EXPLORING THE EFFECT OF SOCIAL MEDIA POPULARITY METRICS ON CURIOSITY

by

ERICA WALDRON

A THESIS

Presented to the Department of Psychology
and the Robert D. Clark Honors College
in partial fulfillment of the requirements for the degree of
Bachelor of Science

May 2022

An Abstract of the Thesis of

Erica Waldron for the degree of Bachelor of Science
in the Department of Psychology to be taken May 2022

Title: Exploring the Effect of Social Media Popularity Metrics on Curiosity

Approved: *Dare Baldwin, PhD*
Primary Thesis Advisor

Social media inundates us with information about popularity; for example, social media posts are accompanied by a count of "likes" as well as comments. Dubey and colleagues (2020) recently demonstrated that such indicators of popularity influence people's curiosity to learn more about specific topics. If so, this is one unexpected, beneficial side effect of social media popularity metrics. However, the way in which they manipulated popularity via Reddit-like upvotes may have inadvertently introduced a confound into their findings. Specifically, people were asked to report about an item's popularity immediately before reporting on their curiosity regarding that item. The immediate juxtaposition of these two questions may have led participants to assume that popularity was relevant to curiosity, thereby creating a demand characteristic that could have contaminated their findings. My thesis research attempts to replicate that of Dubey and colleagues' while avoiding this potential demand characteristic. People rated curiosity first and were asked about popularity only at the end of the survey. Analyses modeled after Dubey et al. (2020) indicate that I partially replicated their findings. That is, when accuracy in recalling item popularity is considered, people are indeed more curious about items with a high number of upvotes than those with a low number of upvotes. These findings confirm that indicators of popularity can elicit curiosity, which sets the stage for deploying popularity as a curiosity-trigger in a range of possible contexts, such as in curiosity research and educational settings.

Acknowledgements

I would like to thank my thesis committee for their support and expertise throughout this process. I am immensely grateful to my primary thesis advisor Dr. Dare Baldwin for introducing me to the topic of curiosity and spending countless hours guiding me through this research. Thank you to Dr. Jennifer Ablow, my second reader, for inspiring me to get involved in developmental psychology research and offering her expert input on this document. I would also like to extend my sincere gratitude to my third reader, Doctoral Student Diana DeWald, for her assistance and taking the time to serve on my committee. Finally, I would like to extend special thanks to my family for their endless encouragement throughout my college career.

Table of Contents

List of Figures

List of Tables

Introduction

Curiosity's status as an elusive, difficult to define concept cements its position as an area of great interest to those within the realm of psychology and beyond. As a concept that continually influences the way we interact with the world, researchers have conducted a wide range of studies to try and parse out curiosity. These investigations into curiosity have covered a variety of avenues, including how to spark curiosity, the purpose it serves for humans, and curiosity's neural correlates. Decades of research have resulted in a general consensus that curiosity has the ability to influence people's behaviors and the manner in which they process the world around them (Loewenstein, 1994; Schmitt & Lahroodi, 2008; Kang et al., 2009; Kidd & Hayden, 2015). Despite this consensus, the specifics of how curiosity affects attentional processing has been largely left unexplored. How does our degree of curiosity determine how we choose to focus on one idea over the other? Does our curiosity depend on what we think other people find interesting? This gap in knowledge necessitates more experimental investigation into the particularities of how curiosity might alter the amount of overall attention focused on an event and influence which portions of information are attended to over others. The results of this research could have future implications concerning the ways that we think and learn. In the world of education, this type of research could allow educators to harness curiosity for students' benefit. By implementing curiosity-inducement techniques, educators could potentially draw attention to specific subjects, promote students' exploration of new topics, and generally enhance their learning experience.

Curiosity

Curiosity is a contentious topic within the scientific community, evident in the fact that a single, widely accepted definition of this concept has yet to emerge. Several theories have attempted to classify what does and does not fit under the umbrella of curiosity. One account of curiosity posits that it is not all that different from the concept of interest (Kashdan & Silvia, 2009). Kashdan and Silvia (2009) assert that there is not a large differentiation to be made between curiosity and interest because both act as sources of intrinsic motivation for exploratory action. Another theory classifies curiosity as a "state of cognitively induced deprivation" concerning knowledge (Loewenstein, 1994). According to Lowenstein (1994), curiosity is what occurs when a person is confronted by an information disparity, consciously recognizes this gap in their knowledge, and is subsequently intrinsically motivated to close this gap. Additionally, Kidd and Hayden (2015) conceptualize curiosity as a generalized, information-seeking behavior. Under this framework, curiosity is more loosely defined as a drive-state for information, free from comparisons to other drive-states, such as interest, or specificity of motivational origins (Kidd & Hayden, 2015). For the purposes of the research at hand, this final definition of curiosity was adopted.

Despite the lack of consensus over the exact definition of curiosity, it is widely agreed upon that it is a fundamental motivational state associated with positive behaviors. Researchers have found curiosity to be an important aspect of education and learning, with curiosity acting as a predictor of academic achievement (Dubey & Griffiths, 2020; Kidd & Hayden, 2015; Loewenstein, 1994). Along this same line, curiosity has even been found to be associated with an improvement in memory recall when it comes to remembering novel information (Kang et al., 2009). Additionally, curiosity is thought to promote typical cognitive development and exploratory behaviors in infants by aiding them in effectively deploying their limited attentional resources (Kidd & Hayden, 2015; Kashdan & Silvia, 2009). Past research has also explored the neural mechanisms and visible behaviors prompted by curiosity's presence.

Some studies have delved into the manners in which curiosity shapes our neural activity, physiological reactions, and processing of information or events. One such study was performed by Kang and colleagues. Kang et al. (2009) sought to explore the behavior of brain activity and pupil dilation in the presence of curiosity through the utilization of fMRI and pupillometry technology. In order to induce curiosity, Kang et al. (2009) exposed participants to various trivia questions meant to introduce an information gap and subsequent drive for knowledge. Participants were then asked to self-rate their curiosity of answers to the questions while undergoing fMRI. Results of their first study reveal that brain activity within the caudate and left pre-frontal cortex, regions associated with reward, was correlated with levels of curiosity (Kang et al., 2009). Kang and colleagues' second study utilized an eye tracker to record pupil dilation responses before and after the reveal of the questions' answers. Pupil dilation was found to be positively correlated with levels of curiosity, with high levels of curiosity showing the largest pupil dilation response just prior to the presentation of an answer (Kang et al., 2009). This realm of research provides a basis for the assumption that curiosity's effect on processing can be visibly recorded and measured. Another potential vehicle for measuring how curiosity might reorganize attentional patterns and processing is the dwell-time paradigm.

The Dwell-Time Paradigm

The dwell-time paradigm was first introduced by Hard and colleagues in their 2011 study investigating how attention is deployed to organize streams of activity. This mode of measurement involves advancing through self-paced slideshows, constructed of slides extracted from a video at a regular frames per second interval, while recording the time spent on each slide. Hard et al.'s (2011) initial study utilized slideshows that depicted large-scale everyday activities such as making a bed. The results of their study revealed a pattern of sensitivity to event structure wherein dwell-times surged at boundary slides, the frames which depict the end of an action or the beginning of another. Subsequent studies utilizing the dwell-time paradigm have both replicated Hard and colleagues' original pattern of dwelling and discovered new findings of interest.

Sage and Baldwin's 2014 study explored whether the original dwell-time patterns would also be found with slideshows depicting small-scale action sequences, such as sleight-of-hand tricks. In addition to replicating Hard et al.'s pattern of surges in dwell-times at breakpoints in action, Sage and Baldwin (2014) found that participants paid more attention to frames depicting causally relevant information about how the sleight-of-hand maneuver was performed, reflected in increases in dwell-time at these points. Another study conducted by Kosie and Baldwin in 2018 utilized the dwell-time paradigm to determine how novelty affects processing and whether multiple viewings of new information would change dwell-time patterns. The slideshows in this study respectively depicted the typical way of tying shoelaces with loops and a novel way involving twists. Similar to previous studies, increases in dwell-time were found at the boundary points in both shoelace tying event streams. Notably, new attentional reorganization patterns emerged as well. On the first viewings of both methods, dwell-times increased for slides depicting content that distinguished the methods from one another (i.e., the use of loops versus twists). However, upon the second viewings, attention to these distinctive regions remained high, and actually increased, only for the slideshow depicting the novel twist method. Kosie and Baldwin (2018) suggest this pattern emerges as a result of participants identifying the portions of the slideshow that are critical to learning more about the novel method and correspondingly adjusting their attentional resources. Each of these studies provide information regarding the regions within slideshows that draw the most attention and the subsequent expected dwell-time patterns.

Dubey, Mehta, and Lombrozo's 2020 Study

One study that has explored the motivators of curiosity while viewing curiosity itself under the definition of a drive state is Dubey et al.'s 2020 work. Dubey et al.'s (2020) study examined how a social intervention involving perceived popularity can induce different levels of curiosity. Their study utilized different amounts of Reddit upvotes, a count that signifies how many people approve of or support a post, in order to provide a cue about social popularity to participants while they were presented with questions about science. Participants were also asked to rate their curiosity about the answers to these questions. Results of this study show that participants rated the questions with a high number of upvotes as more popular and self-reported higher levels of curiosity for those same questions when compared with questions that had lower numbers of upvotes. Dubey et al.'s (2020) findings also reveal that participants were more likely to choose to reveal the answers to questions originally accompanied by a high number of upvotes than a low number. Additionally, through the comparison to a no upvote, baseline condition, their results indicate that popularity may primarily act as a suppressor of curiosity when it comes to the answers to less popular questions rather than as a motivator of curiosity regarding the more popular ones.

Dubey and colleagues also examined the extent to which other factors may have mediated the upvotes' effect on curiosity. They accomplished this by asking participants to rate their surprise concerning question popularity, confidence in knowing the question's answer, and their opinions about the social utility and future usefulness of having the answer to the question. Results showed that their popularity manipulation significantly affected ratings of surprise, social utility, and future usefulness, but not participant confidence (Dubey et al., 2020).

While Dubey et al.'s (2020) findings postulate new and fascinating information regarding the social influence on curiosity, their methods include a potential demand characteristic. After participants were shown a question paired with upvote information, they were immediately asked to rate how popular they believed the question to be on the social forum from which it was drawn. The very next question presented to participants had them rate their level of curiosity regarding the answer to the question. The placement of this popularity manipulation check just prior to the question ascertaining self-rated curiosity may have made participants explicitly privy to the researcher's aim of studying social popularity's effect on curiosity, leading to biased or skewed results. It is imperative to discover whether the findings of a significant curiosity manipulation are still present when this priming popularity information is not front and center.

Current Research

This study utilized the curiosity manipulation introduced by Dubey et al. (2020) paired with an updated version of their methods and the addition of the dwell-time paradigm to investigate their findings and further examine curiosity's effect on attentional processing. This study sought to examine the extent to which Dubey and colleagues' original pattern of findings replicated after the removal of the potential demand characteristic. This study was conducted with future work in mind that would examine ways in which upregulating curiosity might alter downstream dwell-time patterns for related content. While dwell time data were also collected, they will not be presented in this thesis due to time constraints.

I hypothesize that Dubey and colleagues' original patterns of findings will still be present even when controlling for the demand characteristic afflicting their findings. More specifically, I predict that a high number of upvotes will lead to higher ratings of curiosity and a low number of upvotes will lead to lower ratings of curiosity.

Method

Participants

I recruited 350 college aged adults through the University of Oregon's Human Subject Pool. Each participant received one hour of course credit as compensation for their participation in the survey. Participants had to complete the entire survey, or they would be eliminated from the data. These parameters excluded 72 participants from the dataset. Subsequently, this study examined the data of 278 participants ranging in age from 18 to 42 years old (M = 19.64, SD = 2.01). Out of the participants included, 75.90% of participants identified as Caucasian or White, 8.99% as Asian American or Asian, 2.88% as African American or Black, and 9.71% as other with 1.44% electing not to answer. Additionally, 65.11% of participants identified as women, 31.29% as male, 1.80% as gender fluid, and 0.72% as transgender, with 1.08% preferring not to provide their gender identity.

Materials

Replication Stimuli

Two slightly different editions of a survey were used to mount my replication efforts of Dubey and colleagues' original 2020 study. These differences will be expanded upon in the procedure.

This study utilized twenty why-, how-, or what- questions randomly selected from the question bank provided by Dubey et al. (2020) (See Table 1). Additionally, four new questions from the *Explain Like I'm Five* subreddit were utilized as stimuli. These questions were selected using Dubey et al.'s (2020) original criteria, in that they were moderately popular in the subreddit with 200-600 upvotes apiece. The four additional questions were:

1. Why don't penguins get cold feet?
2. Why do octopi not suffer brain damage when they squeeze through tiny openings?
3. Why does a medical doctor check your knee for reflexes?
4. Why do panda bears have fangs like carnivores when all they eat is bamboo?

Dwell-Time Stimuli

Four videos depicting information related to four of the questions were selected to act as the sources of dwell-time slideshows to be utilized in future research. The videos respectively depicted penguins interacting and moving around on the arctic ice, an octopus squeezing through a transparent pipe, a doctor performing a reflex test on a female patient, and a panda bear eating bamboo and then walking away. Still image slides were extracted from each video at a rate of two frames per second with the four complete slideshows ranging from 26 to 60 slides in length. Each slideshow was specifically formatted to include slides depicting both relevant and irrelevant information to their respective question's answer.

Procedure

Participants were first randomly assigned to either Edition 1 or Edition 2 of the survey. Regardless of survey edition, they then viewed a page explaining the origin of the study's stimuli. This page contained the verbatim explanation given to the Dubey et al. (2020) participants:

Upvote Condition:

> “On the following pages, we will show you 8 questions people have asked and a number of upvotes those questions got on a popular online forum. Note that the upvotes were given by members of the online community who viewed just the question, not the answer. The upvotes were based only on the questions and not the answers to those questions. For each question, we will ask you to make a series of judgements.”

Baseline/No Upvote Condition:

> “On the following pages, we will show you 8 questions people have asked on a popular online forum. For each question, we will ask you to make a series of judgements.”

Eight questions were then presented to participants one at a time, four of which were the questions with attached dwell-time stimuli. The other four questions shown to participants were randomly selected from the bank of twenty.

Each of the eight questions was accompanied by upvote information in the upvote condition. Due to the limitations of the survey platform, the extent of randomization within the surveys’ presentation of questions was limited. The dwell-time-related questions were always presented together as were the four questions from the bank, but the order in which these groups were presented was randomized. Two of the four dwell-time-related questions were randomly selected to be accompanied by a high amount of upvotes and two by a low amount. Similarly, two of the four questions drawn from the bank were accompanied by a high amount of upvotes and two by a low amount. Whether questions were accompanied by a low or high amount of upvotes depended on survey edition. In Edition 1 of the survey, as numbered in the bank shown in Table 1, questions 1-10 were always accompanied by a high amount of upvotes, and questions 11-20 by a low amount. In Edition 2 of the survey, questions 1-10 were always accompanied by a low amount of upvotes, and questions 11-20 by a high amount.

For those assigned to Edition 1 of the survey, when shown the dwell-time-related questions, participants first saw two questions accompanied by a low number of upvotes then the two accompanied by a high number. When shown the four questions randomly selected from the bank of questions, the first two seen were always accompanied by a high number of upvotes and the final two by a low number. Conversely, for those assigned to Edition 2 of the survey, when shown the dwell-time-related questions, participants first saw two questions accompanied by a high number of upvotes then the two accompanied by a low number. When shown the questions from the bank, the first two seen were always accompanied by a low number of upvotes and the final two by a high number. No upvote information was provided in the baseline condition. However, as in the upvote condition, the dwell-time-related questions were always presented together as were the random four questions, but the order in which these groups were presented was randomized.

Participants were asked to make a series of four judgements for each question on a scale from 0 (not at all) to 6 (very): curiosity ("How curious are you to know about the question and its answer?"), confidence ("How confident are you that you know the correct answer to this question?"), social utility ("To what extent would knowing the answer to this question be useful to you in a social setting?"), and usefulness ("To what extent would knowing the answer to this question be useful to you in the future?"). The judgement of surprise, prompted by the question "how surprised are you by the popularity of this question", was not included in my surveys due to its relationship to

Dubey et al.'s (2020) potential demand characteristic concerning the rating of popularity.

As in Dubey et al. (2020) study, participants then choose four of the eight questions they wanted to reveal the answers to. They were presented with all eight questions, with upvote information present if applicable for that condition, and were able to click "yes" for the four that they want to learn more about, and "no" for the remaining four. Participants were then told they would watch four slideshows related to some of the questions they just viewed. Each participant then dwelled through the same four slideshows. After viewing all the slideshows, participants completed a memory and attention check. This check consisted of four prompts that appeared in the slideshows with the same number of distractor items (See Table 2).

Following this memory check, participants were asked to rate the popularity of the eight questions they had been shown on a 7-point scale from 0 (not at all popular) to 6 (very popular). For this task, participants assigned to the upvote condition viewed the prompt: "To the best of your ability, please report from memory the popularity (the level of upvotes) associated with each item below" before assigning their ratings. Those in the baseline condition viewed the prompt: "Please report your guess of the popularity associated with each item below". This step reintroduced the popularity manipulation check that was originally placed at the outset of Dubey et al.'s (2020) study.

After rating popularity, participants filled out the Wender Utah Rating Scale meant to retroactively evaluate ADHD symptoms experienced in childhood (Ward, 1993). Participants were then asked to rate how well they adhered to all the instructions provided to them in the survey on a 5-point scale from "not at all" to "closely". Data collected from the Wender Utah Rating Scale and the adherence probe were not analyzed as they are relevant to dwell-time analyses and thus fall beyond the scope of this thesis. Next, participants reported on which type of device and web browser they utilized to complete the survey. After completing these steps, participants were presented with the answers to the questions they previously selected to have revealed. Once they finished reading the answers, participants were shown a page debriefing them on the study’s background and goals.

Results

I hypothesized that even with the removal of the potential demand characteristic, Dubey and colleagues' findings would still replicate. More specifically, I predicted that a comparison of ratings would reveal higher levels of curiosity for high upvote questions than for low upvote questions. In order to approach these predictions about replication, I first performed a series of tests similar to those run by Dubey and colleagues. Following these tests, I ran exploratory multiple regression analyses to test for the unique contribution of variables such as popularity rating accuracy, upvote type, confidence, future use, and social utility for predicting curiosity, popularity, and whether answers were chosen to be revealed.

Preliminary Analyses

Within the dataset analyzed, 136 participants were assigned to Edition 1 of the survey and 142 to Edition 2. Across both surveys, 143 participants were assigned to the upvote condition and 135 to the baseline condition. Due to the limitations of the survey-hosting platform, 47 participants were able to select either greater or fewer than four questions for which they later wanted to reveal the answers. Analyses were undertaken to test the extent to which survey edition and the number of answers chosen to be revealed affected ratings of popularity, curiosity, confidence, social utility, usefulness, and only in the case of survey edition, whether to reveal an answer. An analysis of variance revealed that there were no main effects or interaction effects of survey edition on any of the judgements made by participants. Therefore, all variables were collapsed across both survey editions for all future analyses. However, the number of answers participants chose to reveal did significantly affect curiosity ratings. In subsequent analyses comparing the effect of low versus high upvotes on participants' judgments, number of answers chosen was controlled for, and the pattern of findings did not ultimately change. Additionally, an error in the survey coding led to participants who were assigned to both Edition 2 of the survey and the upvote condition not being shown the popularity rating question for "Why do panda bears have fangs like carnivores when all they eat is bamboo?". As a result, 70 points of data were not included in analyses rated to the perceived popularity of upvote-accompanied questions. Given the sizable number of questions probed and the resulting large data set (1,053 datapoints), this minor implementation error resulted in a small number of missing datapoints which likely would have had minimal effect on analytic outcomes.

Replication Analyses

Curiosity Differences Between Upvote Type

I first tested whether upvote type influenced how participants rated their curiosity. A paired samples t-test revealed that there was no significant difference between the mean curiosity rating for high upvote-accompanied questions (M = 3.648, SD = 0.1.308) versus that of low upvote-accompanied questions (M = 3.616, SD = 1.301), t (141) = -.371, p = .711. These results failed to replicate Dubey and colleagues' original findings of an upvote-based curiosity difference.

Popularity Difference Between Upvote Type

Next, I examined the extent to which the presence of either a low or high amount of upvotes influenced perceived popularity when upvote information was not readily available. A paired samples t-test revealed that the mean popularity of questions with high upvotes (*M* = 3.482, *SD* = 0.938) was significantly higher than that of questions assigned a low number of upvotes (*M* = 3.025, *SD* = 0.918), $t(141) = -4.697$, $p < .001$. This suggests that Dubey and colleagues' popularity manipulation was still registered by participants even when they were probed without being able to readily reference upvotes.

Effect of Upvote Type on Answer Selection

I then assessed whether participants were more likely to choose to reveal the answers to questions accompanied by a high number of upvotes than those accompanied by a low number. A binomial test revealed that the answers to high upvote questions were only chosen to be revealed 50.6% of the time, which is not significantly different than the chance value of 50%. This finding fails to replicate the original results found by Dubey and colleagues regarding answer selection.

Upvote vs Baseline

Next, I conducted a series of independent samples t-tests to explore how judgments made by participants in the upvote condition differed from those in the baseline condition. For each judgement, I created two mean ratings, one formulated from questions accompanied by low upvotes, and another from those accompanied by high upvotes (See Table 3 and 4). These two sets of means were compared to the mean ratings of judgments gathered in response to the questions with no upvotes attached (i.e., baseline condition responses).

An independent samples t-test revealed that baseline curiosity ratings (*M* = 3.11, *SD* = 1.22) did not significantly differ from low upvote ratings (*M* = 3.51, *SD* = 1.86), *t*(205) = -1.88, *p* = .062. However, mean curiosity for high upvote questions (*M* = 3.61, *SD* = 1.79) was significantly higher than mean baseline curiosity, *t*(204) = -2.34, *p* = .001 (See Figure 1). These findings do not replicate those reported in Dubey et al.'s (2020) study, in which the mean rating for low upvote curiosity was found to be significantly lower than mean baseline curiosity, and mean rating for high upvote curiosity was higher than but not significantly different from baseline.

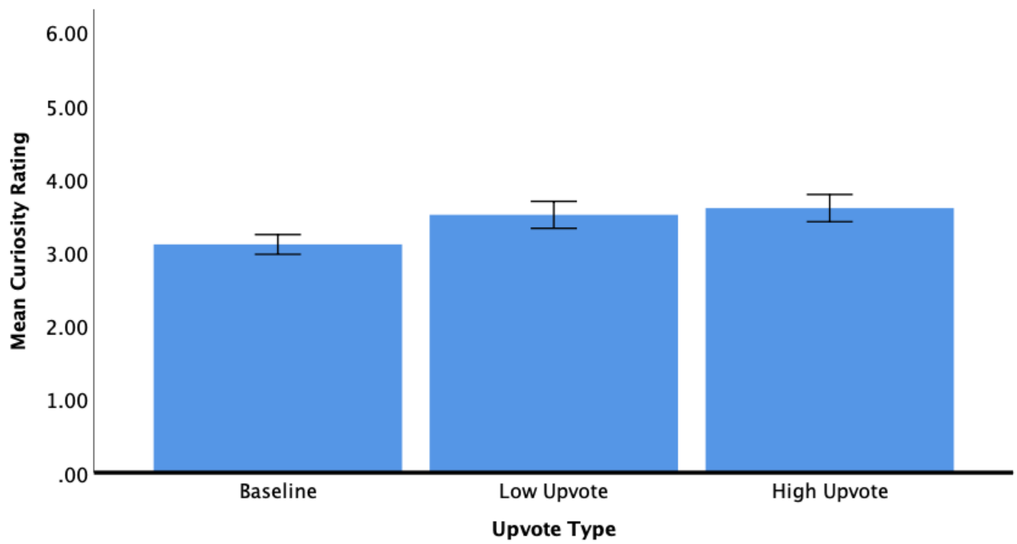

Figure 1: Mean Curiosity Ratings for Baseline, Low Upvote, and High Upvote Items

Note. Error bars: +/- 1 SE.

This trend of non-replication continued for the next three judgment ratings. Mean low upvote confidence was significantly higher than that of baseline confidence, *t*(205) = -4.78, *p* < .001, as was mean high upvote confidence, *t*(204) = -4.38, *p* < .001. Baseline ratings of social utility were not significantly different from low upvote social utility ratings, *t*(205) = -1.65, *p* = .100, nor from high upvote social utility ratings, *t*(204) = -1.56, *p* =.120. Similarly, baseline ratings of usefulness did not differ significantly from low upvote ratings, $t(205) = -1.27$, $p = .204$, nor from high upvote ratings, $t(204) = -1.92$, $p = .056$.

Baseline popularity ratings did not significantly differ from low upvote ratings, $t(205) = .326$, $p = .744$. However, popularity ratings of high upvote questions were significantly higher than those of baseline questions, $t(204) = -3.16$, $p = .002$. These popularity means were the only judgment ratings to replicate Dubey et al.'s (2020) finding of a baseline mean that falls between those of low upvote and high upvote questions. In general, baseline means fell lower than those calculated in response to low upvote questions, and quite a bit lower than those calculated from high upvote questions (See Figure 2).

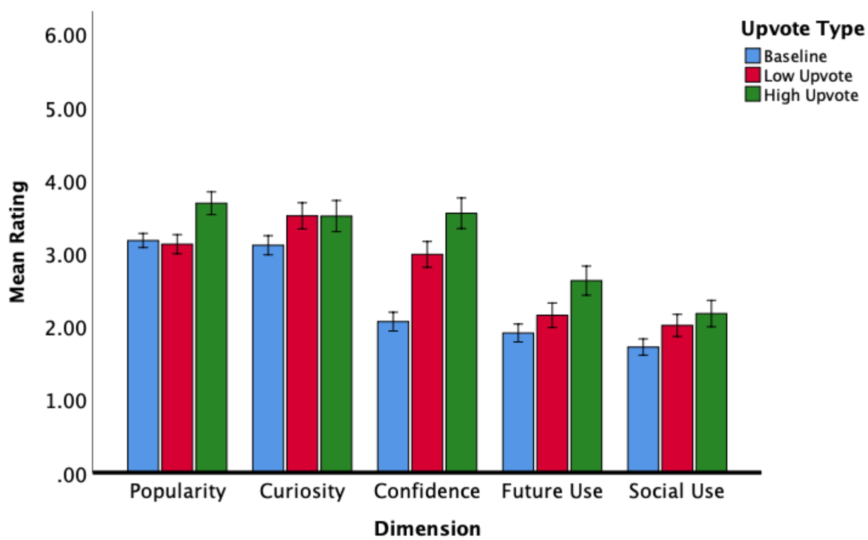

Figure 2: Comparison of Mean Judgment Ratings Across Upvote Types

Note. Error bars: +/- 1 SE.

This first grouping of replication analyses offers little support in the way of Dubey et al.'s (2020) original findings. While participants still registered the upvote manipulation, there was not a significant difference in curiosity levels between low upvote and high upvote questions. Additionally, participants were not more likely to want to reveal the answers to high upvote questions. Baseline judgments fell below those concerning low or high upvote questions, failing to replicate Dubey et al.'s (2020) pattern of findings.

Multiple Regression Analyses

I then conducted a series of multiple regressions to test for the different predictors of curiosity and whether answers were revealed in both the upvote and baseline conditions. Due to the nature of the baseline condition, upvote type was not included as a predictor in any of the baseline models. See Table 6 for the correlations between variables in the upvote condition and Table 7 for those in the baseline condition.

For the upvote condition, the regression showed that a model including upvote type, confidence, social utility, usefulness, and popularity was a significant predictor of curiosity, $F(5, 1067) = 80.53$, $R^2 = .25$, $p < .001$. While social utility ($b = .370$, $p < .001$), usefulness ($b = .141$, $p < .001$), and popularity ($b = .237$, $p < .001$) contributed significantly to the model, upvote type ($b = -.120$, $p = .207$) and confidence ($b = -.020$, $p = .440$) did not. The results of the regression for the baseline condition were quite similar, indicating that overall, the model was a significant predictor of curiosity, $F(4, 1073) = 89.20$, $R^2 = .25$, $p < .001$. Similar to the upvote results, social utility ($b = .273$, $p < .001$), usefulness ($b = .181$, $p < .001$), and popularity ratings ($b = .332$, $p < .001$) contributed significantly to the model, while confidence did not ($b = -.002$, $p = .952$).

When examining the contribution of variables for predicting whether an answer was revealed in the upvote condition, the regression demonstrated that a model including upvote type, confidence, curiosity, social utility, usefulness, and popularity was a significant predictor of answer selection, $F(6, 1066) = 37.47$, $R^2 = .174$, $p < .001$. While confidence ($b = -.023$, $p = .003$), curiosity ($b = .082$, $p < .001$), and popularity ($b = .055$, p < .001) contributed significantly to the model, upvote type ($b = -.010$, $p = .716$), social utility ($b = .007$, $p = .574$), and usefulness ($b = .016$, $p = .175$) did not. The regression for the baseline condition indicated that the model was a significant predictor of whether an answer was chosen, $F(5, 1072) = 39.14$, $R^2 = .154$, $p < .001$. While confidence ($b = -.021$, $p = .008$), curiosity ($b = .061$, $p < .001$), usefulness ($b = .025$, $p = .028$), and popularity ($b = .064$, $p < .001$) contributed significantly to the model, social utility did not ($b = .009$, $p = .492$).

Exploratory Analyses

I stepped beyond performing the same multiple regressions as Dubey et al. (2020) to test for the contribution of variables predicting perceived popularity. For the upvote condition, the regression demonstrated that a model including upvote type, confidence, curiosity, social utility, and usefulness was a significant predictor of popularity, $F(5, 1067) = 24.471$, $R^2 = .103$, $p < .001$. While upvote type ($b = .429$, $p < .001$) and curiosity ($b = .227$, $p < .001$) contributed significantly to the model, confidence ($b = .024$, $p = .344$), social utility ($b = .040$, $p = .357$), and usefulness ($b = .009$, $p = .826$) did not. The results of the regression for the baseline condition indicated that the model was a significant predictor of popularity, $F(4, 1073) = 51.22$, $R^2 = .257$, $p < .001$. While curiosity ($b = .278$, $p < .001$), social utility ($b = .273$, $p < .001$), and usefulness ($b = -.084$, $p = .020$) contributed significantly to the model, confidence did not ($b = .004$, $p = .859$).

I also examined whether a curiosity difference between upvote was present when participants' accuracy in rating question popularity was taken into account. To accomplish this, curiosity ratings given to questions accompanied by a low number of upvotes were reverse scored so that lower ratings were treated as more accurate. For high upvote items, higher ratings were treated as more accurate. Additionally, I computed a difference score between participants' curiosity ratings to high upvote questions vs low upvote questions. This allowed me to examine if accuracy in remembering which questions had a high number of upvotes and which had a low number predicted this difference score. A curve estimation regression revealed that a model including popularity accuracy was a significant predictor of curiosity ratings, ($b = .492$, $p = .001$), $F(1, 141) = 11.925$, $R^2 = .078$, $p = .001$. This means the extent to which participants noticed and remembered the amount of upvotes assigned to a question predicted whether there was a difference in curiosity ratings for high and low upvote items (See Figure 3). This finding of popularity significantly affecting curiosity is consistent with the results of the original Dubey et al. (2020) study.

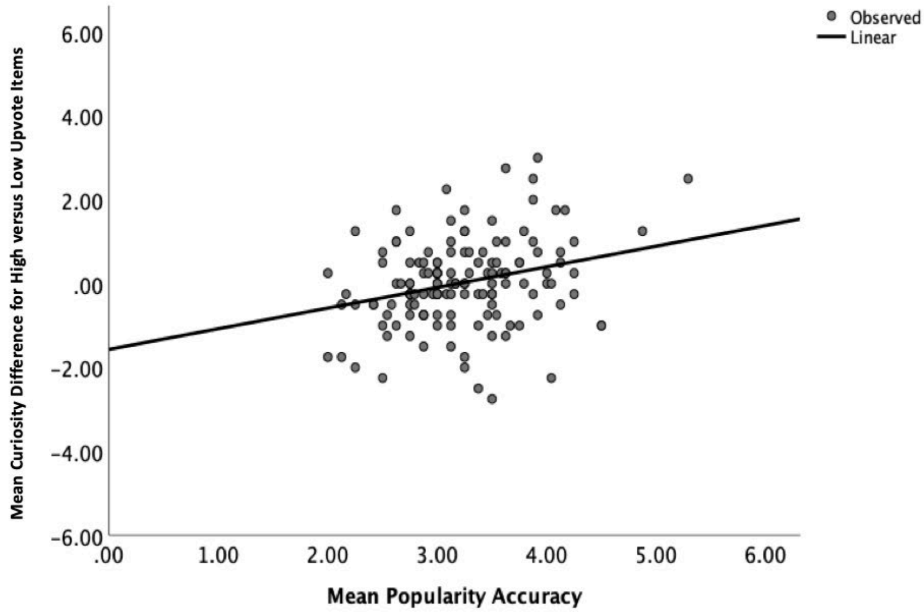

Figure 3: Scatterplot Depicting the Correlation Between Popularity Rating Accuracy and the Curiosity Difference for High versus Low Upvote Items

Discussion

The two primary goals of this study were to carry out a replication of Dubey, Mehta, and Lombrozo's (2020) study while correcting for a potential demand characteristic. A tertiary goal was to establish a foundation for related dwell-time research to investigate ways in which curiosity might influence downstream processing. My thesis presented data on just the first two parts of this three-part goal. Regarding the first two goals, I predicted that Dubey et al.'s (2020) results would still replicate even with the removal of the confound. I specifically hypothesized that the curiosity difference between low and high upvote questions would remain even when popularity was queried at the end, rather than immediately preceding the curiosity query as in Dubey et al.'s study.

In the interest of undertaking direct replication of the Dubey et al. findings, I first performed analyses that directly mirrored those they presented to test for a possible influence of popularity (upvotes) on participants' curiosity and choices regarding the questions they would like to see answers to. These analyses failed to reveal a statistically significant effect of popularity on participants' curiosity. Moreover, items with high upvotes were no more likely to be chosen for revealing the answer than items with low upvotes. Thus, on the face of it, my findings seemed to fail to replicate those of Dubey et al.

However, in my thesis research, information about popularity would have been considerably less salient to participants when they made judgements about curiosity, confidence, and the like, than it would have been for Dubey et al.'s participants. Therefore, it is perhaps not surprising that analyses directly following Dubey et al.'s analytic approach failed to show a significant effect of popularity on curiosity ratings. I realized that I could use participants' popularity ratings collected from near the end of the survey as an indicator of the extent to which they had registered, and accurately recalled, the upvote information about popularity. Interestingly, when I performed analyses that considered participants' memory accuracy for the popularity information provided, I saw a clear influence of upvotes on curiosity ratings. When participants took notice of the upvotes and accurately remembered which questions had been assigned a low or high amount, high popularity was associated with increased curiosity. This points to my research having replicated Dubey and colleagues' original finding in spirit. Taken together, my findings combined with those of Dubey et al. confirm that providing indicators of popularity can indeed enhance curiosity regarding the content of popular items.

I also undertook a series of exploratory analyses to better understand which factors were uniquely associated with curiosity, popularity, and the choice to reveal the answer to a given question. By and large, results from the three multiple regression analyses testing the contributions of various factors on perceived popularity, curiosity levels, and whether an answer was revealed or not looked similar across the upvote and baseline conditions. This fact points to robustness in my findings. The regressions revealed that ratings of social utility, usefulness, and perceived popularity all significantly, and uniquely, predicted curiosity in the baseline and upvote conditions. The results from these regressions replicated Dubey et al.'s findings of social utility being a significant predictor of curiosity, as well as confidence, curiosity, and popularity acting as significant predictors of answer selection. These findings point to perceived popularity being just one among several factors that influenced participants' curiosity about the Reddit questions they encountered in this study.

Unlike Dubey and colleagues, I also used multiple regression to probe which factors were uniquely associated with participants' popularity ratings. It turned out that curiosity was a significant predictor of popularity across both the upvote and baseline conditions, with upvote type also significantly predicting popularity in the upvote condition, and social utility and usefulness becoming additional significant predictors in the absence of upvotes in the baseline condition. These findings point to there having been multiple influences on participants' popularity ratings and suggest that popularity and curiosity may have bidirectional influence on one another.

The last set of regressions focused on which factors uniquely predicted whether an answer was chosen to be revealed. Confidence, curiosity, and popularity significantly predicted which answers were revealed in both conditions, with future usefulness only acting as a significant predictor in the baseline condition. By and large, these findings replicated those of Dubey and colleagues, and they again point to such choices tending to be jointly determined by multiple factors.

Limitations

Although some minor departures from precise counterbalancing occurred during data collection (i.e., a small number of data points were missing) the large data set arising from a substantial number of participants answering a sizable and diverse set of Reddit questions means that the power of this research to investigate these questions was strong. However, the generalizability of my findings is limited on the basis of the participants having been drawn from a convenience sample of university students.

Undergraduate students are often immersed in "Western, Educated, Industrialized, Rich, and Democratic (WEIRD)" societies (Henrich et al., 2010, p. 61). Past research has demonstrated that people from WEIRD regions are often outliers on many measures and are subsequently not representative of populations that fall outside the purview of these labels (Henrich et al., 2010).

Another methodological concern involves the fact that the popularity manipulation utilized could be viewed as rather artificial. The manipulation may have presented as a powerful enhancer of curiosity in this contrived experimental situation. It is not clear yet the extent to which popularity information in real-life contexts, such as observable peer consensus or celebrity endorsement may act as a stronger or weaker triggers of curiosity.

It is also worth noting that this study explored popularity as a curiosity enhancer only with respect to *Explain Like I'm Five* Reddit questions. The twenty-four questions used as stimuli were limited to a focus on physical, biological, and business-practice related topics. Curiosity about a variety of other possible domains, such as the emotional well-being of another, was not probed. Curiosity may be affected differently by popularity information when pertaining to these unexplored topics, especially if they are self-generated areas of interest.

Broader Implications

In this thesis, I investigated the extent to which a manipulation of popularity—via high versus low levels of upvotes—influenced participants' curiosity about a variety of questions. When participants registered and remembered the high upvote indicator of popularity, this information indeed caused enhanced curiosity about the question. This finding sets the stage for using popularity manipulations to trigger curiosity. However, my finding also makes it clear that if a researcher's main goal is to use popularity to enhance curiosity, constructing a strong popularity manipulation is paramount.

While causal conclusions can be drawn about popularity's effect on curiosity, it must be recognized that my findings also demonstrate a possible bidirectional effect between curiosity and popularity. By performing the multiple regression analyses to examine the predictors of perceived popularity and curiosity, I found that both variables acted as significant predictors of one another while controlling for other influential factors such as social and future utility. This means that the more popular a question seemed to someone, the higher they rated their curiosity about its answer and the more curious they were about a question's answer, the more popular they found the question. This finding goes beyond the regression model suggested in Dubey et al.'s (2020) study. Their model only attempts to explain how curiosity could arise from popularity, but it does not attempt to explain the reverse directionality that my findings seem to suggest exists. If this relationship holds up as a causal directionality in future experimental manipulations, a richer model than what Dubey et al. (2020) presented would be necessitated to explain interactions between curiosity and popularity.

This study also aides in the understanding of what curiosity is and how it functions. My findings are consistent with Dubey and colleagues' when it comes to demonstrating that curiosity acts as a significant predictor of whether someone seeks out the answer to a question. This finding lends further support to the classification of curiosity as a drive state for information. Additionally, curiosity is a vital component of learning. Its stimulation is thought to promote the desire for knowledge, whether it be to fill in the gaps created by novel information or expand upon areas of personal interest (Loewenstein, 1994; Schmitt & Lahroodi, 2008). Curiosity's importance to this fundamental force for learning behooves us to gain further understanding about the relationship between popularity, or other triggers, and curiosity. This knowledge could eventually allow for educators to more effectively harness curiosity to motivate learning.

Future Research Directions

My findings combined with those from Dubey et al. (2020) firmly suggest that it is possible to influence people's curiosity through a popularity manipulation. This fact will be helpful in the future for exploring the original third goal of my thesis: examining the extent to which curiosity reorganizes people's processing of information related to the things they are curious about. This goal can be approached by analyzing the dwell time output originating from the four question-related slideshows shown to participants. Regarding this avenue of research, I hypothesize that dwell-time patterns will differ between participants exposed to the low, high, and baseline conditions of upvotes. Specifically, I predict that a high number of upvotes will lead to higher dwell-times overall, with an especially large increase in dwell-time for question relevant portions of the slideshows. Additionally, I predict that a low number of upvotes will reduce dwell-times overall, with little time difference between question relevant and irrelevant portions of the slideshows.

Techniques beyond the Dwell-Time Paradigm could also be used to investigate the possibility that popularity manipulates curiosity beyond in the moment questions about a puzzling topic to the point that it affects how a person expends their attention when taking in visual input. One alternative approach is pupillometry, a technique that records pupil dilation over time. As previously mentioned, pupil dilation has been found to be positively correlated with curiosity, with high levels of curiosity leading to the largest dilation responses (Kang et al., 2009). This technique is also well-suited to this venture due to its status as a valuable method of for probing event processing (e.g., Tanaka, 2018).

In the future, it will also be of interest to explore the extent to which different triggers of curiosity induce comparable downstream consequences for processing. Other known triggers to curiosity—beyond popularity—include the introduction of a) an information gap and b) a violation of expectations experience (Loewenstein, 1994). It seems plausible that varying curiosity triggers might induce differing alterations to processing. Put another way, perhaps curiosity is a multi-faceted phenomenon which comes in different forms that generate unique consequences for subsequent learning.

Conclusion

When considered together with Dubey, Mehta, and Lombrozo's previous research, my findings demonstrate that information concerning the popularity of an item can trigger enhanced curiosity about questions regarding puzzling scenarios. These findings provide information about the drive state mechanisms of curiosity, which is an important motivation for learning and development. My partial replication also sets the stage for future research to examine in more detail how curiosity reshapes learning and attentional processes.

Appendix A

1	Why aren't other animals as freaked out by bugs and creepy crawlies as humans?
2	How do scientists know what the global temperature was millions of years ago?
3	Why is CPR for drowning different than CPR for people who collapse from heart problems?
4	What make some objects "bouncier" than others?
5	What is the difference between beat, bar, steps, tempo, tact, and rhythm?
6	Why do we toss and turn/constantly reposition ourselves during sleep?
7	All energy transfer in nature from one point to another happens in waves. Which fundamental property of nature is responsible for wave-like nature?
8	What is the difference between forward and reverse osmosis?
9	How do breeders ensure diversity among their animals' offspring?
10	If rockets use controlled explosions to propel forward, why can't we use a nuclear reaction to launch and fly our rockets?
11	How can alcohol withdrawal or detox kill you?
12	Why do tongues get weird bumps when burnt or after eating something really sweet or really salty?
13	Do multivitamins and Omega-3 pills actually do anything?
14	Why does giving someone a transfusion of my blood not give them my immunity?
15	Why can't the asteroid belt accumulate into one rocky planet?
16	How do earphones produce adequate bass despite their size?
17	An anechoic chamber at Orfield Laboratories in Minnesota has negative decibel levels (lower than -9db). How is this possible?
18	How do car dealerships make money when they claim the markup on new cars is only a few hundred dollars?
19	Why do US based airlines lag behind in service and quality, especially in their premium cabins?
20	What happens that makes beer taste terrible after it warms up and then is re-chilled?

Table 1: Bank of 20 Questions

Correct items	1. The doctor checked the patient's reflexes 2. A baby penguin appeared 3. The octopus crawled through a tube 4. A panda chewed on bamboo
Distractor items	1. The octopus changed color 2. A panda ate a pumpkin 3. The doctor weighed the patient 4. A penguin walked into the water

Table 2: Attention Check Prompts

Appendix B

	upvtyp	N	Mean	Std. Deviation	Std. Error Mean
bothconfdnc	.00	135	2.0657	1.06009	.09124
	1.00	72	2.9861	1.70742	.20122
bothcrsty	.00	135	3.1102	1.22113	.10510
	1.00	72	3.5139	1.86134	.21936
bothscl	.00	135	1.7176	.93765	.08070
	1.00	72	2.0139	1.64011	.19329
bothftr	.00	135	1.9093	.97334	.08377
	1.00	72	2.1528	1.78144	.20994
bothanswr	.00	135	.5074	.12194	.01050
	1.00	72	.4028	.49390	.05821
bothpoplrty	.00	135	3.1759	.84879	.07305
	1.00	72	3.1250	1.39352	.16423

Table 3: Descriptive Statistics for Baseline and Low Upvote Judgments

	upvtyp	N	Mean	Std. Deviation	Std. Error Mean
bothconfdnc	.00	135	2.0657	1.06009	.09124
	2.00	71	3.0563	2.19016	.25992
bothcrsty	.00	135	3.1102	1.22113	.10510
	2.00	71	3.6056	1.79267	.21275
bothscl	.00	135	1.7176	.93765	.08070
	2.00	71	1.9859	1.52590	.18109
bothftr	.00	135	1.9093	.97334	.08377
	2.00	71	2.2817	1.81410	.21529
bothanswr	.00	135	.5074	.12194	.01050
	2.00	71	.4507	.50111	.05947
bothpoplrty	.00	135	3.1759	.84879	.07305
	2.00	51	3.6863	1.27264	.17820

Table 4: Descriptive Statistics for Baseline and High Upvote Judgments

Appendix C

	Mean	Std. Deviation	N
Answer	.4874	.50007	1073
Confidence	2.3075	1.90064	1073
Curiosity	3.6039	1.80657	1073
Social	2.1594	1.64071	1073
Future	2.3821	1.80674	1073
UpVoteType	1.5005	.50023	1073
Popularity	3.2544	1.59217	1073

Table 5: Descriptive Statistics for Multiple Regressions

Correlations

		Confidence	Curiosity	Social	Future	Answer	Popularity
Confidence	Pearson Correlation	1	.119**	.237**	.280**	-.013	.071*
	Sig. (2-tailed)		.000	.000	.000	.673	.020
	N	1144	1144	1144	1144	1144	1073
Curiosity	Pearson Correlation	.119**	1	.466**	.407**	.364**	.285**
	Sig. (2-tailed)	.000		.000	.000	.000	.000
	N	1144	1144	1144	1144	1144	1073
Social	Pearson Correlation	.237**	.466**	1	.731**	.216**	.179**
	Sig. (2-tailed)	.000	.000		.000	.000	.000
	N	1144	1144	1144	1144	1144	1073
Future	Pearson Correlation	.280**	.407**	.731**	1	.209**	.157**
	Sig. (2-tailed)	.000	.000	.000		.000	.000
	N	1144	1144	1144	1144	1144	1073
Answer	Pearson Correlation	-.013	.364**	.216**	.209**	1	.265**
	Sig. (2-tailed)	.673	.000	.000	.000		.000
	N	1144	1144	1144	1144	1144	1073
Popularity	Pearson Correlation	.071*	.285**	.179**	.157**	.265**	1
	Sig. (2-tailed)	.020	.000	.000	.000	.000	
	N	1073	1073	1073	1073	1073	1073

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

*. Correlation is significant at the 0.05 level (2-tailed).

Table 6: Correlational Table for the Upvote Condition

Correlations

		Confidence	Curiosity	Social	Future	Answer	Popularity
Confidence	Pearson Correlation	1	.120**	.246**	.304**	-.008	.066*
	Sig. (2-tailed)		.000	.000	.000	.797	.031
	N	1076	1076	1076	1076	1076	1076
Curiosity	Pearson Correlation	.120**	1	.400**	.346**	.328**	.371**
	Sig. (2-tailed)	.000		.000	.000	.000	.000
	N	1076	1076	1076	1076	1076	1076
Social	Pearson Correlation	.246**	.400**	1	.640**	.206**	.273**
	Sig. (2-tailed)	.000	.000		.000	.000	.000
	N	1076	1076	1076	1076	1076	1076
Future	Pearson Correlation	.304**	.346**	.640**	1	.185**	.153**
	Sig. (2-tailed)	.000	.000	.000		.000	.000
	N	1076	1076	1076	1076	1076	1076
Answer	Pearson Correlation	-.008	.328**	.206**	.185**	1	.299**
	Sig. (2-tailed)	.797	.000	.000	.000		.000
	N	1076	1076	1076	1076	1076	1076
Popularity	Pearson Correlation	.066*	.371**	.273**	.153**	.299**	1
	Sig. (2-tailed)	.031	.000	.000	.000	.000	
	N	1076	1076	1076	1076	1076	1076

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

*. Correlation is significant at the 0.05 level (2-tailed).

Table 7: Correlational Table for the Baseline Condition

Bibliography

Dubey, R., & Griffiths, T. L. (2020). Reconciling novelty and complexity through a rational analysis of curiosity. Psychological Review, 127(3), 455.

Dubey, R., Mehta, H., & Lombrozo, T. (2020). Curiosity is contagious: A social influence intervention to induce curiosity. https://doi.org/10.31234/osf.io/zfjkg

Hard, B. M., Recchia, G., & Tversky, B. (2011). The shape of action. *Journal of experimental psychology: General*, *140*(4), 586. https://psycnet.apa.org/doi/10.1037/a0024310

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world?. *Behavioral and brain sciences*, *33*(2-3), 61-83. https://doi.org/10.1017/S0140525X0999152X

Kang, M. J., Hsu, M., Krajbich, I. M., Loewenstein, G., McClure, S. M., Wang, J. T. Y., & Camerer, C. F. (2009). The wick in the candle of learning: Epistemic curiosity activates reward circuitry and enhances memory. *Psychological science*, *20*(8), 963-973. https://doi.org/10.1111%2Fj.1467-9280.2009.02402.x

Kashdan, T. B., & Silvia, P. J. (2009). Curiosity and interest: The benefits of thriving on novelty and challenge. Oxford handbook of positive psychology, 2, 367–374.

Kosie, J. E., & Baldwin, D. (2019). Attention rapidly reorganizes to naturally occurring structure in a novel activity sequence. *Cognition*, *182*, 31-44. https://doi.org/10.1016/j.cognition.2018.09.004

Kidd, C., & Hayden, B. Y. (2015). The psychology and neuroscience of curiosity. *Neuron*, *88*(3), 449-460. https://doi.org/10.1016/j.neuron.2015.09.010

Liquin, E. G., & Lombrozo, T. (2020). A functional approach to explanation-seeking curiosity. *Cognitive psychology*, *119*, 101276. https://doi.org/10.1016/j.cogpsych.2020.101276

Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. Psychological Bulletin, 116, 287–317.

Sage, K.D. & Baldwin, D. (2014). Looking to the hands: Where we dwell in complex manual sequences. *Visual cognition*, *22*(8), 1092-1104. http://dx.doi.org/10.1080/13506285.20 14.962123

Schmitt, F. F., & Lahroodi, R. (2008). The epistemic value of curiosity. *Educational theory*, *58*(2), 125-148. https://doi.org/10.1111/j.1741-5446.2008.00281.x

Tanaka, Y. (2018). *An implicit measure of event segmentation using pupillometry* [Unpublished master's thesis]. University of Oregon.

Ward, Mark F. "The Wender Utah Rating Scale: an aid in the retrospective diagnosis of childhood attention deficit hyperactivity disorder." *American journal of Psychiatry* 150 (1993): 885-885. https://doi.org/10.1176/ajp.150.6.885