about co-parenting as a good thing for a child because their parents are still working together, yet there can be a lot of negative mental health and behavioral consequences that could arise (Parry et al., 2020; Wolchik et al., 2013). After a divorce occurs, it is common for things to go well for the family and people start to get along. However, as time passes the relationships change and it can lead to different problems as the child grows up in terms of the child potentially facing mental health problems, substance use, onset of sexual activity, and even physical health problems (Wolchik et al., 2013). Some of the factors that can influence whether or not a child may develop any mental health problems, such as anxiety and depression, involve examining the conflict between the parents, the number of nights the child spends at each place, the quality of time spent with each parent, the parent's coping strategies, and the child's coping strategies (O'Hara et al., 2019; Parry et al., 2020; Wolchik et al., 2013). Quality time being spent with parents includes positive interactions between the parent and child that instills a sense of support, connection, and so much more. If this is lacking, children could be at risk for developing mental health problems that they are unable to handle (O'Hara et al., 2019; Parry et al., 2020; Wolchik et al., 2013). One of the big things that is important for children to understand and master are coping mechanisms. Many children are unable to cope with difficult situations right away due to not having the tools to properly process adverse experiences, which can lead to varying negative mental health implications for the child, such as anxiety, depression, and other mental disorders. Divorce and co-parenting can also have negative behavioral consequences for the child involved. There are many negative externalizing behaviors experienced by the child that can be attributed to divorce, lack of control, and evolving parenting instruction. Some of the most common negative behaviors that are associated with divorce are substance use (e.g., alcohol, nicotine, and drugs), temper tantrums, destroying things, lying, stealing, and different sexual actions (Parry et al., 2020; Wolchik et al., 2013). Often, these behavioral and mental health problems occur after overnight visits with each parent, especially if there were any conflicts between child and parent (O'Hara et al., 2019). Since children have a harder time coping with different situations, emotions are often shown through actions (O'Hara et al., 2019). It is very important for parents to recognize different changes within their children, because often there are lasting implications that could carry over into adulthood and manifest in their child's relationships.

Court Disputes

After a divorce, a custody battle of the child takes place. Though only one custody battle takes place the majority of the time, sometimes multiple battles take place if the parents choose to refile for custody or if a substantial change in circumstances occurs after the initial decision is made regarding custody. Even if the family decides to co-parent, often there will be some sort of court or custody dispute that will still occur. Co-parenting can cause those negative implications if one of the parental figures ultimately wants to gain custody over the child. Many of the conflicts involved in co-parenting surround the child custody battle that usually ensues (Donner, 2006; Sbarra & Emery, 2005; Sbarra & Emery, 2008). During child custody disputes, different feelings often arise, and the dispute becomes less about the child and their well-being, and more about expressing their anger and aggression toward the opposite parent which ultimately hurts the child (Donner, 2006; Sbarra & Emery, 2005). Another reason why custody battles tend to take place is due to the parent's own need to "win" and can often be used after a divorce as a way to emotionally contest the ending of a marriage (Donner, 2006; Sbarra & Emery, 2008). Custody battles tend to become difficult and there can be negative implications that come from it that affect both the child and the parents if a decision is made that does not best benefit the child (Donner, 2006). If a family is trying to co-parent, it is important to not look at it as an opportunity to fight for custody, but to look at what is best for the child's well-being.

Conclusion

Co-parenting sounds like it would be a good way to keep the relationship between child and parents in-tact, but it is important to understand that there are also plenty of negative consequences that can come from engaging in co-parenting. Often divorces occur due to some sort of marital discord, but that conflict also tends to continue after a divorce and into the co-parenting relationship. There tends to be frequent conflicts between the parents whether, it is due to communication issues, differing opinion on treatment of the child, or even physical violence (Bornovalova et al., 2013; Parry et al., 2020). This can influence the child in more ways than one. Mental health and behavior issues are very important consequences that can come from divorce and the co-parenting relationships. Since children do not have very good coping mechanisms, problems will often arise in their mental health. Some of the main issues that may arise include internalizing factors, which include depression and anxiety. There are also externalizing behaviors that often occur such as temper tantrums, substance abuse, or different sexual activities (Wolchik et al., 2013). These problems can even follow the child into adulthood. The final negative implication of divorce and co-parenting is regarding different custody battles that may ensue. While it may start out with both parents being supportive of each other, often one of the parents will try to gain full custody, which causes issues between the parents and the children (Sbarra & Emery, 2005). It is very important to understand the negative consequences of co-parenting when deciding what is best for the child, parents, and relationships after a divorce.

Positive Implications of Co-Parenting

Premise

Co-parenting has significantly increased in recent years. It is important to note that the way a child is impacted by co-parenting is circumstantial. Some children are negatively impacted by the way that their parents co-parent. However, many laypersons don't realize the numerous benefits co-parenting has for children, especially those who have parents that work better apart. Forgiveness between co-parents has significant benefits for the children involved (Yárnoz-Yaben et al., 2016). Co-parenting also fosters a cooperative atmosphere where parents promote positive parenting and involvement for both parties (Diogo et al., 2016; Feinberg et al., 2010; Yárnoz-Yaben et al., 2016). Co-parenting can positively impact children psychologically and behaviorally, through internalizing and externalizing behaviors (Diogo et al., 2016; Feinberg et al., 2010). Children, whose parents separated due to discord, experience less conflict at home by witnessing less discord between co-parents due to physical proximity (Diogo et al., 2016). When co-parents are able to cooperate and have agreement in making decisions for their children, their children see more benefits than children whose parents don't agree.

Forgiveness and Cooperation

Forgiveness is a process that requires the passage of time in order to allow both partners' emotions to cool and the memory of the transgression to fade (Yárnoz-Yaben et al., 2016). Forgiveness is often a process that co-parents must go through to best benefit their relationship between their child and the other co-parent. When parents experience a substantial amount of relational and marital discord, but continue to remain in the relationship, the home environment can become pretty hostile for the child. However, parents who choose to separate due to this discord, could be opening themselves up to forgiveness and closing themselves to rumination. Children will get the most out of their parent-child relationship when co-parents learn to forgive one another. The reappraisal of potentially painful and

stressful experiences through forgiveness is an opportunity for the co-parents to improve, grow, and render life meaningful and is associated with less depression, more positive emotions and relationships, self-esteem, optimism, and overall well-being (Yárnoz-Yaben et al., 2016).

When co-parents forgive each other for past transgressions, they are more likely to cooperate with one another when it comes to making decisions directly impacting their children (Yárnoz-Yaben et al., 2016). It is important to note that separation or divorce is more about dissolving the couple, not the family. Forgiveness helps divorced parents share their parental responsibilities and take care of their children after divorce (Yárnoz-Yaben et al., 2016). Cooperation between co-parents is one of the biggest factors that contributes the most to the children's harmonious development, even after the divorce or separation has occurred (Yárnoz-Yaben et al., 2016). Cooperation for childrearing has been found to be beneficial for children because parents are able to still uphold consistent set of rules and expectations (Yárnoz-Yaben et al., 2016). Even though the co-parents are separated or divorced, they are still able to keep a united front this way. Parents who are supportive of one another and agree on child-rearing tasks are actually more likely to model and assist children in emotion regulation, which is important for exploration and learning in social contexts (Cabrera et al., 2012). Buchanan and colleagues (1991) have found that adolescents whose parents were actively cooperating were less likely to feel "caught in the middle," mediate between parents, less likely to manifest negative adjustment outcomes, and less likely to experience loyalty dilemmas.

Co-parents who were able to forgive post-divorce was positively associated with trust and acceptance of the divorce (Kluwer et al., 2020). This suggests that co-parents were more likely to forgive and trust the other co-parent following acceptance of divorce. Following the results of their study, Kluwer et al. (2020) suggest that accepting the divorce is an important ingredient of post-divorce interventions. Co-parents who willingly chose to participate in intervention programs regarding forgiveness and cooperation to avoid putting their child in the middle of disputes between co-parents, were also more likely to forgive post-divorce (Kluwer et al., 2020).

Children's Psychological and Behavioral Benefits

When co-parents are able to cooperate and effectively communicate with one another, children are more likely to experience psychological and behavioral benefits. Some intervention programs that teach co-parents how to best serve their children have provided multiple indirect benefits to children. Feinberg et al. (2010) introduces Family Foundations, a transition to parenthood program for couples that focuses on the co-parenting relationship as a point of entry and has been shown to have a positive impact on families with an infant at 6 and 12 months after birth. Parents that participated in this study reported on their own behavior and the behavior of their children aged 6 and 12 months old. Children whose parents enrolled in this program demonstrated better social competency and boys evidenced lower levels of internalizing and externalizing problems (Feinberg et al., 2010). This could be due to the fact that parents of these infants in this program reported less harsh, physical, and overreactive parenting than control parents, while also simultaneously reporting less lax, permissive parenting (Feinberg et al., 2010). Competency, internalization, and externalization was measured by a mail-in questionnaire regarding the child's behavior completed by the mother of the child. An in-home visit also occurred when the children were 12 months old and another visit at 36 months old by a researcher to observe behavior of the child and parents (Feinberg et al., 2010). Diogo and colleagues (2016) also found that children of cooperative co-parents reported lower overall externalizing and internalizing problems. Programs like Family Foundations could help indirectly improve children's psychological and behavioral functioning by directly improving co-parent relationships.

Children with co-parents who are cooperative can experience a multitude of benefits, such as social competency, emotion regulation, and positive behavior (Cabrera et al., 2012; Diogo et al., 2016; Feinberg et al., 2010). Adolescents with cooperative parents significantly reported lower levels of behavior problems than did adolescents with undermining and high conflict co-parents (Diogo et al., 2016; Feinberg et al., 2010). Co-parents who communicate about their child on a regular basis and share in making decisions regarding children's health, nutrition, or childcare may signal to their children parental harmony and a working relationship, which can result in children having feelings of security and competence. Children feeling a sense of security and competence can positively affect their learning behavior (Cabrera et al., 2012). What many laypersons do not realize is how home life and relationships between parents can greatly affect school performance for their children. Cabrera and other researchers (2012) conducted a study where parent interviews and child assessments were taken at the 9-month, 24-month, and 48-month marks within the study. This sample included parents and their children aged from infancy to

kindergarten. Parents were asked to report their child's behavior regarding social, literacy, and math skills. According to Cabrera et al. (2012), mothers who reported talking frequently with their partners about their children had children who scored higher on school readiness skills. They also found that higher reports of co-parenting communication at the 24-month mark was related to their children scoring higher math scores, higher literacy scores, and better social skills than at the 9-month mark.

Less Conflict Exposure

Children whose parents get along better apart than together, can experience the benefit of less conflict exposure. A home environment where parents consistently do not get along is a very hostile and unstable environment for children to grow. Acrimonious parental interactions and disagreements about childrearing can be detrimental to young children's development (Cabrera et al., 2012). Children whose parents just do not get along when in a relationship, but can cooperatively co-parent from afar, reap the benefits of having a more stable home environment. Effective co-parenting relationships between ex-spouses may be favorable for children's psychological adjustment, since they reduce children's exposure to conflict, increase parents' cooperation in childrearing, promote positive parenting and involvement of the nonresidential parent, and contribute to positive parental psychological adjustment (Diogo et al., 2016). It is important to note that reduced conflict exposure is circumstantial. Children benefit more from less conflict exposure when parents put behind them past transgressions, effectively communicate, and cooperate compared to children whose parents still openly have conflict with one another. When contrasted with high conflict co-parent and undermining co-parent groups, cooperative co-parents reported significantly higher levels of post-divorce family functioning and low levels of both types of co-parenting conflict, overt (conflict exposure to the child) and covert (undermining the co-parent) (Diogo et al., 2016). As discussed in the section highlighting negative implications of co-parenting, intimate partner violence (IPV) can still occur after a divorce, the children in the family are not as exposed to the abuse between the parents due to the separation of the parents. Though it is not an ideal situation, whether or not the parents are together or apart, the child is still less exposed to the abuse and the separation gives the co-parents a chance to start working separately on their own negative behavior in order to benefit the child. As stated previously, it is important to note that reduced conflict exposure is circumstantial due to the risk of abuse still being present. However, having co-parents that do not openly and consistently have conflict with one another improves child development (Diogo et al., 2016).

Conclusion

In conclusion, there can be a lot of benefits that children experience when parents decide to co-parent. When children have cooperative co-parents, they can experience more benefits than just having two birthday cakes and Christmases. They can experience better school performance including better social competency, math skills, literacy skills, etc. when their parents take the initiative to become cooperative parents (Cabrera et al., 2012; Feinberg et al., 2010). Children of co-parents can experience emotional stability, less conflict exposure, modeling of forgiveness and effective communication, and emotional regulation as a result of cooperative co-parenting as well. Instilling intervention programs, similar to Family Foundations, could help co-parents find a way to be cooperative ones, so children of co-parents can all experience these benefits (Feinberg et al., 2010). The positive shaping of the child's life should be the most important part of parenting regardless if the parents are together or apart. Though the children's parents may be apart, they are still very much a part of their child's lives. Co-parenting is not the end of a family, it's just an adjustment of one.

Discussion

As discussed throughout this paper, there are many positive implications that can come from co-parenting after a divorce. Benefits of co-parenting include higher levels of cooperation and forgiveness among parents after a divorce (Yárnoz-Yaben et al., 2016), positive psychological and behavioral health among the children involved in co-parenting, and less conflict between families that use cooperative co-parenting (Diogo et al., 2016). However, there are also plenty of negative implications to look at before deciding to co-parent including high levels of conflict and discord among the parents after a divorce (Hardesty et al., 2017), negative mental health and behaviors that can come out of co-parenting due to lower levels of coping skills (O'Hara et al., 2019), and the possibility of different custody disputes that often take place after a divorce (Donner, 2006; Sbarra & Emery, 2005). The research at hand shows the negative and positive impact that can come from co-parenting after divorce. When a family is deciding to get divorced, it is important for the family to understand both the benefits and negative implications that can come from co-parenting. When considering whether or not to co-parent, it is important to make sure that the child's well-being is the main focus in whatever decision is being made.

References

Bornovalova, M. A., Blazei, R., Malone, S. H., McGue, M., & Iacono, W. G. (2013). Disentangling the relative contribution of parental antisociality and family discord to child disruptive disorders. Personality Disorders: Theory, Research, and Treatment, 4(3), 239–246. https://doi.org/10.1037/a0028607

Buchanan, C. M., Maccoby, E. E., & Dornbusch, S. M. (1991). Caught between parents: Adolescents' experience in divorced homes. Child Development, 62(5), 1008-1029. https://doi.org/10.2307/1131149

Cabrera, N. J., Scott, M., Fagan, J., Steward, S. N., & Chien, N. (2012). Coparenting and children's school readiness: A mediational model. Family Process, 51(3), 307–324. https://doi.org/10.1111/j.1545-5300.2012.01408.x

Centers for Disease Control and Prevention (2020, May 5). Marriages and divorces. National Center for Health Statistics. https://www.cdc.gov/nchs/fastats/marriage-divorce.htm

Diogo, L., Figueiredo, B., Bastos, A., & Feinberg, M. (2016). Typologies of post-divorce co-parenting and parental well-being, parenting quality and children's psychological adjustment. Child Psychiatry & Human Development, 47(5), 716–728. https://doi.org/10.1007/s10578-015-0604-5

Donner, M. B. (2006). Tearing the child apart: The contribution of narcissism, envy, and perverse modes of thought to child custody wars. Psychoanalytic Psychology, 23(3), 542–553. https://doi.org/10.1037/0736-9735.23.3.542

Feinberg, M. E., Jones, D. E., & Kan, M. L. (2010). Effects of family foundations on parents and children: 3.5 years after baseline. Journal of Family Psychology, 24(5), 532–542. https://doi.org/10.1037/a0020837

Hardesty, J. L., Crossman, K. A., Khaw, L., & Raffaelli, M. (2016). Marital violence and co-parenting quality after separation. Journal of Family Psychology, 30(3), 320–330. https://doi.org/10.1037/fam0000132

Hardesty, J. L., Ogolsky, B. G., Raffaelli, M., Whittaker, A., Crossman, K. A., Haselschwerdt, M. L., Mitchell, E. T., & Khaw, L. (2017). Coparenting relationship trajectories: Marital violence linked to change and variability after separation. Journal of Family Psychology, 31(7), 844–854. https://doi.org/10.1037/fam0000323

Kluwer, E. S., van der Wal, R. C., Visser, M., & Finkenauer, C. (2020). Predictors of forgiveness among divorced parents. Journal of Family Psychology. https://doi.org/10.1037/fam0000799.supp

Murray, C., Crowe, A., & Akers, W. (2016). How can we end the stigma surrounding domestic and sexual violence? A modified Delphi study with national advocacy leaders. Journal of Family Violence, 31(3), 271–287. https://doi-org.vproxy.cune.edu/10.1007/s10896-015-9768-9

O'Hara, K. L., Sandler, I. N., Wolchik, S. A., Tein, J.-Y., & Rhodes, C. A. (2019). Parenting time, parenting quality, interparental conflict, and mental health problems of children in high-conflict divorce. Journal of Family Psychology, 33(6), 690–703. https://doi.org/10.1037/fam0000556

Parry, L. Q., Davies, P. T., Sturge-Apple, M. L., & Coe, J. L. (2020). Co-parental discord and children's behavior problems: Children's negative family representations as an explanatory mechanism. Journal of Family Psychology, 34(5), 523–533. https://doi.org/10.1037/fam0000638

Sbarra, D. A., & Emery, R. E. (2005). Coparenting conflict, nonacceptance, and depression among divorced adults: Results from a 12-Year follow-up study of child custody mediation using multiple imputation. American Journal of Orthopsychiatry, 75(1), 63–75. https://doi.org/10.1037/0002-9432.75.1.63

Sbarra, D. A., & Emery, R. E. (2008). Deeper into divorce: Using actor-partner analyses to explore systemic differences in co-parenting conflict following custody dispute resolution. Journal of Family Psychology, 22(1), 144–152. https://doi.org/10.1037/0893-3200.22.1.144

Wolchik, S. A., Sandler, I. N., Tein, J.-Y., Mahrer, N. E., Millsap, R. E., Winslow, E., Vélez, C., Porter, M. M., Luecken, L. J., & Reed, A. (2013). Fifteen-year follow-up of a randomized trial of a preventive intervention for divorced families: Effects on mental health and substance use outcomes in young adulthood. Journal of Consulting and Clinical Psychology, 81(4), 660–673. https://doi.org/10.1037/a0033235

Yárnoz-Yaben, S., Garmendia, A., & Comino, P. (2016). Looking at the bright side: Forgiveness and subjective well-being in divorced Spanish parents. Journal of Happiness Studies, 17(5), 1905–1919. https://doi.org/10.1007/s10902-015-9677-x

Author Note
Correspondence concerning this article should be addressed to Tayler D. Gipe, Concordia University, Nebraska. Email: tayler.gipe@cune.org

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 will be accepted 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 versus 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.
- 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: Is homosexuality incompatible with military service?
        - 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.