Perceptions of Allegations of Repeated Victimization: The Roles of Event Frequency, Language Specificity, and Disclosure Delay

Mark D. Snow[1] , **Sonja P. Brubacher[2],**
Lindsay C. Malloy[1], and Kirk Luther[3]

Abstract

Although many forms of victimization are repeated (e.g., domestic violence), we know relatively little about the perceived credibility of adult claimants who allege repeated maltreatment. We examined the effects of Event Frequency (Single vs. Repeated), Language Specificity (Episodic vs. Generic), and Disclosure Delay (Immediate vs. Delayed) on laypersons' perceptions of claimant credibility. Participants ($N = 649$) read a mock interview transcript and provided subjective ratings (e.g., credibility, likelihood of suspect guilt, claimant responsibility). When the alleged abuse occurred a single time (vs. repeatedly), participants rated the interviewee as less blameworthy but no more (or less) credible. Exploratory findings indicated that female participants viewed the interviewee as more credible and less responsible than did male participants.

Keywords

repeated events, victimization, credibility, domestic violence, disclosure, victim blame

[1]Ontario Tech University, Oshawa, Canada
[2]Centre for Investigative Interviewing, Griffith Criminology Institute, Griffith University, Australia
[3]Carleton University, Ottawa, Canada

Corresponding Author:
Lindsay C. Malloy, Faculty of Social Science and Humanities, Ontario Tech University, 2000 Simcoe Street North, Oshawa, ON L1G 0C5, Canada.
Email: Lindsay.Malloy@ontariotechu.ca

Many forms of abusive behavior and misconduct involve recurrent victimization (Buzawa et al., 2017; Connolly & Read, 2006). Laycock (2001) described domestic violence (DV) as "the quintessential repeat crime" (p. 67). Although a large body of research has shown that repeated maltreatment can influence victims' ability to retrieve and report event-related details (for recent reviews, see Brubacher & Earhart, 2019; Snow et al., 2020), research on the perceived credibility of claimant allegations as a function of event frequency remains relatively scarce.

Victims of repeated maltreatment face multiple disadvantages. Continuous victimization can result in a range of personal and interpersonal consequences, including (but not limited to) physical injury, trauma symptomatology, and variable mental health symptoms (e.g., Herman, 1992; Mechanic et al., 2008). Victims of repeated maltreatment may also face barriers to successful prosecution (e.g., Hartwig et al., 2012). To ensure fairness to the accused, the criminal justice system generally requires the "particularization" of allegations—constituting a discrete event at a specific time and place (Guadagno et al., 2006; R. v. B. [G.], 1990; S. v. R., 1989; Woiwod & Connolly, 2017). Although the generic components of a recurring event— what *usually* happens—are often well-remembered, detailed memory reports of specific episodes are comparatively difficult to generate (Brubacher & Earhart, 2019; Snow et al., 2020). The difficulty of generating detailed memory reports of specific episodes is a substantive concern for many police investigators (e.g., when interviewing victims of repeated DV; Hartwig et al., 2012). Some research suggests that external observers perceive repeated-event memory reports as less credible than that of a single event (e.g., Connolly et al., 2008; Weinsheimer et al., 2017). Existing research in this area, however, remains relatively limited as no study to date has disentangled the effects of event frequency (single- vs. repeated-event), event report specificity (episodic vs. generic language), and the timing of event disclosure (immediate vs. delayed) on perceptions of claimant credibility.

Event Frequency

A handful of experimental investigations demonstrate that reporters of a repeated experience are disadvantaged when laypeople judge the credibility of single- and repeated-event memory reports. A community sample viewed children who reported a repeated play session as less credible than those who reported a single session, although this effect may have been driven by inconsistencies within the former's accounts rather than frequency of experience, per se (Connolly et al., 2008). Similarly, adults who reported a repeated event have been perceived as less cognitively competent, honest, and credible compared to

adults who reported a single event (food-tasting event; Weinsheimer et al., 2017), and less credible compared to adults reporting either a single event or a lie (healthy lifestyle session; Deck & Paterson, 2020). However, the effect of event frequency may depend upon the veracity of the speaker (Deck & Paterson, 2021). Importantly, the individuals evaluated in all but one study (Deck & Paterson, 2021) reported innocuous—arguably positive—experiences. There are several considerations unique to the victimization context that raise questions about the generalizability of these laboratory findings. For one, the perceived role of the claimant in causing, failing to stop, or delaying disclosure of the event(s) under investigation may be an important component of credibility assessment and legal decision-making (e.g., Ellison & Munro, 2009).

Elsewhere, researchers have examined perceptions of child abuse victims' credibility and responsibility regarding multiple allegations. Rogers et al. (2007) found no effect of abuse history on perceptions of victim culpability but did report an interaction effect such that female respondents rated the perpetrator as more culpable in the first-time offense as compared to the repeated-offense condition. Pozzulo et al. (2010) found no main effect of abuse frequency on ratings of credibility (but see Danby et al., 2021 who found that abuse frequency and disclosure affected credibility perceptions of children alleging physical abuse). However, Theimer and Hansen (2017) found that participants viewed a 15-year-old child sexual abuse (CSA) victim as more responsible for her abuse when she was described as having experienced five instances of abuse as compared to a single instance. The authors concluded that participants may have blamed the victim because they believed she should have done something immediately after the first incident to prevent further abuse. This pattern may be even more pronounced within an adult victim context; adults may be viewed as being more equipped, socially and developmentally, to protect themselves from repeated abuse. Results from outside the laboratory have been less consistent; however, with some retrospective studies of CSA cases finding that single abuse incidents were more likely than multiple abuse incidents to be deemed credible (e.g., Hershkowitz et al., 2018; Melkman et al., 2017), whereas other studies have found the opposite (e.g., Wang et al., 2019). Our aim was to extend this research by examining perceptions of an individual (adult) recounting single or repeated victimization.

Language Specificity

The linguistic representation of event memory differs as a function of event frequency (Schneider et al., 2011). People tend to use present tense, generic language when recalling repeated events (e.g., "I'm usually in bed when it happens") and past tense, episodic language when recalling single events ("I

was in bed when it happened"; Brubacher & Earhart, 2019). An analysis of field interviews with 5- to 13-year-olds alleging sexual abuse showed that as frequency of abuse increased, so did use of generic language about what usually happened (Brubacher et al., 2013). External observers often view repeated event memory reports as less compelling than single event reports (e.g., Weinsheimer et al., 2017), and this may be due to the differential linguistic structure of the reports. Thus, to some extent, the effects of event frequency on credibility perceptions may be confounded with language differences between single and repeated event speakers. Interestingly, Connolly et al. (2008) did not find that generic ("script") language was correlated with event frequency or overall credibility scores when undergraduate students and community laypersons rated the accounts of 4- to 7-year-olds who experienced a repeated event. Yet, the speech style of adult witnesses—such as their language specificity—may play a greater role in their perceived credibility compared to children (e.g., Ruva & Bryant, 2004).

Disclosure Delay

Many forms of victimization (e.g., DV, sexual assault) are underreported to the police (Conroy & Cotter, 2017; Morgan & Oudekerk, 2019). Among those who do report, delays are common (e.g., London et al., 2005; Ullman, 1996). Sexual assault victims often disclose weeks, months, or years after the event(s) (see Lanthier et al., 2018), and there is some evidence among CSA victims that increased abuse frequency is associated with increased disclosure delay (e.g., Smith et al., 2000; Wallis & Woodworth, 2020).

The extent to which disclosure is delayed may influence perceptions of claimant credibility (e.g., Danby et al., 2021; Ellison & Munro, 2009; Franiuk et al., 2020; Pozzulo et al., 2010). Balogh et al. (2003), for instance, examined the effects of delay on perceptions of sexual harassment disclosure using an audio recorded mock trial scenario. Participants viewed the claimant more positively and the alleged perpetrator as more likely to be guilty when she reported the harassment immediately as compared to after an 18-month delay. Similarly, Thompson et al. (2021) observed, within a sexual assault scenario, lower ratings of victim believability after a 10-year reporting delay, as compared to a 1-year delay. This pattern (of viewing delayed reports more negatively) may not hold across much longer delays (e.g., multiple decades), however (e.g., Fraser et al., 2021). Craig (2011) argued that although the Supreme Court of Canada has ruled against drawing adverse inferences regarding claimant credibility based on delayed disclosure (R. v. D. D., 2000), the assumption that "real" victims do not delay disclosure has persisted in many trial court decisions. Given that increased event frequency would generally require increased passage of time, it may be particularly

difficult in cases of repeated victimization to ascertain the relative impact of disclosure delay and event frequency on perceptions of credibility.

The Current Research

Criminal allegations may involve repeated events spanning a period of weeks, months, or even years. To date, research on perceptions of allegations of repeated versus single event victimization remains relatively scarce. Our goal was to advance understanding in this area by examining, within the context of a mock investigative interview, how allegations of assault (physical and sexual) by an intimate partner are perceived by third parties as a function of event frequency (i.e., whether the violence allegedly occurred a single time or repeatedly), language specificity (i.e., whether the interviewee recounts the events using generic or episodic language), and disclosure delay (i.e., whether the alleged victim discloses immediately or after a delay). Data and research materials are available on the Open Science Framework (https://osf. io/vafyj/?view_only=0fbe035d07cd477f9b5457a648f6d6f0).

Hypotheses

We expected participants to view the interviewee as less credible and more responsible when reporting repeated versus single victimization. Second, we predicted that the use of generic language would be associated with lower credibility than the use of episodic language. Third, we hypothesized that participants would view the interviewee as less credible and more responsible when she disclosed after a delay compared to immediately. Fourth, we predicted an Event Frequency × Language Specificity interaction such that the effect of language specificity on overall credibility would hold only for repeated event claimants. Fifth, we predicted an Event Frequency × Disclosure Delay interaction such that participants would rate the victim in the repeated event, delayed disclosure condition as more responsible than in all other conditions. We also conducted exploratory analyses to test for potential differences in perceptions of the claimant's cognitive ability, and various reporting and event factors (e.g., perceived stressfulness of the event, whether the reporting timeframe was reasonable).

Method

Participants

An *a priori* power analysis (Superpower package in R; Lakens & Caldwell, 2021) revealed a per group sample size of 80 participants to be the minimum

sample size needed to detect small main effects (Cohen's $f=.10$) at 80% power (alpha$=.05$) assuming large standard deviations ($SD\sim5$; power to detect main effects increases with less variability in the scores). This sample size also had the power (.80) to detect small-to-medium two-way interactions (Cohen's $f=.125$). Increasing the sample size in these simulations did not result in significant changes in the power to detect similar sized interaction effects (e.g., $n=100$ per group, Cohen's $f=0.115$, power of .80). Rather, a substantially larger sample size per group (e.g., $n=180$ or $N=1,440$) was needed to detect slightly smaller interaction effects (Cohen's $f=.10$, power of .81). Given the exploratory nature of the hypothesized interaction effects and the substantial additional cost in doubling the overall sample size for a small reduction in effect size, our final *a priori* sample was $N=640$ (80 per group).

Of the 813 responses recorded in Qualtrics, 649 were individuals who completed the survey successfully and passed the required manipulation and attention checks (i.e., questions 12, 19, and 20, see Supplemental Appendix). Thus, our final sample consisted of 649 participants ($M_{age}=39.49$, 358 males, 286 females, three "other", and two did not specify gender). Participants were adults from the general US population, the majority of whom self-identified as White ($n=504$), Asian ($n=64$), Black or African American ($n=59$), or Hispanic or Latino ($n=43$); participants could select more than one category. Regarding participants' highest completed level of education, the modal response was "undergraduate degree" ($n=277$). Participants' modal household income was $45,001 to $90,000 ($n=265$).

Materials and Design

We used a 2 (Event Frequency: Single vs. Repeated) $\times$ 2 (Language Specificity: Episodic vs. Generic) $\times$ 2 (Disclosure Delay: Immediate vs. Delayed) between-participants design. The experimental stimuli consisted of an online survey hosted on Qualtrics.com. The first page of the survey was an informed consent form that provided a basic overview of the present research. The second page contained several questions regarding participants' demographic characteristics. On the third page participants received the following instructions (bracketed text differed between conditions):

> The following is a brief excerpt of a police interview with a 28-year-old woman who has alleged that she was [repeatedly] physically and sexually assaulted by her boyfriend [with the last time being] [approximately 1 year ago/this past weekend]. Please take your time and read everything carefully.

The fourth page contained one of eight mock interview transcript excerpts. The transcript consisted of a police interview with a woman who alleged that she had been sexually and physically assaulted by her (now ex-) boyfriend. The incident(s) was described as happening either repeatedly or a single time, reported either immediately or after a 1-year delay, and in either episodic or generic language (i.e., generic conditions involved the use of present-tense and you-pronouns). When the woman alleged a single incident using generic language, her responses were in present tense (e.g., "I'm kind of sensing like something's wrong and it starts making me anxious"). All other aspects of the interview transcripts were held constant.

After reading the randomly assigned interview transcript, participants completed a survey assessing on 6-point scales their perceptions of the interviewee, including honesty and cognitive competence dimensions (items modified from past research; Connolly et al., 2008; Weinsheimer et al., 2017) as well as two items regarding the likelihood the suspect was guilty of physical assault and sexual assault. These survey questions appeared in a randomized order and contained an attention check item (i.e., "please select two on the scale below").

For the second part of the survey, participants received instructions that they should assume for the following questions 100% certainty that the events described in the interview really did occur and to rate the responsibility of the victim and perpetrator. This set of questions also appeared in a randomized order. Next, participants responded in randomized order to several manipulation check questions and rated the realism of the scenario on two separate items. All survey items are included in the Supplemental Appendix.

Procedure

We recruited participants via Cloud Research (formerly TurkPrime, https://www.cloudresearch.com/), an online participant sourcing platform. Cloud Research recruits participants from various panel providers (e.g., Amazon Mechanical Turk, Prime Panels) while using technical security measures (e.g., reCaptcha) and screening questions to prevent fraudulent respondents and bots. Interested individuals could enter via Amazon Mechanical Turk and select the present study if they wished to take part. After providing informed consent, they completed the survey outlined in the Materials and Design section. Finally, participants received debriefing information and compensation through Cloud Research. We preregistered the current study on the Open Science Framework (https://osf.io/5ekyf/?view_only=d99e6377f3ee443782 5d0eaee3ad3de5).

Results

Descriptive statistics as a function of experimental condition are reported in our Supplementary Materials.

Preliminary Analyses

To examine potential gender and age differences between groups we conducted chi-square analyses (gender) and a between-subjects ANOVA (age). Gender and age did not differ significantly as a function of experimental condition (ps $>$ .05). We also found no significant differences across conditions in participants' ratings of the seriousness of the event ($M = 5.80$, $SD = 0.52$) nor in their ratings of the extent to which events such as what was described in the interview often occur in real life ($M = 5.52$, $SD = 0.84$, ps $>$ .05).

The survey consisted of 17 questions about the interviewee's perceived credibility in the interview, and perceived responsibility for the events she described. We analyzed all 17 questions using principal axis factoring with Promax rotation (attention and manipulation checks excluded). As shown in Table 1, we obtained three factors: (1) Eight items loaded on the first factor, henceforth referred to as *Credibility* (all item loadings $>$ .40; Cronbach's $\alpha = .95$). Note that two additional items (perceived accuracy and stress) loaded on this factor but were removed due to cross-loadings; (2) Four items loaded on the second factor, henceforth referred to as *Cognitive Ability* (all item loadings $>$ .40; Cronbach's $\alpha = .75$). Note that one additional item (perceived accuracy) loaded on this factor but was removed due to cross-loading; (3) Two items loaded on the third factor, henceforth referred to as *Victim Blame* (all item loadings $>$ .40; Cronbach's $\alpha = .80$).

Main Analyses

There are two main deviations from our pre-registration plan. First, our analysis plan did not permit an explicit test of our hypotheses. Thus, our analysis plan has changed from between-subject ANOVAs to independent samples *t*-tests, the latter of which will allow us to test our hypotheses. Second, reviewers suggested that we conduct Bayesian analyses to determine whether there was support for the null hypotheses. Bayes factors were conducted using JASP 0.15.0.0 (obtained from https://jasp-stats.org) using the default Cauchy prior scale (0.707). Among its benefits, Bayes factors provide the relative evidence for one hypothesis/model over another, as well as the ability to quantify evidence in favor of the null hypothesis, and are not severely biased against the null hypothesis (Wagenmakers et al., 2018). Bayes factors were interpreted based on commonly used categories (see Lee & Wagenmakers, 2013).

Table 1. Factor Loadings for Survey Items.

Item	Factor		
	Credibility	Cognitive Ability	Victim Blame
How likely is it that the boyfriend (Mick) is guilty of sexual assault?	**.832**	-.090	.116
How likely is it that the boyfriend (Mick) is guilty of physical assault?	**.825**	-.052	.056
How believable was the interviewee (Ella)?	**.796**	.144	.017
How truthful was the interviewee (Ella)?	**.795**	.196	-.049
How honest was the interviewee (Ella)?	**.784**	.189	-.042
Overall, how credible was the interviewee (Ella)?	**.737**	.253	-.053
How likely is it that the interviewee (Ella) made up the event?	**-.649**	-.109	-.179
How responsible is the boyfriend (Mick) for what happened?	**.519**	-.260	.313
How accurately did the interviewee (Ella) recall the event(s)?	**.508**	**.478**	-.170
How stressful was the experience described by the interviewee (Ella)?	**.406**	-.131	.385
How intelligent was the interviewee (Ella)?	-.053	**.729**	.090
How likeable was the interviewee (Ella)?	.183	**.596**	.007
How likely is it that the interviewee (Ella) made some mistakes in her account?	.077	**-.545**	-.241
How confused was the interviewee (Ella)?	-.105	**-.527**	-.011
The interviewee (Ella) could have prevented the abuse from happening.	.033	-.179	**-.776**
How responsible is the interviewee (Ella) for what happened?	-.028	-.057	**-.732**
The interviewee (Ella) waited too long before reporting.	-.021	-.248	-.351

Note. Values are pattern matrix factor loadings. Loadings ≥ .40 are bolded.

Hypothesis 1. The effect of Event Frequency on Credibility, $t(647)=-1.46$, $p=.072$, $d=-0.12$, 95% CI [-0.27, 0.04], was not significant. A Bayesian independent samples t-test revealed strong evidence in favor of the null hypothesis ($BF_{01}=27.06$); specifically, the data are 27 times more likely under the null hypothesis (i.e., no significant difference in Credibility ratings as a function of event frequency) than the alternative hypothesis.

The effect of Event Frequency on Victim Blame, $t(647)=-3.08$, $p=.001$, $d=-0.24$, 95% CI [-0.40, -0.09], was significant. Further, Bayesian analysis revealed strong evidence in favor of the alternative hypothesis ($BF_{10}=17.65$). Participants rated the interviewee as more blameworthy when the alleged abuse occurred repeatedly ($M=1.94$, $SD=1.24$) compared to a single incident ($M=1.67$, $SD=1.04$).

Hypothesis 2. The effect of Language Specificity on Credibility, $t(647)=1.42$, $p=.922$, $d=0.11$, 95% CI [-0.04, 0.27], was not significant. A Bayesian independent samples t-test revealed strong evidence for the null hypothesis ($BF_{01}=26.55$); specifically, the data are nearly 27 times more likely under the null hypothesis (i.e., no significant difference in Credibility ratings as a function of language specificity) than the alternative hypothesis.

Hypothesis 3. The effect of Disclosure Delay on Credibility, $t(647)=1.61$, $p=.054$, $d=0.13$, 95% CI [-0.03, 0.28], was not significant; Bayesian analysis provided inconclusive evidence for either hypothesis ($BF_{10}=0.58$); more data are needed before a conclusion can be drawn about potential effects. The effect of Victim Blame, $t(647)=-0.81$, $p=.210$, $d=-0.06$, 95% CI [-0.22, 0.09], was not significant. Bayesian analysis provided moderate evidence for the null hypothesis ($BF_{01}=5.27$); that is, the data are five times more likely under the null hypothesis (i.e., no meaningful difference in Victim Blame ratings as a function of disclosure delay) than the alternative hypothesis.

Hypothesis 4. A 2 (Event Frequency) $\times$ 2 (Language Specificity) between-subjects ANOVA indicated that the interaction was not significant, $F(1, 645)=.001$, $p=.977$, $\eta_p^2=.00$. The Bayesian analysis revealed a $BF_{excl}=27.64$, indicating there is 27.64 times higher probability for support of the null effect than the alternative hypothesis.

Hypothesis 5. A 2 (Event Frequency) $\times$ 2 (Disclosure Delay) between-subjects ANOVA indicated that the interaction was not significant, $F(1, 645)=.34$, $p=.561$, partial $\eta_p^2=.00$. The Bayesian analysis revealed a $BF_{excl}=18.01$, indicating there is 18.01 times higher probability for support of the null effect than the alternative hypothesis.

Exploratory Analyses

Cognitive ability. To explore potential differences in perceptions of cognitive ability, we conducted a 2 (Event Frequency) × 2 (Disclosure Delay) × 2 (Language Specificity) between-subjects ANOVA. Results revealed that cognitive ability ($M = 4.44$, $SD = 0.90$) did not differ significantly as a function of experimental conditions ($ps > .05$). Bayes factors (BF_{01}) indicated moderate evidence in favor of the null hypotheses for the main effects of Language Specificity ($BF_{01} = 4.49$), Event Frequency ($BF_{01} = 5.83$), and Disclosure Delay ($BF_{01} = 8.12$). For the interaction between Event Frequency and Language Specificity, the Bayesian analysis revealed a $BF_{excl} = 135.43$, indicating there is 135.43 times higher probability for support of the null effect than alternative hypothesis; the interaction between Event Frequency and Disclosure Delay revealed a $BF_{excl} = 206.50$; the interaction between Disclosure Delay and Language Specificity revealed a $BF_{excl} = 178.35$; and the three-way interaction revealed a $BF_{excl} = 37203.25$. Across conditions, participants provided moderately high ratings of the interviewee's cognitive ability (means ranged from 4.28 to 4.58).

Other individual items. Participants rated the claimant's recall accuracy as being relatively high overall ($M = 4.91$, $SD = 0.90$), and ratings did not differ significantly as a function of experimental condition ($ps > .05$). The Bayes factors indicated that there is moderate evidence in favor of the Disclosure Delay null model ($BF_{01} = 4.35$), and strong evidence in favor of the Language Specificity and Event Frequency null models ($BF_{01} = 10.29$ and 10.68, respectively). For the interaction between Event Frequency and Disclosure Delay, the Bayesian analysis revealed a $BF_{excl} = 230.51$, indicating there is 230.51 times higher probability for support of the null effect than alternative hypothesis; the interaction between Event Frequency and Language Specificity revealed a $BF_{excl} = 407.79$; the interaction between Disclosure Delay and Language Specificity revealed a $BF_{excl} = 196.73$; and the three-way interaction revealed a $BF_{excl} = 57881.30$.

Participants rated the experience described by the interviewee as being highly stressful ($M = 5.72$, $SD = 0.65$), and ratings did not differ significantly as a function of experimental condition ($ps > .05$). The Bayes factors indicated that there is moderate evidence in favor of Disclosure Delay and Event Frequency null models ($BF_{01} = 3.98$ and 9.02, respectively), and strong evidence in favor of the Language Specificity null model ($BF_{01} = 10.29$). For the interaction between Event Frequency and Language Specificity, the Bayesian analysis revealed a $BF_{excl} = 174.63$, indicating there is 174.63 times higher probability for support of the null effect than alternative hypothesis; the

interaction between Event Frequency and Disclosure Delay revealed a $BF_{excl} = 176.63$; the interaction between Disclosure Delay and Language Specificity revealed a $BF_{excl} = 202.08$; and the three-way interaction revealed a $BF_{excl} = 36316.81$.

Participants generally did not believe that the interviewee waited too long before reporting ($M = 2.73$, $SD = 1.77$). However, consistent with our hypotheses, there were significant main effects of Disclosure Delay, $F(1, 641) = 66.80$, $p < .001$, $BF_{10} = 255.22$, $d = 0.64$ and Event Frequency, $F(1, 641) = 8.66$, $p = .003$, $BF_{10} = 3.67$, $d = 0.22$. Participants were less likely to think the interviewee waited too long before reporting when she reported after a shorter delay (i.e., this past weekend; $M = 2.18$, $SD = 1.56$) as compared to a longer delay (i.e., 1 year ago; $M = 3.26$, $SD = 1.81$). Participants were also less likely to think she waited too long when the alleged abuse occurred only a single time ($M = 2.53$, $SD = 1.78$) as compared to when it was repeated ($M = 2.92$, $SD = 1.75$). The effects of Disclosure Delay and Event Frequency were small to medium. No other main effects or interactions emerged ($ps > .05$).

Perceptions of language specificity. Participants rated the extent to which the interviewee's account was episodic versus generic. We sought to explore whether such ratings differed as a function of Language Specificity. Interestingly, ratings of the interviewee's specificity did not differ as a function of the Language Specificity manipulation ($p = .91$, $BF_{01} = 11.42$), but they differed as a function of Event Frequency, $F(1, 641) = 12.80$, $p < .001$, $BF_{10} = 47.06$, $d = 0.29$. Participants rated the interviewee's account as having greater specificity when the alleged abuse occurred a single time ($M = 5.20$, $SD = 0.82$) than when it occurred repeatedly ($M = 4.96$, $SD = 0.85$). This was a small effect.

Participant gender and age. We also conducted exploratory analyses to examine potential relations between participant gender, age, and their survey responses. We conducted a series of independent samples *t*-tests with participant gender as the independent variable (three people who self-identified as other and two who did not specify their gender could not be included in these comparisons). On the first factor (Credibility), we found that female participants ($M = 5.43$, $SD = 0.68$) rated the interviewee as more credible than did male participants ($M = 5.27$, $SD = 0.81$, $d = 0.22$, 95% CI [0.06, 0.37]), $t(642) = 2.71$, $p = .007$, $BF_{10} = 3.15$. On the second factor (cognitive ability), we found that female participants ($M = 4.55$, $SD = 0.87$) provided higher ratings of the interviewee's cognitive ability than did male participants ($M = 4.37$, $SD = 0.91$, $d = 0.20$, 95% CI [0.05, 0.36]), $t(642) = 2.53$, $p = .012$, $BF_{10} = 2.01$. Regarding the third factor (Victim Blame), results revealed that female

participants ($M=1.63$, $SD=1.00$) tended to blame the interviewee to a lesser extent than did male participants ($M=1.93$, $SD=1.23$, $d=0.26$, 95% CI [0.11, 0.42]), $t(642)=3.29$, $p=.001$, $BF_{10}=17.01$. Note that all the gender differences were small effects.

Regarding the individual items, no significant gender differences emerged in participants' ratings of perceived accuracy ($p=.070$, $BF_{01}=2.28$) or stress ($p=.153$, $BF_{01}=4.16$). However, results revealed a significant gender difference in participants' ratings of whether the interviewee waited too long before reporting, $F(1, 642)=15.04$, $p<.001$, $BF_{10}=128.10$. Female participants ($M=2.42$, $SD=1.63$) disagreed to a greater extent than did male participants ($M=2.96$, $SD=1.84$, $d=-0.31$, 95% CI [−0.46, −0.15]) that the interviewee had waited too long before reporting. This was a small-medium effect. Finally, we conducted exploratory correlational analyses to examine potential associations between participant age and survey responses. Except for a small correlation between participant age and ratings of the likelihood that the claimant made some mistakes in her account ($r=.10$, $p=.012$), participant age was not significantly correlated with participants' survey ratings (ps > .05).

Discussion

As various forms of victimization occur repeatedly over time (e.g., Buzawa et al., 2017), allegations of repeated events are of frequent concern in criminal investigations and victim interviews (Hartwig et al., 2012). Our aim was to shed light on the independent effects of event frequency, language specificity, and disclosure delay, on laypeoples' perceptions of credibility and responsibility.

Contrary to research where adult participants described healthy lifestyle sessions and food tastings (Deck & Paterson, 2020; Weinsheimer et al., 2017)—and despite increased perceptions of victim blame—we did not find evidence that event frequency significantly affected perceptions of credibility when vignettes described an adult female claiming sexual and physical assault. However, our findings are in line with what Pozzulo et al. (2010) found, with no main effect of abuse frequency on ratings of credibility. Our findings extend beyond those reported by Pozzulo and colleagues; specifically, while the null hypothesis cannot be proven true (Harms & Lakens, 2018), our data suggest strong evidence that there is no meaningful difference in credibility as a function of event frequency, at least for the particular scenario involving the alleged sexual assault of an adult. Yet, participants generally found the interviewee more responsible for the event(s) —and expressed stronger agreement that she waited too long to report—when the event(s)

were described as having occurred repeatedly compared to just once, in partial support of our first hypothesis and consistent with research with sexual abuse vignettes involving a 15-year-old victim (Theimer & Hansen, 2017).

Participants appeared to be more influenced by descriptions of victimization frequency than the language used by the interviewee. Language specificity had no systematic effects on participants' responses. Ratings of language specificity did not differ significantly between episodic and generic conditions, suggesting that participants were not consciously aware of the linguistic differences in the claimants' reports between conditions, yet participants rated the claimant's description as more specific (and less generic) in the single relative to the repeated event condition. We argue against a simple manipulation failure interpretation for three reasons: (1) the manipulation was as strong as it could be, in that we changed instances of specific past tense language to present tense and incorporated you-pronoun use, (2) there is mounting evidence that language specificity is not what drives credibility differences between repeated and single event speakers when they are children (Connolly et al., 2008) or adults describing a pleasant event (Deck & Paterson, 2020; Weinsheimer et al., 2017), and (3) the Bayesian analysis of the main effect of language specificity on credibility revealed that the data are nearly 27 times more likely supporting the null hypothesis of no significant difference in credibility ratings based on language specificity. By explicitly manipulating language specificity in vignettes while holding constant report consistency and other features, we can be relatively confident that this is a true null effect; nevertheless, future replication efforts may help to further elucidate the boundaries of this effect.

Many victims report abuse after a delay (e.g., Ullman, 1996). We did not find that delay affected perceptions of credibility or victim responsibility. Participants were, however, less likely to agree with the statement that the interviewee waited too long before reporting when the delay was just a few days compared to a year (see also Lucarini et al., 2020). Even in the one-year delay condition, however, participants' average level of agreement was moderate ($M=3.26$) on the 6-point scale. As victims of abuse may delay their disclosure for years or decades, further research may wish to address longer delays (Lanthier et al., 2018).

In summary, allegations of repeated victimization and a relatively lengthy delay to disclosure had relatively minor effects on perceptions of the interviewee in terms of somewhat higher blame attribution and propensity to think she waited too long to report. Promisingly, our analyses (i.e., Bayes factors, examination of effect sizes) suggest that there were few significant differences regarding how claimants were perceived in terms of their credibility or blameworthiness based on whether they experienced single or repeated

events, disclosed immediately or with a delay, and the specificity with which they describe the event(s). There was, however, one factor that had more impact on interviewee perceptions than any other—participant gender. Participants identifying as male viewed the interviewee as less credible, less cognitively competent, and more responsible for what happened than participants identifying as female. This gender difference is consistent with existing findings in the literature, and a multilevel meta-analysis including studies from 1976 to 2019 indicates it is persistent (Persson & Dhingra, 2020).

Limitations and Future Directions

To increase the generalizability of our findings, we encourage researchers to examine these issues in future studies using scenarios that introduce greater variation in credibility (e.g., interviewee is a drug dealer rather than a female university student). The use of additional scenarios would help to further clarify the potential credibility effects of event frequency and related variables in future research. We also note that none of our observed effects were particularly large. Thus, it is important not to overstate the role of the present independent variables in real-world cases. The generalizability of the current study may also be limited by the characteristics of the current sample (i.e., predominantly White North American adults) and characteristics of the claimant (i.e., woman in a heterosexual relationship). In future studies, researchers may wish to examine additional demographic and cultural factors (e.g., socioeconomic status, sexual orientation, urban vs. rural residence) that may moderate participants' perceptions of DV and claimant credibility.

Conclusion

Our data raise some concerns regarding the perceived blameworthiness of individuals who report repeated (relative to single) victimization and report after a delay and provide further evidence of the persistence of gender differences in credibility perceptions of females alleging sexual assault. On a positive note, although extant literature indicated lower credibility perceptions for reporters of repeated versus single innocuous events (Connolly et al., 2008; Deck & Paterson, 2020; Weinsheimer et al., 2017), such findings were not evident in the present study for an alleged victim of physical and sexual assault—instead, our findings suggest that claimants will be viewed as credible and not responsible for their victimization, regardless of several key case characteristics. We would like to see replication of our results, as well as extensions of the current research (e.g., using other scenarios), before

providing concrete recommendations to practitioners based on our findings. Preliminarily, our results can be interpreted by professionals in the legal system as suggesting that factors of repeated victimization and delayed reporting may not negatively impact complainants' credibility, at least under some circumstances.

ORCID iD

Mark D. Snow https://orcid.org/0000-0002-5073-0752

Supplemental Material

Supplemental material for this article is available online.

References

Balogh, D. W., Kite, M. E., Pickel, K. L., Canel, D., & Schroeder, J. (2003). The effects of delayed report and motive for reporting on perceptions of sexual harassment. *Sex Roles: A Journal of Research*, *48*(7–8), 337–348. https://doi.org/10.1023/A:1022990530657

Brubacher, S. P., & Earhart, B. (2019). Investigative interviewing about repeated experiences. In J. J. Dickinson, N. S. Compo, R. N. Carol, B. L. Schwartz, & M. R. McCauley (Eds.), *Evidence-based investigative interviewing: Applying cognitive principles* (pp. 216–233). Routledge.

Brubacher, S. P., Malloy, L. C., Lamb, M. E., & Roberts, K. P. (2013). How do interviewers and children discuss individual occurrences of alleged repeated abuse in forensic interviews? *Applied Cognitive Psychology*, *27*(4), 443–450. https://doi.org/10.1002/acp.2920

Buzawa, E. S., Buzawa, C. G., & Stark, E. D. (2017). *Responding to domestic violence: The integration of criminal justice and human services* (5th ed.). SAGE.

Connolly, D. A., Price, H. L., Lavoie, J. A. A., & Gordon, H. M. (2008). Perceptions and predictors of children's credibility of a unique event and an instance of a repeated event. *Law and Human Behavior*, *32*(1), 92–112. https://doi.org/10.1007/s10979-006-9083-3

Connolly, D. A., & Read, J. D. (2006). Delayed prosecutions of historic child sexual abuse: Analyses of 2064 Canadian criminal complaints. *Law and Human Behavior*, *30*(4), 409–434. https://doi.org/10.1007/s10979-006-9011-6

Conroy, S., & Cotter, A. (2017). *Self-reported sexual assault in Canada, 2014*. Canadian Centre for Justice Statistics. https://www150.statcan.gc.ca/n1/en/pub/85-002-x/2017001/article/14842-eng.pdf?st=rK1X3za4

Craig, E. (2011). The relevance of delayed disclosure to complaint credibility in cases of sexual offense. *Queen's Law Journal*, *36*, 551–586.

Danby, M. C., Sharman, S. J., & Klettke, B. (2021). Factors influencing the perceived credibility of children alleging physical abuse. *Psychiatry, Psychology and Law*, *29*(3), 456–470. https://doi.org/10.1080/13218719.2021.1917012

Deck, S. L., & Paterson, H. M. (2020). Liars are perceived as more credible than truth-tellers who recall a repeated event. *Applied Cognitive Psychology*, *34*(3), 643–653. https://doi.org/10.1002/acp.3665

Deck, S. L., & Paterson, H. M. (2021). The perceived credibility of repeated-event witnesses depends upon their veracity. *Psychiatry, Psychology and Law*, *29*(4), 577–592. https://doi.org/10.1080/13218719.2021.1956382

Ellison, L., & Munro, V. E. (2009). Reacting to rape: Exploring mock jurors' assessments of complainant credibility. *British Journal of Criminology*, *49*(2), 202–219. https://doi.org/10.1093/bjc/azn077

Franiuk, R., Luca, A., & Robinson, S. (2020). The effects of victim and perpetrator characteristics on ratings of guilt in a sexual assault case. *Violence Against Women*, *26*(6–7), 614–635. https://doi.org/10.1177/1077801219840439

Fraser, B. M., Pica, E., & Pozzulo, J. D. (2021). The effect of delayed reporting on mock-juror decision-making in the era of #MeToo. *Journal of Interpersonal Violence*, *37*(13–14), NP11791–NP11810. https://doi.org/10.1177/0886260521997464

Guadagno, B. L., Powell, M. B., & Wright, R. (2006). Police officers and legal professionals' perceptions regarding how children are, and should be, questioned about repeated abuse. *Psychiatry, Psychology and Law*, *13*(2), 251–260. https://doi.org/10.1375/pplt.13.2.251

Harms, C., & Lakens, D. (2018). Making 'null effects' information: Statistical techniques and inferential frameworks. *Journal of Clinical and Translational Research*, *30*(3), 382–393. https://doi.org/10.18053/jctres.03.2017s2.007

Hartwig, M., Dawson, E. C., Wrede, O., & Ask, K. (2012). Interviewing victims of repeated domestic violence: Investigators' beliefs and strategies. *Psychiatry, Psychology and Law*, *19*(5), 672–681. https://doi.org/10.1080/13218719.2011.615814

Herman, J. L. (1992). Complex PTSD: A syndrome in survivors of prolonged and repeated trauma. *Journal of Traumatic Stress*, *5*(3), 377–391. https://doi.org/10.1002/jts.2490050305

Hershkowitz, I., Melkman, E. P., & Zur, R. (2018). When is a child's forensic statement deemed credible? A comparison of physical and sexual abuse cases. *Child Maltreatment*, *23*(2), 196–206. https://doi.org/10.1177/1077559517734059

Lakens, D., & Caldwell, A. R. (2021). Simulation-based power analysis for factorial analysis of variance designs. *Advances in Methods and Practices in Psychological Science*, *4*(1), 1–14. https://doi.org/10.1177/2515245920951503

Lanthier, S., Du Mont, J., & Mason, R. (2018). Responding to delayed disclosure of sexual assault in health settings: A systematic review. *Trauma, Violence & Abuse*, *19*(3), 251–265. https://doi.org/10.1177/1524838016659484

Laycock, G. (2001). Hypothesis based research: The repeat victimization story. *Criminology & Criminal Justice*, *1*(1), 59–82. https://doi.org/10.1177/1466802501001001004

Lee, M. D., & Wagenmakers, E. J. (2013). *Bayesian data analysis for cognitive science: A practical course*. Cambridge University Press.

London, K., Bruck, M., Ceci, S. J., & Shuman, D. W. (2005). Disclosure of child sexual abuse: What does the research tell us about the ways that children tell? *Psychology, Public Policy, and Law*, *11*(1), 194–226. https://doi.org/10.1037/1076-8971.11.1.194

Lucarini, A., Suitner, C., Brown, R., Craig, M. A., Knowles, E. D., & Casara, B. G. S. (2020). The# MeTooLate Effect: Victim blame and trust denial for sexual harassment not immediately reported. *Personality and Individual Differences*, *167*, 110240. https://doi.org/10.1016/j.paid.2020.110240

Mechanic, M. B., Weaver, T. L., & Resick, P. A. (2008). Mental health consequences of intimate partner abuse: A multidimensional assessment of four different forms of abuse. *Violence Against Women*, *14*(6), 634–654. https://doi.org/10.1177/1077801208319283

Melkman, E. P., Hershkowitz, I., & Zur, R. (2017). Credibility assessment in child sexual abuse investigations: A descriptive analysis. *Child Abuse & Neglect*, *67*, 76–85. https://doi.org/10.1016/j.chiabu.2017.01.027

Morgan, R. E., & Oudekerk, B. A. (2019). *Criminal victimization, 2018 (NCJ 253043)*. US Department of Justice. https://www.bjs.gov/content/pub/pdf/cv18.pdf

Persson, S., & Dhingra, K. (2020). Attributions of blame in stranger and acquaintance rape: A multilevel meta-analysis and systematic review. *Trauma, Violence, & Abuse*, *23*(3), 795–809. https://doi.org/10.1177/1524838020977146

Pozzulo, J. D., Dempsey, J. L., & Crescini, C. (2010). Factors affecting juror decisions in historic child sexual abuse cases involving continuous memories. *Criminal Justice and Behavior*, *37*(9), 951–964. https://doi.org/10.1177/0093854810373587

Rogers, P., Josey, N., & Davies, M. (2007). Victim age, attractiveness and abuse history as factors in the perception of a hypothetical child sexual abuse case. *Journal of Sexual Aggression*, *13*(2), 121–137. https://doi.org/10.1080/13552600701644835

Ruva, C. L., & Bryant, J. B. (2004). The impact of age, speech style, and question form on perceptions of witness credibility and trial outcome. *Journal of Applied Social Psychology*, *34*(9), 1919–1944. https://doi.org/10.1111/j.1559-1816.2004.tb02593.x

R. v. B. (G.), 2 S. C. R. 30 (1990).

R. v. D. D., S. C. J. No. R. 275 (2000).

Schneider, L., Price, H. L., Roberts, K. P., & Hedrick, A. M. (2011). Children's episodic and generic reports of alleged abuse. *Applied Cognitive Psychology*, *25*(6), 862–870. https://doi.org/10.1002/acp.1759

Smith, D. W., Letourneau, E. J., Saunders, B. E., Kilpatrick, D. G., Resnick, H. S., & Best, C. L. (2000). Delay in disclosure of childhood rape: Results from a national survey. *Child Abuse & Neglect*, *24*(2), 273–287. https://doi.org/10.1016/S0145-2134(99)00130-1

Snow, M. D., Malloy, L. C., Brubacher, S. P., & Sutherland, J. E. (2020). Memory for repeated sexual misconduct: Does repetition matter? In J. Pozzulo, E. Pica, & C. Sheahan (Eds.), *Memory and sexual misconduct: Psychological research for criminal justice* (pp. 42–70). Taylor & Francis.

S. v. R., 168 CLR 266 (1989).

Theimer, K., & Hansen, D. J. (2017). Attributions of blame in a hypothetical child sexual abuse case: Roles of behavior problems and frequency of abuse. *Journal of Interpersonal Violence*, *35*(11–12), 2142–2163. https://doi.org/10.1177/0886260517716943

Thompson, L. E., Pica, E., & Pozzulo, J. (2021). Jurors' decision making in a sexual assault trial: The influence of victim age, delayed reporting, and multiple allegations. *American Journal of Forensic Psychology*, *39*(2), 19–46.

Ullman, S. E. (1996). Social reactions, coping strategies, and self-blame attributions in adjustment to sexual assault. *Psychology of Women Quarterly*, *20*(4), 505–526. https://doi.org/10.1111/j.1471-6402.1996.tb00319.x

Wagenmakers, E. J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., Selker, R., Gronau, Q. F., ŠMíra, M., Epskamp, S., Matzke, D., Rouder, J. N., & Morey, R. D. (2018). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review*, *25*(1), 35–57. https://doi.org/10.3758/s13423-017-1343-3

Wallis, C. R. D., & Woodworth, M. D. (2020). Child sexual abuse: An examination of individual and abuse characteristics that may impact delays of disclosure. *Child Abuse & Neglect*, *107*, 104604. https://doi.org/10.1016/j.chiabu.2020.104604

Wang, L., Hung, Y., Chow, P. C., Chu, C., Li, H., Lu, T., & Tsai, C. (2019). Credibility judgment predictors for child sexual abuse reports in forensic psychiatric evaluations. *Psychiatry Investigation*, *16*(2), 139–144. https://doi.org/10.30773/pi.2018.12.03

Weinsheimer, C. C., Coburn, P. I., Chong, K., MacLean, C. L., & Connolly, D. A. (2017). Perceptions of credibility for a memory report of a single versus repeated event. *Applied Cognitive Psychology*, *31*(4), 414–423. https://doi.org/10.1002/acp.3340

Woiwod, D. M., & Connolly, D. A. (2017). Continuous child sexual abuse: Balancing defendants' rights and victims' capabilities to particularize individual acts of repeated abuse. *Criminal Justice Review*, *42*(2), 206–225. https://doi.org/10.1177/0734016817704700

Author Biographies

Mark D. Snow, MSc, is a PhD Candidate in Forensic Psychology at Ontario Tech University. His research interests are broadly concerned with investigative interviewing and his dissertation focuses on the effects of emotional memory in eyewitness contexts.

Sonja P. Brubacher, PhD, is an Adjunct Senior Research Fellow at the Centre for Investigative Interviewing, Griffith University. Her research interests include memory for repeated experiences, developmental differences in the effectiveness of theoretically-based interview procedures (including supportive behaviors and interview preparatory tasks), and the evaluation of those practices in the field.

Lindsay C. Malloy, PhD, is an Associate Professor at Ontario Tech University specializing in developmental and forensic psychology. Her research addresses questions concerning children's and adolescents' disclosure of negative or traumatic experiences, investigative interviewing and interrogation techniques, and implications of research findings for the legal system.

Kirk Luther, PhD, is an Assistant Professor at Carleton University. His research interests include safeguarding legal rights for adults and youth, and advancing research and practice on investigative interviewing.