Modern Psychological Studies

Volume 25 | Number 1 Article 8

January 2020

Evaluation of Antecedent Ambiguity on Identification and Categorization of Behavior in Lunar-Effects Research

Brenda M. Gabbert
University of Alaska Anchorage, bmgabbert@alaska.edu

Mychal A. Machado
University of Alaska Anchorage, mmachado2@alaska.edu

Abstract

We created an online survey to determine the degree to which 114 participants could correctly identify and categorize scripted visual examples of typical and atypical behavior when the antecedent preceding those behaviors was not present. We also asked participants to determine whether our visual examples of behavior occurred in the presence of a full moon, and subsequently evaluated these data in the context of participant's self-reported belief in lunar-effects. Our results show the absence of an antecedent event influenced some participants' identification and categorization accuracy scores, and that participants with a prior belief in lunar effects were more likely to attribute atypical behavior to the presence of a full moon. Future areas for improving measurement in lunar-effects research are discussed.

Keywords: observation, antecedents, ambiguity, atypical behavior, lunar effects, moon

Evaluation of Antecedent Ambiguity on Identification and Categorization of Behavior in Lunar-Effects Research

Humans have a strong desire to understand and think about their world, but we are not always rational in this approach. Science and critical thinking do not come naturally; they take training, experience, and effort; thinking is skilled work (Mander, 1947). However, we live in a diverse world with complex problems, and many of us want certainty, want to control our environment, and want a simple explanation. In the absence of these explanations or when multiple explanations for the same phenomenon exist, myths, superstition, and lore may trump empiricism (Shermer, 2002). For example, a common fallacy is the suggested relationship between phases of the moon and various forms of atypical behavior, including assaults, homicides, road rage, and theft. The origins of this lore and the specifics related to how the moon influences behavior are not clear, but nonetheless it is not uncommon to hear many individuals publicly, including police officers (e.g., Lieber, 1978; Stolzenberg, D'Alessio, & Flexon, 2017) and mental health professionals (e.g., Schafer, Varano, Jarvis, & Cancino, 2010; Parmar et al., 2014) attribute behaviors considered to be aberrant, criminal, or otherwise atypical to the moon.

One explanation for the endurance of lore related to the moon-behavior relationship could be familiarity with research in support of the myth, because (no matter how controversial) these data exist. For example, Purpura (1979) reviewed the number of calls for police service made to three law enforcement agencies over a 59-month period, and found that certain crimes (i.e., breaking and entering) were more likely to occur in the presence of the full moon. More recently, Stolzenberg et al. (2017) investigated the effect of lunar illumination on reported crimes and found as the moon became brighter (i.e., fuller) the number of reported outdoor criminal activities increased simultaneously. Over the past four decades, other researchers have reported a

positive relationship between lunar phases and sleep deprivation (Chaput et al., 2016), aggression (Lieber, 1978; Kazemi-Bajestani, Amisadri, Samari, & Javanbakht; 2011), violent incidents in correctional settings (Pettigrew, 1985), demand for emergency room services (Blackmon & Catalina, 1973; Kazemi-Bajestani et al., 2011), suicide attempts (Taylor & Diespecker, 1972), hospital admissions (Templer & Veleber, 1980; Weiskott & Tipton, 1975), and calls to counseling hotlines (Templer & Veleber, 1980; Weiskott, 1974). Thus, results in support of a moon-behavior relationship exist, span decades, and are often readily available to readers and interested parties.

Despite the existence of data supporting a phenomenon, science requires the assessment of evidence from all sides of an argument, not just those data in favor of a claim (Shermer, 2002). With respect to the moon, most lunar-effect studies have failed to replicate positive findings in at least two ways. First, when evaluators have reexamined previously reported data using control comparisons (Drum, Terry, & Hammonds 1986; Schafer et al., 2010), an increased observation window (Kelly & Martens, 1994; Rotton, Kelly, & Frey, 1983), stricter criteria to determine statistical significance (Roy, Biswas, & Roy, 2017), or a combination of these methods (Bickis, Kelly, & Byrnes, 1995; Rotton & Kelly, 1985; Templer, Veleber, & Brooner, 1982), the obtained findings were contradictory to those previously reported (i.e., positive results turned null). Second, when researchers applied these methodological improvements in their own independent investigations, similar failures to replicate were obtained for lunar effects and physical activity (e.g., Parmar et al., 2014), sleep (e.g., Chaput et al., 2016), health (e.g., Roy et al., 2017), aggression (e.g., Simón, 1998), use of hospital and psychiatric services (e.g., Chapman & Morrell, 2000; Walters, Markley, & Tiffany, 1975), incidents of suicide and suicide attempts (e.g., Biermann et al., 2005; Lester, Brockopp, & Priebe, 1969), use of crisis hotlines

(Wilson & Tobacyk, 1990), demands for emergency services (e.g., Schafer et al., 2010), road rage (Campbell & Beets, 1978; Laverty & Kelly, 1998), and homicide (Pokorny & Jachimczyk, 1974). Thus, many researchers have failed to find a lunar effect, and these researchers have stressed the importance of including comparison samples, appropriate windows for observation, and a careful use of statistics in future investigations.

Another potential consideration for lunar-effect studies is the use of calls for service as a dependent variable. Calls for service (i.e., calls placed to police and emergency departments) are commonly and currently used as an indicator of the frequency with which atypical behavior occurs in an area, and these data are correlated with phases of the lunar cycle to determine the presence (absence) of a moon-behavior relationship (see Kazemi-Bajestani et al., 2011, Parmar et al., 2014, and Stolzenberg et al. 2017 for recent examples). It seems reasonable that atypical behaviors would be more likely to evoke a call for service than typical behavior, however, Klinger and Bridges (1997) examined police dispatching records in 60 neighborhoods and found that calls for service records often underestimated levels of crime, overestimated types of crime (e.g., assaults), and varied in number systematically across spaces. There was also substantial error in the accuracy with which calls for service were made; when calls for service were compared to street officer records (i.e., reports and radio calls), 20% of calls for service initially classified as crime (n = 185) involved no crime at all, and 19% of calls initially classified as not involving a crime (n = 165) did involve a situation requiring police involvement. Together, these data suggest that caller inaccuracy can bias aggregate calls for service crime counts and subsequently impact analyses based on these data, which suggests that identifying an alternative variable to calls for service would be ideal. However, it is unlikely researchers would be motivated to select an alternative variable because of the ease with which calls for service data

can be obtained (Klinger & Bridges, 1997). Thus, identifying the conditions that impact caller accuracy would be beneficial to future lunar-effect studies.

A reasonable first step in evaluating caller accuracy suggested by the clear connection in prior research between calls for service data and atypical behaviors (e.g., crime, aggression, etc.) could be to determine the degree to which an individual can accurately distinguish (categorize) behavior that is atypical, which should warrant a call for service, from behavior that is typical, which should not warrant a call for service. Although a seemingly easy distinction, accuracy might vary under conditions where the cause of a behavior is unknown to the observer. That is, if the antecedent event (i.e., an event preceding behavior) is observed, an individual might be able to categorize the behavior as typical or atypical accurately. However, if the same behavior is observed in the absence of that antecedent event, an individual's response might be inaccurate. For example, if Brenda observes Katya pick-up a backpack from under an empty chair in a public space, she might be more likely to categorize the behavior as typical (i.e., a person came back to retrieve a thing they forgot) and not engage in a call for service. In contrast, if Brenda had observed Trixie get-up from the chair to throw something away, and then sees Katya pick-up the backpack from under the empty chair, she might be more likely to categorize the behavior as atypical (i.e., someone is trying to steal something) and subsequently respond differently. Thus, the absence of an antecedent event might differentially impact an observer's accuracy with respect to identifying and categorizing behavior, which might then change the probability of accurately placing a call for service.

We are unaware of any evaluations that have sought to determine whether the absence of a discernable antecedent event impacts the accuracy with which individuals identify and categorize typical and atypical behaviors. Such information would be beneficial, because

behaviors leading to a call for service are most likely observed under conditions where the cause of behavior is unknown (Reiss, 1971). Therefore, it would be beneficial to know whether witnessing behavior "happen out of the blue" would impact an individual's accuracy, because errors could influence the probability of false-positive and false-negative reports to authorities. Thus, the purpose of the current study was to assess the degree to which participants could correctly identify and categorize scripted visual examples of typical and atypical behavior when the antecedent preceding those behaviors was not present. In addition, because lunar research continues to focus on atypical behavior, we also evaluated whether atypical behaviors were more likely to be attributed to the presence of the full moon than typical behaviors, and how, if at all, those data related to an individual's prior belief in lunar-effects. We hypothesized that participants would have difficulty accurately categorizing behavioral examples as typical or atypical when the antecedents proceeding those behavioral examples were unclear. We also hypothesized that participants reporting a prior belief in lunar effects would be more likely to attribute atypical behaviors to the presence of the full moon than typical behaviors.

Method

Participants and Setting

Participants were recruited for an online survey from a small, public university in Alaska through an online survey system available to psychology students and from online posts announcing the study located on several social media sites (i.e., Facebook, Twitter, and Google+). After removing data from participants who did not complete the study (N = 57), 114 participants remained. All participants were 18 years or older, and most participants (41.2%) reported ages between 18 and 29 years (n = 47), followed by 30 to 45 years (22.8%, n = 26), 46 to 60 years (21.1%, n = 24), and 60 years or more (14.9%, n = 17). Men made up 30.7% of the

sample (n = 35) but the majority (69.3%) were women (n = 79); all participants reported as cisgender. Finally, most participants (93.9%, n = 107) reported being raised primarily in the United States. We did not inquire about the specific age of participants nor the ethnicities to which participants identified because our study was focused on obtaining responses from the general adult population, regardless of age and ethnic identification. All participants completed the online survey, on average, in 7.14 min (range, 5.08 min to 9.70 min) at a time that was convenient for them and in a self-identified location that had an internet connection.

Materials and Procedures

Programmed typical and atypical behaviors. Participants were asked to view several video vignettes featuring examples of behaviors generally considered typical or atypical and subsequently decide whether the behaviors viewed occurred in the presence of a full moon. Our selection of behaviors for each category was determined quasi-systematically. First, as a group, the research team generally defined typical and atypical behavior. Second, an initial list of behaviors that could potentially fit each category definition was generated by the first author using examples from previous lunar-effects studies. Third, 10 members of an undergraduate social science seminar course and the second author reviewed the list and came to a consensus (by vote) on the four examples they agreed would best represent each category of behavior. Through these procedures, we selected the following definitions and behaviors for the categories: *typical behavior* was defined as any behavior one might expect to be evoked given the limited context presented, and included (1a) walking to a rubbish bin, (2a) playing a game on a smartphone, (3a) asking to borrow money, and (4a) having a conversation; *atypical behavior* was defined as behavior one would not expect to be evoked given the same or similar context, and

included (1b) sprinting back-and-forth to a rubbish bin, (2b) kicking over a piece of furniture, (3b) stealing a wallet, and (4b) engaging in a physical altercation (i.e., a fight).

Vignettes. After selecting our typical and atypical behavior examples, we created eight video vignettes (four vignettes each of typical and atypical behavior) using the camera on an iPhone 7 and 11 student role players. The recording area of the camera was large enough to cover a standard sized classroom (i.e., 4.5 m by 7.6 m). To decrease the likelihood that participants would identify our examples as being contrived, we arranged each vignette to look as if it were taken from a security camera by positioning the camera up high and angled down toward the target area. After filming was complete, we removed the audio and color from each vignette using iMovie (10.1.4). We opted not to include a timestamp on the vignettes, because we wanted to prevent participants from being able to research the date and time displayed to check the status of the moon that day. Each vignette was brief ($\leq$ 30 s), contained a single programmed example of typical or atypical behavior, and featured one to seven role players (see Table 1 for details). We scripted each example to increase the accuracy with which the role players engaged in programmed actions. It is also important to note that none of the vignettes were filmed during a full moon.

Online Survey. After creating our vignettes of typical and atypical behaviors, we created our online survey using Qualtrics (https://www.qualtrics.com/). Participants were directed to the survey through an online survey system available to psychology students and through online posts on several social media sites. Following an informed consent form, participants were asked to respond to several demographic questions related to age range, gender, and country of origin. Next, participants were shown eight video vignettes, consisting of four typical and four atypical behavior examples, in a random order. After each vignette, we asked participants to identify the

behavior displayed in the vignette (Identification), label whether they believed the behavior displayed was typical or atypical behavior (Categorization), and then indicate whether they believed the behavior displayed occurred in the presence of a full moon (Attribution). Finally, after all vignettes were viewed, participants were asked to indicate their general belief in the moon's effects on behavior and the natural world (General Belief).

Following completion of the primary survey, participants were directed to a short debriefing form, which explained that all vignettes had been scripted and created by the research team. Participants who were university students were then directed to an additional survey where they could enter their name and the title of a psychology course for which they wanted to earn extra credit as compensation. Participants who were not university students were subsequently directed to an additional survey where they could enter their name into a raffle for a $25 gift card to the coffee shop of their choice. All names, course information, and contact information obtained from these supplemental surveys were kept separate from the primary survey data to maintain participant confidentiality.

Dependent Measures

Identification. Identification was defined as a participant correctly selecting the target behavior programmed in each vignette from a list of multiple-choice options. To assess identification after each vignette, participants were presented with the question "Which behavior was displayed in the video you just watched?" along with four response options. Response options consisted of the target behavior, two non-target behaviors, and a null response (i.e., "none of the above"); the programmed target behavior and a null response were always included in the response options. Participants received a score of "1" or "0" depending on whether they selected the correct or incorrect response option, respectively. We included identification

questions following each vignette because this allowed us to verify whether participants attended to the programmed target behavior included in each vignette. Cronbach's alpha for the 8 identification questions was 0.66.

Categorization. Categorization was defined as the label (typical or atypical) a participant gave to each programmed behavior example. To assess categorization after each vignette, participants were presented with the question "Do you consider the behavior displayed in the video you just watched typical or atypical?". Participants received a score of "1" if their categorization matched the category to which we assigned the behavior. Participants received a score of "0" if their categorization did not match the category to which we assigned the behavior. Cronbach's alpha for the 8 categorization questions was 0.68.

Attribution. Attribution was defined as a participant indicating that a behavior example occurred during a full moon. To assess attribution after each vignette, participants were presented with the statement "The behavior I just observed occurred during a full moon." Participants received a score of "1" if they responded that the behavior did occur in the presence of a full moon, and a score of "0" if they responded that the behavior did not occur in the presence of a full moon. We included these attribution questions so that we could subsequently compare attributions of typical and atypical behaviors to prior beliefs regarding lunar effects (i.e., General Belief). Cronbach's alpha for the 8 attribution questions was 0.62.

General belief. General belief was defined as a participant's self-reported belief in the effects of the full moon. To assess general belief, participants were presented with the question "What is your belief regarding the full moon?" along with three response options after all vignettes had been viewed. Response options included (a) it affects behavior, (b) it affects the tides/other natural phenomena, and (c) it affects none of the above. Participants received a score

of “2” if they indicated that the moon affects behavior. Participants received a score of “1” if they indicated that the moon affects tides and other natural phenomena. Participants received a score of “0” if they indicated the moon affects none of the presented options. Because it is not advised to conduct Cronbach’s alpha to determine reliability with a single item (Aron, Coups, & Aron, 2013), we confirmed the appropriateness of our response options through independent, verbal affirmation from faculty members with backgrounds in psychology (n = 3) and biology (n = 2) prior to conducting our study. These affirmations provided some face validity for the included options.

Results

Identification

Table 2 displays the overall identification scores across typical and atypical video vignettes. As a group, participants (N = 114) correctly identified the programmed target behavior, on average, in 96.5% of the typical behavior vignettes (range, 95.6% to 97.4%) and in 94.3% of the atypical behavior vignettes (range, 90.4% to 97.4%). Although correct identifications were slightly higher for the typical behavior vignettes than the atypical behavior vignettes, the difference in correct identification scores across vignette type was not statistically significant (t [6] = 1.67, p = 0.09). Together, our identification data showed that many of our participants (79%, n = 90) were able to correctly identify our programmed behaviors, regardless of category. However, approximately 21% of our participants (n = 24) incorrectly identified one or more of our programmed behaviors.

Data Trimming

We were interested in assessing whether participants could correctly categorize scripted visual examples of typical and atypical behavior when the antecedent preceding those behaviors

was unclear. To this end, we considered correctly identifying the programmed target behavior in each vignette critical, because, if a participant did not attend to the programmed target behavior, then their subsequent categorization of the behavior would be potentially under the control of something other than our programmed target. In situations where attending to a presented stimulus prior to responding is important, omitting data from trials or participants that did not meet attending criteria is recommended (Aron et al., 2013). Thus, after evaluating identification scores across all the vignettes for each participant, we removed data from participants who had at least one incorrect identification (*n = 24*), which left 90 participants for our remaining analyses.

Categorization

Table 3 displays the overall categorization scores across typical and atypical behavior vignettes. With respect to the individual programmed target behaviors, the programmed behaviors in most vignettes (75%, *n* = 6) were correctly categorized with greater than or equal to 90% accuracy. The only exceptions were Vignette 3a (typical) and Vignette 2b (atypical), which had a correct categorization score of 66.7% and 88.9%, respectively. As a group, participants (*N* = 90) correctly categorized our programmed typical behavior examples, on average, with 90.6% accuracy (range, 66.7% to 100%), and our atypical behavior examples with 95.6% accuracy (range, 88.9% to 100%). The overall difference in categorization scores across typical and atypical behavior vignettes was not statistically significant (t [6] = 0.58, p = 0.30).

Attribution and General Belief

The results for attribution and general belief as they relate to our atypical and typical behavior vignettes are represented in Table 4 and Table 5, respectively. We examined the relationship between belief in lunar effects and attribution of behavior to the presence of the full moon using a 2 x 3 chi-square test of independence; a Bonferroni correction of the *P* values was

used because of our limited sample size. Across participants (N = 90), these analyses showed that the relation between these variables was significant for atypical behaviors, ($\chi 2$ [2, N = 90] = 10.59, p = 0.01), but not for typical behaviors ($\chi 2$ [2, N = 90] = 0.68, p = 0.71). More broadly, participants reported a variety of beliefs in the moon's effects, including it affects behavior (32%, n = 29), it affects the natural world (41%, n = 37), and it affects nothing (27%, n = 24).

Discussion

Calls for service has face validity as an indicator of atypical behavior, and these data are easily obtained from police dispatch, fire department, and emergency service (e.g., crisis hotline) records. However, variability in caller accuracy has been shown to introduce error when calls for service are used as a primary dependent variable, and little research has been conducted on what impacts a caller's ability to accurately report behavior observed. Although most of our participants could identify and categorize our programmed behaviors, we found that 21% of participants in our original sample (N = 114) did not correctly identify the behavior of interest in our programmed vignettes, and, of the participants who correctly identified all behaviors of interest (N = 90), 38% of those participants incorrectly categorized at least one of our examples. We also found that participants who reported a preexisting belief in the moon's effect on behavior were more likely to attribute atypical behavior to the presence of the full moon, and, more broadly, that 59% of our participants reported a belief related to the moon that was not aligned with current empirical research. Together, our preliminary results suggest that attention to a presented behavioral episode, the ambiguity of that behavioral episode, and prior beliefs about the actions in the behavioral episode could be important variables in future analyses of the errors produced by calls for service data. Our results also suggest that efforts to mitigate

perpetuation of the moon-behavior relationship might potentially be achieved through education related to what the moon can (not) affect.

To our knowledge, only two attempts to investigate how accuracy impacts calls for service data have been reported. Reiss (1971) noted that a comparison of police dispatch records and 911 call logs showed that inaccuracy in callers' descriptions of the events they reported led to discrepancies in the overall frequency and type of crime data obtained. The most common error was reporting false negatives (i.e., Type II errors); callers reported a significant number of noncriminal events (i.e., typical behavior) as if they were criminal events (i.e., atypical behavior). Similarly, Klinger and Bridges (1997) examined police dispatching records in 60 neighborhoods and found, in addition to a high number of false negatives, that calls for service records often overestimated types of crime (i.e., produced false positives or Type I errors). We extended the results of Reiss (1971) and Klinger and Bridges (1997) in several ways. First, we independently evaluated accuracy by creating scripted vignettes to serve as the criterion record against which comparisons were made. Second, we evaluated two specific variables that could impact caller accuracy: identification and categorization. Third, we ensured that our analysis of categorization included only those participants who correctly attended to the behaviors of interest; by contrast, Reiss (1971) and Klinger and Bridges (1997) assessed but did not control for correspondence between the events observed and subsequently reported on. Fourth, we evaluated individuals who would likely place a call for service rather than examine 911 operator and police officer call logs. Fifth, we included an analysis of typical as well as atypical behavior. Finally, we evaluated how participant attributions of behavior related to prior belief in lunar effects.

Klinger and Bridges (1997) noted that in addition to identifying ways to correct for the level of error observed in studies using calls for service data, evaluations into the conditions under which callers are likely to be more (less) accurate when reporting atypical behavior would be beneficial. These data could be used to refine future analyses using calls for service, fuel the search for and validation of a different variable for use in these studies, or both. We evaluated whether participants would identify the behavior of interest when a clear antecedent event was not presented, and, if so, whether they would be able to correctly categorize the behavior according to our programming, but we did not ask participants whether they would place a call for service for two reasons. First, the important question that only a few researchers have attempted to answer (Klinger & Bridges, 1997; Reiss, 1971) is whether categorization accuracy varies with respect to typical and atypical behavior. Second, as one consideration with respect to internal validity, we told our participants that our video examples were taken from a university's security system, which might have deceased the probability that participants would place a call for service, because law enforcement is typically already involved in the monitoring of these systems. Nonetheless, it is important to know how identification and categorization of an observed event relates to engagement in calls for service. Future researchers might expand on our evaluation by presenting video vignettes of scripted behavior examples to participants as we modeled, and then evaluating the degree to which observer identification and categorization accuracy impacts the number of calls for service that would be made.

Despite procedures to conceptualize and create clear, salient examples of typical and atypical behavior, some participants incorrectly identified our programmed behaviors, and some participants incorrectly categorized our programmed behaviors despite identifying the programmed behaviors correctly. The most frequently misidentified programmed behaviors in

our evaluation were atypical behaviors, which were subsequently categorized as typical behaviors (data available from second author), and included a student sprinting back-and-forth in a classroom to deposit trash in a rubbish bin (Vignette 1b) and a student kicking over an ottoman in a lounge area (Vignette 2b). We also found that most participants correctly identified our scripted example of someone asking someone else for money to use a vending machine (Vignette 3a), but subsequently categorized the behavior as atypical. These examples of false negative and false positive results replicate previous findings (e.g., Klinger & Bridges, 1997). However, in contrast to previous findings, we found relatively few errors overall in our evaluation. One explanation could be that our vignettes were brief (i.e., $\leq$ 30 s), and this might have increased the probability of correct responding because the information provided in each vignette was minimal. Future research should therefore include evaluations of identification and categorization accuracy over extended (e.g., 1 hr) observation samples.

It might also be helpful for future researchers to consider the rationale for why participants categorize observed behavior as typical or atypical, and how that rationale impacts their categorization accuracy. We did not ask our participants to report why they categorized behavior as they did, but it is likely they responded based on more than just the presented topographies (i.e., how the behavior looked). Research on attribution theory has suggested that determinations of causation can be affected by an observer's perception of whether the behavior occurring is controllable (volitional) or uncontrollable (reflexive) given a presented situation (Fishman & Husman, 2017; Weiner, 2010), and whether the behavior is assumed to be a frequent (stable) or infrequent (unstable) part of an individual's repertoire (Weiner, 2010). Thus, two individuals could observe the same topography (e.g., kicking over an ottoman), but categorize the behavior differently based on their perceptions of why the behavior is occurring. Future

evaluations of categorization should therefore include open-ended questions to capture the reasons for categorizations provided. These data are likely to allow a more nuanced analysis of accuracy.

The design of our vignettes was partially informed by research involving assessments of lunar effects on atypical behavior. However, as noted by Lerman et al. (2010), in naturalistic contexts there is likely to be significant "noise" or ambiguous stimuli that can control observer responding. We did not program ambiguous stimuli because, as a preliminary study, we were primarily interested in whether participants could correctly identify and categorize brief, clear examples of typical and atypical behavior in the absence of an antecedent event. Nevertheless, future researchers should evaluate the generality of our results to live instances of typical and atypical behavior obtained from a variety of sources (e.g., police cameras, traffic cameras, etc.). The influence of naturally occurring stimuli could also be systematically evaluated in a more controlled evaluation, as we modeled.

We evaluated participant identification and categorization accuracy when the target behaviors were not preceded by a clear antecedent event. Although a definitive answer requires experimentation, identifying and categorizing behavior that has a clear antecedent event should be easier given that most of our participants could identify and categorize our programmed behavior examples with high accuracy. Thus, our positive outcomes should give confidence to researchers who evaluate identification and categorization scores using examples of behavior with a known or clear antecedent event.

Researchers should also consider other variables that might impact caller accuracy as it is likely calls for service will continue to be used as a method of gathering data on atypical behavior, regardless of potential limitations, because substantial precedence exists in the

literature (e.g., Parmar et al., 2014; Schafer et al., 2010; Stolzenberg et al., 2017), and these data are relatively easy to obtain (Klinger & Bridges, 1997). We found that identification and categorization accuracy could be two variables of interest, but there are likely additional variables that can influence calls for service data, such as the time of day the observation happens, whether there is visible evidence of the moon's presence, and the recency to exposure of entertainment and media focusing on a belief in lunar influence. Moreover, our results are limited insofar as we obtained a relatively small sample size and included a limited number of observations per dependent variable; the latter potentially explains our low reliability scores. Thus, future replications and extensions with larger sample sizes are warranted before drawing definitive conclusions regarding the need for an alternative to calls for service data.

Beyond calls for service data, our results might have implications for the study of lunar effects in general. We found that participants who reported a preexisting belief in the moon's effect on behavior were more likely to attribute atypical behavior to the presence of a full moon. This is not a surprising outcome given the common connection of atypical behavior and the moon in lore. However, we also found that four of our participants, regardless of their self-reported belief, attributed typical behaviors to the presence of a full moon despite having correctly identified and categorized those behaviors. One explanation could be that each of these participants attributed the same typical behavior example to the presence of the full moon, but, upon further inspection of our results, this was not the case (data available from second author). Another explanation could be that these results were obtained by chance, which would be partially supported by our non-significant chi-square analysis of the relationship between belief and attribution for typical behaviors (Table 5). It could also be that other, idiosyncratic variables contributed to each participant's attribution (e.g., perceptions of control; Fishman & Husman,

2017). We are unaware of any other evaluations that have included a controlled assessment of attribution to the moon with typical behaviors to which we can draw comparisons, nor did we request participants to provide their rationale for their responses. Thus, future research into the conditions under which atypical behaviors are likely to be attributed to the full moon should include typical behaviors for comparative reasons, as well as open-ended response options to gather more specific data on the variables that impact attribution with respect to the moon.

Participants in the current study were not required to provide specific information about their age, ethnicity, or other major demographic variables (e.g., socioeconomic status, education level, etc.), but these variables might provide some explanatory power with respect to attribution in future studies. For example, research by Heine (2001) on cultural differences between East Asians and North Americans revealed that East Asians tend to look at the self with an external locus of control, while North Americans tend to look at the self with an internal locus of control. Individuals with an internal locus of control may be more likely to attribute behavior to the full moon because they believe the moon changed something within another individual (or themselves). In contrast, individuals with an external locus of control may believe something in the environment (e.g., the moon) caused the observed behavior, making it uncontrollable (Weiner, 2010; Fishman & Husman, 2017). Future researchers should therefore consider how demographic variables might influence attributions of behavior to a full moon.

Researchers might also consider more generally evaluating the conditions that support continued belief in lunar effects, which might include evaluations of physics knowledge, psychological biases (e.g., confirmation bias), sensationalism, and the entertainment value of a belief in lunar influence. Iosif and Ballon (2005) hypothesized that these variables might explain why belief in the moon-behavior relationship continues in the modern world, and our results

might partially support this. For example, more than half of our participants (59%) did not report that the moon influences ocean tides, which physics research has validated (e.g., Hansen, 1982; Lambeck, 1975; Zurbenko & Potrzeba, 2013). In contrast, these participants reported that the moon affects behavior (n = 37) or that the moon has no known effects (n = 29), both of which are beliefs unsubstantiated by science. Although potentially innocuous in the context of college students who are still undergoing training, similar results obtained with health and law enforcement officials could suggest that inaccurate beliefs in the moon's effects might influence professional behavior, which would be detrimental to practice.

Acknowledgments

We thank Kyah Ellefson and Serena Leppanen for their assistance gathering our data and for their feedback on an earlier version of this manuscript. We also thank Felicia Glaser, Makenzie Heatherly, Harrison Helton, Skye Johnson, Serena Leppanen, Katelynn Mobley, Bethany Munden, Yen Ngyuen, Patrick Pineda, Mallory Raymore, and Nathan Weber for their assistance in the creation of our video vignettes.

References

Aron, A., Coups, E. J., & Aron, E. N. (2013). *Statistics in psychology* (6th ed). Boston, MA: Pearson.

Bickis, M., Kelly, I. W., & Byrnes, G. F. (1995). Crisis calls and temporal and lunar variables: A comprehensive examination. *Journal of Psychology, 129,* 701-711. doi: 10.1080/00223980.1995.9914940

Biermann, T., Estel, D., Sperling, W., Bleich, S., Kornhuber, J., & Reulback, U. (2005). Influence of lunar phase on suicide: The end of a myth? A population-based study. *Chronobiology International, 22,* 1137-1143. doi: 10.1080/07420520500398114

Blackmon, S., & Catalina, D. (1973). The moon and the emergency room. *Perceptual and Motor Skills, 37,* 624-626.

Campbell, D. E., & Beets, J. L. (1978). Lunacy and the moon. *Psychological Bulletin, 85,* 1123-1129.

Chapman, S., & Morrell, S. (2000). Barking mad? Another lunatic hypothesis bites the dust. *British Medical Journal, 321,* 1561-1563. doi:10.1136/bmj.321.7276.1561

Chaput, J., Weippert, M., LeBlanc, A. G., Hjorth, M. F., Michaelsen, K. F., Katzmarzyk, P. T… Sjodin, A. M. (2016). Are children like werewolves? Full moon and its association with sleep and activity behaviors in an international sample of children. *Frontiers in Pediatrics, 4,* 1-6. doi: 10.3389/fped.2016.00024

Drum, M. W., Terry, C. L., & Hammonds, C. R. (1986). Lunar phase and acting-out behavior. *Psychological Reports, 59,* 987-990. doi: 10.2466/pr0.1986.59.2.987

Fishman, E. J., & Husman, J. (2017). Extending attribution theory: Considering students' perceived control of the attribution process. *Journal of Educational Psychology, 109,* pp. 559-573. doi: 10.1037/edu0000158

Hansen, K. S. (1982). Secular effects of oceanic tidal dissipation on the Moon's orbit and the Earth's rotation. *Reviews of Geophysics, 20,* 457-480. doi: 10.1029/RG020i003p00457

Heine, S. J. (2001). Self as cultural product: An examination of East Asian and North American selves. *Journal of Personality, 69,* 881-906. https://doi-org.proxy.consortiumlibrary.org/10.1111/1467-6494.696168

Iosif, A., & Ballon, B. (2005). Bad moon rising: The persistent belief in lunar connections to madness. *Canadian Medical Association Journal, 173,* 1498-1500. doi: 10.1503/cmaj.051119

Kazemi-Bajestani, S. M. R., Amirsadri, A., Samari, A. A., & Javanbakht, A. (2011). Lunar phase cycle and psychiatric hospital emergency visits, inpatient admissions and aggressive behavior. *Asian Journal of Psychiatry, 4,* 45-50, doi: 10.1016/j.ajp.2010.12.002

Kelly, I. W., & Martens, R. (1994). Geophysical variables and behavior: LXXVIII. Lunar phase and birthrate: an update. *Psychological Reports, 75,* 507-511. doi: 10.2466/pr0.1994.75.1.507

Klinger, D. A., & Bridges, G. S. (1997). Measurement error in calls-for-service as an indicator of crime. *Criminology, 35,* 705-726. doi: 10.1111/j.1745-9125.1997.tb01236.x

Lambeck, K. (1975). Effects of tidal dissipation in the oceans on the Moon's orbit and Earth's rotation. *Journal of Geophysical Research, 80,* 2971-2925. doi: 10.1029/JB080i020p02917

Laverty, W. H., & Kelly, I. W. (1998). Cyclical calendar and lunar patterns in automobile property accidents and injury accidents. *Perceptual and Motor Skills, 86,* 299-302. doi: 10.2466/pms.1998.86.1.299

Lerman, D. C., Tetreault, A., Hovanetz, A., Bellaci, E., Miller, J., Karp, H., . . . Toupard, A. (2010). Applying signal-detection theory to the study of observer accuracy and bias in behavioral assessment. *Journal of Applied Behavior Analysis, 43,* 195-213. doi:10.1901/jaba.2010.43-195

Lester, D., Brockopp, G. W., & Priebe, K. (1969). Association between a full moon and completed suicide. *Psychological Reports, 25,* 598.

Lieber, A. L., (1978). Human aggression and the lunar synodic cycle. *Journal of Clinical Psychiatry, 39,* 385-393.

Mander, A. E. (1947). *Logic for the millions*. New York, NY: Philosophical Library.

Parmar, V.S., Talikowska-Szymczak, E., Downs, E., Szymczak, P., Meiklejohn, E. & Groll, D. (2014). Effects of full-moon definition on psychiatric emergency department presentations. *International Scholarly Research Notice Emergency Medicine, 2014,* 1-6. doi: 10.1155/2014/398791

Pettigrew, C. G. (1985). Seasonality and lunacy of inmate behavior. *The Southern Psychologist, 2,* 41-46.

Pokorny, A. D., & Jachimczyk, J. (1974). The questionable relationship between homicides and the lunar cycle. *American Journal of Psychiatry, 131,* 827-829.

Purpura, P. P. (1979). Police activity and the full moon. *Journal of Police Science and Administration, 7,* 350-353.

Reiss, A. J. (1971). A policeman's lot. In *The Police and the Public* (pp. 1-62). New Haven, CT: Yale University Press.

Rotton, J. & Kelly, I. W. (1985). Much ado about the full moon: A meta-analysis of lunar-lunacy research. *Psychological Bulletin, 97*(2), 286-306. http://dx.doi.org.proxy.consortiumlibrary.org/10.1037/0033-2909.97.2.286

Rotton, J., Kelly, I. W., & Frey, J. (1983). Geophysical variables and behavior: X. Detecting lunar periodicities: Something old, new, borrowed, and true. *Psychological Reports, 52,* 111-116. Doi: 10.2466/pr0.1983.52.1.111

Roy, A., Biswas, T., & Roy, A. K. (2017). A structured review of relation behavior full moon and different aspects of human health. *SM Journal of Biometrics and Biostatistics, 1,* 1007-1010.

Schafer, J. A., Varano, S. P., Jarvis, J. P., & Cancino, J. M. (2010). Bad moon on the rise? Lunar cycles and incidents of crime. *Journal of Criminal Justice, 38,* 359-367. doi: 10.1016/j.jcrimjus.2010.04.003

Shermer, M. (2002). *Why people believe weird things: Pseudoscience, superstition, and other confusions of our time* (2nd ed.). New York, NY: St. Martin's Griffin.

Simón, A. (1998). Aggression in a prison setting as a function of lunar phases. *Psychological Reports, 31,* 110.

Stolzenberg, L., D'Alessio, S. J. & Flexon, J. L. (2017). A hunter's moon: The effect of moon illumination on outdoor crime. *American Journal of Criminal Justice, 42,* 188-197. doi: 10.1007/s12103-016-9351-9

Taylor, L. J., & Diespecker, D. D. (1972). Moon phase and suicide attempts in Australia. *Psychological Reports, 31,* 110.

Templer, D. I., & Veleber, D. M. (1980). The moon and madness: A comprehensive perspective. *Journal of Clinical Psychology, 36,* 865-868.

Templer, D. I., Veleber, D. M., & Brooner, R. K. (1982). Geophysical variables and behavior: VI. Lunar phase and accident injuries: A difference between night and day. *Perceptual and Motor Skills, 55,* 280-282. doi: 10.2466/pms.1982.55.1.280

Walters, E., Markley, R. P., & Tiffany, D. W. (1975). Lunacy: A type I error? *Journal of Abnormal Psychology, 84,* 715-717.

Weiner, B. (2010). The development of an attribution-based theory of motivation: A history of ideas. *Educational Psychologist, 45*(1), 28-36. https://doi.org/10.1080/00461520903433596

Weiskott, G. N. (1974). Moon phases and telephone counseling calls. *Psychological Reports, 35,* 752-754.

Weiskott, G. N., & Tipton, G. B. (1975). Moon phases and state hospital admissions. *Psychological Reports, 37,* 486.

Wilson, J. E., II, & Tobacyk, J. J. (1990). Lunar phases and crisis center telephone calls. *Journal of Social Psychology, 130,* 47-51. doi: 10.1080/00224545.1990.9922932

Zurbenko, I. G. & Potrzeba, A. L. (2013). Tides in the atmosphere. *Air Quality, Atmosphere, & Health, 6*(1), 39-46. DOI:10.1007/s11869-011-0143-6

Table 1

Location, Context, and Target Behaviors Programmed Across Video Vignettes

	Vignette #	Location	Context	Target Behavior
Typical Behaviors	1a	Classroom	Seven students sit at a table	Female student walks to rubbish bin and deposits a piece of paper
	2a	Lounge Area	Two male students sit on opposite couches with a small ottoman between them	Students play games on phones
	3a	Vending Machine	Male student standing at vending machine without money	Male student at vending machine asks student walking by for money
	4a	Common Area	Male and female student sit at a table	Students have a conversation
Atypical Behaviors	1b	Classroom	Seven students sit around a table	Female student sprints to the rubbish bin three times
	2b	Lounge Area	Two male students sit on opposite couches with a small ottoman between them	One student kicks over ottoman
	3b	Vending Machine	Female student stands at vending machine with wallet visible in her purse	Female student who walks by steals visible wallet from purse
	4b	Common Area	Male and female student sit at table	Female student hits male twice

Table 2

Number and Percentage of Correct and Incorrect Target Behavior Identifications

	Vignette #	Target Behavior	Number		Percentage	
			Correct	Incorrect	Correct	Incorrect
Typical Behaviors	1a	Female student walks to rubbish bin and deposits a piece of paper	109	5	95.6	4.4
	2a	Students play games on phones	110	4	96.5	3.5
	3a	Male student at vending machine asks student walking by for money	111	3	97.4	2.6
	4a	Students have a conversation	110	4	96.5	3.5
Atypical Behaviors	1b	Female student sprints to the rubbish bin three times	103	11	90.4	9.6
	2b	One student kicks over ottoman	106	8	93.0	7.0
	3b	Female student who walks by steals visible wallet from purse	111	3	97.4	2.6
	4b	Female student hits male twice	110	4	96.5	3.5

Note. The total number of presentations for each vignette was 114 (i.e., one presentation per participant).

Table 3

Number and Percentage of Correct and Incorrect Target Behavior Categorizations

	Vignette #	Target Behavior	Number		Percentage	
			Correct	Incorrect	Correct	Incorrect
Typical Behaviors	1a	Female student walks to rubbish bin and deposits a piece of paper	90	-	100	-
	2a	Students play games on phones	89	1	98.9	1.1
	3a	Male student at vending machine asks student walking by for money	60	30	66.7	33.3
	4a	Students have a conversation	87	3	96.7	3.3
Atypical Behaviors	1b	Female student sprints to the rubbish bin three times	90	-	100	-
	2b	One student kicks over ottoman	80	10	88.9	11.1
	3b	Female student who walks by steals visible wallet from purse	87	3	96.7	3.3
	4b	Female student hits male twice	87	3	96.7	3.3

Note. The total number of presentations for each vignette was 90 (i.e., one presentation per participant who had an identification score of 100%).

Table 4

Attribution to the Moon by General Belief in Lunar Effects for Atypical Behaviors

Attribution to Moon		Believe Moon Affects…			
		Behavior (*n* = 29)	Natural World (*n* = 37)	Nothing (*n* = 24)	Total
Yes (*n* = 71)					
	Row	36.6%	32.4%	31.0%	100%
	Column	89.7%	62.2%	91.7%	
No (*n* = 19)					
	Row	15.8%	73.7%	10.5%	100%
	Column	10.3%	37.8%	8.3%	
Total		100%	100%	100%	

Note. χ^2 (2, *N* = 90) = 10.59, p = 0.01

Table 5

Attribution to the Moon by General Belief in Lunar Effects for Typical Behaviors

Attribution to Moon		Believe Moon Affects…			
		Behavior (*n* = 29)	Natural World (*n* = 37)	Nothing (*n* = 24)	Total
Yes (*n* = 4)					
	Row	50.0%	25.0%	25.0%	100%
	Column	6.9%	2.7%	4.2%	
No (*n* = 86)					
	Row	31.4%	41.9%	26.7%	100%
	Column	93.1%	97.3%	95.8%	
Total		100%	100%	100%	

Note. χ^2 (2, *N* = 90) = 0.68, p = 0.71