Implausibility Inhibits But Does Not Eliminate False Autobiographical Beliefs

Alan Scoboria
Giuliana Mazzoni
Josée Jarry
Dana Shapero

1 Department of Psychology, University of Windsor, 401 Sunset, Windsor, ON, N9B 3P4, Canada (scoboria@uwindsor.ca, jjarry@uwindsor.ca, shapero@uwindsor.ca)
2 Department of Psychology, University of Hull, Hull, UK HU6 7RX (g.mazzoni@hull.ac.uk)

In Press: Canadian Journal of Experimental Psychology (as of September, 2012). This version is not the publication of record, and may differ slightly from the published version.
Abstract

The role of the plausibility of suggested events in the formation of false autobiographical beliefs and memories has long been debated. In two studies, the shape of the relationship between pre-suggestion personal plausibility and the development of post-suggestion false beliefs was examined. Participants rated personal plausibility and autobiographical belief for childhood events. They later received a suggestion that an unlikely event occurred during their childhood and provided post-suggestion ratings. The best fit was a curvilinear relationship between plausibility and belief, with the lowest risk for false belief at the plausibility scale floor. Above this threshold, the risk for false belief increased sharply and remained similar across all other levels of plausibility. A minority of those who initially viewed the event as highly implausible showed increased beliefs; this was accompanied by large increases in personal plausibility. We conclude that only extreme implausibility inhibits suggestion induced false autobiographical beliefs, unless suggestions cause increases in plausibility ratings.
Implausibility Inhibits But Does Not Eliminate False Autobiographical Beliefs

The plausibility of suggested events is associated with the formation of false autobiographical beliefs and memories (Scoboria, Mazzoni, Kirsch, & Jimenez, 2006; van Golde, Sharman, & Candel, 2010). While it is known that plausibility can change and that small increases in plausibility can potentiate increases in false beliefs (Mazzoni, Loftus, & Kirsch, 2001), the degree of plausibility required to yield false beliefs and memories as well as the shape of the relationship between initial plausibility and false beliefs remains unknown. Theories of memory formation are consistent in arguing that events must be viewed as *sufficiently plausible* before being attributed to past experience (Pezdek, Finger, & Hodge, 1997; Scoboria, Mazzoni, Kirsch, & Relyea, 2004; Hyman & Loftus, 1998), but what ‘sufficiently plausible’ means has yet to be defined.

Pezdek and colleagues (1997) and Pezdek and Hodge (1999) proposed that suggestions result in false memories only when events are plausible. In their studies, false memories were more likely for a plausible event (i.e., lost in a mall) than an implausible event (i.e., receiving an enema). Pezdek, Blandón-Gitlin and Gabbay (2006) reported that when people are told that an event occurs infrequently, the suggestion that the event occurred is ineffective. The implicit argument is that the likelihood of false beliefs and memories increases with the plausibility of events. Because subjective plausibility was not measured in these studies, this proposed relationship has not been empirically examined.

Sharman and Scoboria (2009) examined the effects of imagining events selected by the researchers as low, moderate, or high in plausibility on belief and memory ratings. Memory ratings were more likely to increase for high plausibility, followed by moderate, and finally by low plausibility events. Because plausibility was selected by the researchers, how subjective plausibility ratings relate to suggestion induced beliefs and memories is unclear.

Another line of research has used population surveys to evaluate the degree to which
plausibility is a risk factor for false memories. Rubin and Berntsen (2007) asked participants if it was plausible that people needing psychotherapy could have been victims of childhood sexual abuse that was subsequently forgotten. Most (82.2%) rated plausibility above the mid-point on a four-point scale, which was interpreted as demonstrating a widely held conviction that abuse could be forgotten but still affect behavior, and as a risk for the development of false memories. Pezdek and Blandón-Gitlin (2009) argued that this study measured general plausibility, whereas risk is tied to the plausibility of the event for the person (see Scoboria et al., 2004, for more on general versus personal plausibility). Their participants rated general and personal plausibility for the same scenario. While most (80%) rated the event above the scale mid-point on general plausibility, fewer (25%) rated the event at this level on personal plausibility. They argued this shows that fewer people are at risk for false memories of abuse (although some argue that 1 in 4 remains substantial; see Rubin & Berntsen, 2009). Most relevant here, both studies defined the risk threshold as falling at the mid-point of the plausibility scale, implying that false memories are more likely when plausibility is moderate.

Mazzoni (2007) found that belief ratings about an unlikely event (witnessing a demonic possession) were made the most rapidly when plausibility was rated very low (at the scale floor), while plausibility ratings at any higher point occurred more slowly and at the same speed. She argued that highly implausible events are assessed rapidly and rejected without any need to search memory or decide on occurrence of the event because further processing is cognitively inefficient, a phenomenon which she termed the “plausibility heuristic”. Extending this to the relationship between plausibility and the effects of suggesting events, the risk of false belief formation should be lowest when personal plausibility is very low, with the likelihood of belief increasing when the rating is just slightly above the bottom of the scale. In other words, only an extreme level of implausibility hinders the development of false belief. This is termed here the “event rejection” model.
In this view, plausibility plays an important role during efforts to recall unremembered events. In the absence of memory and belief, plausibility is evaluated to determine whether an event could have occurred (Mazzoni & Kirsch, 2002). Plausibility indicates the degree to which an event is possible, and therefore whether information in memory should be sought. This argument is consistent with other ideas about the role of plausibility in recall (e.g. Blank, 2009; Ghetti & Alexander, 2004; Johnson & Raye, 1981; Johnson, Hashtroudi, & Lindsay, 1993; Lampinen & Odegard, 2006; Scoboria et al., 2004), but in addition predicts which initial level of plausibility hinders the creation of false beliefs.

**The Current Studies**

In two studies, pre-suggestion plausibility was assessed and a false event was later suggested in order to examine the relationship between plausibility and the development of false autobiographical beliefs. False beliefs occur when an objectively false event comes to be viewed as genuinely having occurred, whether or not there is an associated episodic memory. False beliefs are interesting because they are partly distinct from episodic memory (Mazzoni et al., 2010; Clark, Nash, Fincham & Mazzoni, 2012), because they are a precursor to complete false memories (Hyman & Kleinecht, 1999), and because false beliefs alone can influence behavior (Geraerts, Bernstein, Merckelbach, Linders, Rayemaekers, & Loftus, 2008). Beliefs are also more amenable to study because they develop rapidly whereas the creation of false memories requires more extensive procedures (Hyman & Pentland, 1996; Otgaar, Scoboria & Smeets, in press; Scoboria et al., 2007).

The most parsimonious model to relate pre-suggestion plausibility and post-suggestion belief is a linear model by which false belief increases uniformly with plausibility. An alternate model is to assume a critical plausibility threshold above which the likelihood of belief is higher. If a threshold model applies, prior work suggests that the location of the threshold might fall either at the scale midpoint or immediately above the plausibility scale.
floor. In either case, a function with one point of inflection (e.g., quadratic, inverse) should produce a better fit than the linear model. If the event rejection model applies, it was not expected that belief would never change when plausibility was initially low because suggestions can increase plausibility.

Because events such as abuse cannot be suggested for ethical reasons, Study 1 used a relatively benign but still negative event (illness after eating spoiled food), and Study 2 used medical screenings that do not occur to children. In Study 1, participants provided baseline ratings of personal plausibility, belief, and memory. They were later told that they were ill due to spoiled peach yogurt as a child and completed post-suggestion measures. In Study 2, previously published data that used similar methods were reanalyzed.

Method

Participants

The participants were 42 undergraduates (85% female, ages 17-23, $M = 19.19$, $SD = 1.27$) who received course credit.

Measures

The Autobiographical Beliefs and Memory Questionnaire (ABMQ; Scoboria et al., 2004) assesses general plausibility, personal plausibility, autobiographical belief, and memory for events. The plausibility items query whether events could have occurred to others and to the person making the rating; the belief item queries whether events did in fact occur; and the memory item asks about the strength of memory. Each item is rated on an 8-point scale, anchored ‘Not at all plausible’ and ‘Extremely plausible’ for plausibility; ‘Definitely did not happen’ and ‘Definitely happened’ for belief; and ‘No memory for event at all’ and ‘Clear and complete memory for event’ for memory. Participants rated six events taken from prior studies (lost in mall, breaking a toy, choking on object, sick after eating spoiled peach yogurt, sick after eating a pickle, breaking a window with a hand).
Other measures were included to add credibility to the suggestion. These included food preference and food history inventories; a survey of common phobias; and a 24-item Life Events Inventory (LEI; Garry, Manning, Loftus, & Sherman, 1996).

Procedure

All procedures received ethical review. Potential participants completed an on-line screening (ABMQ, FPI, FHI, LEI, phobias). Eligible participants (between age 17 and 23, no memory or belief for childhood dairy related illness, non-extreme preferences for peach yogurt, normal eating characteristics; see Scoboria, Mazzoni & Jarry, 2008, for more on the screen) were invited to participate in what they were told were two separate studies. The data here are from the first session, as the second session could not affect the current results.

Participants attended an experimental session three weeks after the screening ($M = 19.29$ days, $SD = 1.23$). This session was administered by an assistant who was blind to the hypotheses. Participants were randomly assigned to suggestion ($n = 22$) or control groups ($n = 20$). Suggestion participants were told that the purpose of the study was to investigate the use of psychological tests to predict the occurrence of childhood events, and that a profile depicting the likelihood that certain events had occurred during their childhood had been generated from their screening responses. They were told that this was a valid way of detecting past events, and that the study was examining the accuracy of the reports. Participants received a bogus profile that showed how their screening responses allegedly matched those of others who had experienced similar events. Four of ten events, including the target ‘sick on peach yogurt,’ were depicted as above a critical threshold. Participants were told that events above the threshold occurred during their childhood, and that occurrence of those below the threshold was unknown. The high score on the target event was pointed out, and participants were told that the researchers were interested in this event because it had happened to many people, but was remembered by few.
As “evidence” of the event, participants next read a false health report containing population prevalence information. The report resembled a government health alert, and was dated when participants were between three and six years old. The report indicated that many people became ill due to peach yogurt contaminated by the *E.coli.* bacterium, noted the source of the contamination, and described common symptoms.

Participants were asked if they recalled or had heard about the event (none did). Participants then completed a guided visualization and described this imagined narrative aloud. Specifically, participants were asked to close their eyes, concentrate, bring the event into focus, and to imagine being back in the situation. They were encouraged to focus on details of the event, such as feelings, thoughts, what it was like to experience the event, season of the, and who else was present. Participants then imagined a second control event (getting sick on pickles as a child). Finally, participants provided post-suggestion plausibility, belief and memory ratings. Control participants imagined and described the control (pickle) event and completed the post-suggestion measures.

**Results**

The first set of analyses examined whether the manipulations produced changes in belief and memory ratings, as seen in prior studies (see Table 1 for descriptive data). The suggestion did influence belief and memory ratings, relative to control participants and the control event. Univariate ANCOVAs were used to examine change in belief controlling for baseline belief, and ANOVAs were used to examine change in memory (no covariate was needed because all participants had the same baseline memory score). The suggestion led to significantly greater change in both belief and memory for the target event; belief, $F(1, 39) = 4.40, p = .042$, Cohen’s $d = 0.66$; memory, $F(1, 40) = 5.23, p < .05, d = 0.72$. Examination of the control event revealed significant main effects of suggestion for belief, $F(1, 39) = 4.97, p < .05$, and memory, $F(1, 39) = 4.97, p < .05$, which were because belief scores *decreased* in
the control group, and remained unchanged in the suggestion group. These findings confirm that the suggestion produced results similar to those seen in preceding research.

Prior to examining the role of plausibility, we verified that a sufficient number of personal plausibility ratings occurred at the scale floor; 50% of suggestion and 25% of control participants initially rated personal plausibility at this level. Multiple regression analyses tested linear, inverse, quadratic, and cubic models, predicting post-suggestion belief for the target event from pre-suggestion personal plausibility. This analytic approach is appropriate when scales have five or more points. For suggestion participants, an inverse model, $Adj \ R^2 = .37$, $F (1, 19) = 12.44$, $p < .01$, 95% CI for $R^2 .33$ to .41, explained more variance than a linear model, $Adj \ R^2 = .21$, $F (1, 19) = 6.28$, $p < .05$, 95% CI for $R^2 .17$ to .25. As depicted in Fig. 1, post-suggestion belief was lowest when plausibility was at the scale floor and increased sharply at just one point higher. For controls, no model predicted significant variance (linear model, $Adj \ R^2 = .05$, $F (1, 18) = 1.02$, $p > .25$, 95% CI .04 to .06). Models relating plausibility and post-suggestion scores for the control (pickle) event were non-significant (linear, $Adj \ R^2 = .003$, $F (1, 38) = .11$, $p > .25$, 95% CI for $R^2 .00$ to .01). This shows that change in belief was due to the suggestion, that the effect was specific to the target event, and that the effects of the suggestion were predicted by pre-suggestion plausibility.

To quantify the threshold relationship in terms of group differences, average post-suggestion belief was compared for suggestion participants who initially rated plausibility at the scale floor (n =11) and suggestion participants who rated plausibility at any higher point (n =11; see Figure 2). A univariate ANCOVA revealed a main effect of group, $F (1, 18) = 5.02$, $p < .05$. The pre-post effect size was $d = .99$ when plausibility was at the scale floor, and was $d = 2.15$ when plausibility was above baseline.

Another analytic approach used in studies of false memory to reduce the influence of cases with large increases in ratings is to examine proportions of individuals who increase
their ratings. The impact of the suggestion on autobiographical belief was greater when initial plausibility was not at baseline. Of those who rated plausibility above the scale floor, 91% increased their belief rating two or more points, compared to 40% of those who rated plausibility at the scale floor; \( \chi^2 (1) = 6.11, p < .05 \).

**Discussion**

This study provides initial evidence that the relationship between pre-suggestion plausibility and post-suggestion false belief is non-linear, with a threshold for increased belief toward the lower end of the plausibility scale. Consistent with the event rejection hypothesis, initial implausibility was associated with lower likelihood of post-suggestion false belief.

While events rated as highly implausible were less likely to result in false beliefs, belief ratings increased in a minority of individuals who initially rated the event as highly implausible. This result is not unexpected, because previous research (Mazzoni et al., 2001) has found that suggestion can increase plausibility for events initially rated as highly implausible. This opens the road for the creation of beliefs.

While the findings of Study 1 are consistent with the event-rejection model, several issues limit this conclusion. The possibility of other thresholds could not be examined because this study only allowed comparisons of initial low plausibility with all other levels of plausibility. Second, the results may be influenced by variation in baseline belief ratings. Third, whether the suggestion led to changes in plausibility that then impacted belief could not be examined due to sample size. Addressing these issues requires a larger sample.

**Study 2**

To address these concerns, previously published data were reanalyzed (Scoboria et al., 2007; described below). This larger dataset refers to different target events, and similar but not identical suggestive procedures. In addition to examining the hypotheses tested in Study 1, the larger sample permitted examination of increases in plausibility and the potential role
of baseline belief ratings.

Study 2 Method

The studies in Scoboria et al. (2007) examined whether a combination of false prevalence information and a rationale normalizing the forgetting of events would result in the development of false beliefs but not false memories. In two studies, one of two false events (skin sample taken by a nurse, bone density screening) was suggested. The design consisted of three experimental groups. One group received prevalence information and a rationale normalizing forgetting, one received just prevalence information, and the third group was a no-suggestion control. We focus here on the first group (which most resembles Study 1) and the control group.

Participants

The data analyzed here were collected from 116 university students (79% female; ages 17-27, $M = 20.42$, $SD = 2.19$).

Measures

The ABMQ was applied to 10 events representing a range of plausibility. The target event was always presented fourth in the series of events.

Procedure

Participants were contacted via email and invited to participate in a two-part experiment. They completed informed consent and provided baseline ABMQ ratings on-line. Two weeks later they visited the lab, and were randomized to conditions and target events. The study was completed at a computer terminal. Suggestion participants ($n = 74$) viewed five pages of information. The first page introduced the study. The second page provided false prevalence information, which indicated that the target false event (bone density or skin sample) occurred commonly to children in Ontario at a time when the participants was a child. Participants answered questions about the content of the narrative that served as a
manipulation check and encouraged encoding of the information. The third page provided an unrelated medical narrative followed by similar questions; this was included as a filler task. The fourth page provided a rationale regarding forgetting, which indicated that forgetting is common, that even emotional events are frequently forgotten, and that events can often be recalled at a later time. Finally, participants filled out the post-suggestion ABMQ. Control participants \((n = 42)\) did not view the prevalence information or the forgetting rationale.

**Study 2 Results**

There were no significant differences in effects across events and the two original studies, and the analyses reported below collapse across both.

The results of linear, cubic, and inverse models are presented in Table 2 and Figure 3, and a scatterplot is provided in Figure 4. The inverse and cubic models proved equally good fits to the data; inverse: \(\text{Adj } R^2 = .33, F (1, 72) = 32.41, p < .001, 95\% \text{ CI for } R^2 .28 \text{ to } .39\); cubic, and \(\text{Adj } R^2 = .33, F (1, 70) = 11.66, p < .001, 95\% \text{ CI for } R^2 .29 \text{ to } .38\), respectively. Both models predicted more variance than the linear model, \(\text{Adj } R^2 = .20, F (1, 72) = 17.69, p < .001, 95\% \text{ CI for } R^2 .16 \text{ to } .25\). The plots of the inverse and cubic models show that post-suggestion belief was predicted to be lowest when baseline personal plausibility was at the scale floor, and that predicted belief increased sharply when plausibility was one point above the scale floor. The inverse and cubic models showed substantial similarities across the lower half of the plausibility scale.

Comparisons between suggestion participants who rated plausibility at the scale floor and suggestion participants who rated plausibility above the scale floor showed the same pattern as in Study 1. Plausibility above the scale floor led to a larger increase in beliefs as compared to plausibility at the scale floor. The pre-post effect size for belief ratings was \(d = 0.75\) when plausibility was at baseline, and \(d = 1.47\) when plausibility was above baseline.

The primary distinction between the inverse and cubic models lay at the upper end of
the plausibility scale. The cubic model depicts an accelerating positive relationship between plausibility and belief ratings at the scale ceiling, while the inverse model does not. The low number of observations at the upper end of the scale limited the ability to determine which model was the better fit. Because the goals of the analysis were focused on the lower half of the plausibility scale and the two models were consistent in this regard, this interesting difference is noted and examination of the upper end of the plausibility scale is left to future research. Similar analyses on control participants’ data revealed a non-significant linear effect (see Figure 3), \( \text{Adj } R^2 = .052, \ F(1, 40) = 3.24, \ p < .10, \ 95\% \ CI \ for \ R^2 \ .03 \ to \ .08 \). Controls did not show a meaningful relationship between plausibility and post-suggestion belief.

In the suggestion group, individuals initially at the plausibility scale floor were less likely to increase belief scores. At the scale floor, 29% increased their score two or more points on the belief scale, versus 58% of those above the scale floor; \( \chi^2 (1) = 4.68, \ p < .05 \).

**Change in plausibility ratings.** We examined whether the suggestion facilitated increases in plausibility which in turn impacted beliefs. For suggestion participants who started at the plausibility scale floor, those who increased their belief ratings increased plausibility significantly more than those who did not increase belief, \( t(22) = 2.15, \ p < .05, \ d = 1.01; M \text{ change} = 4.43, \ SD = 1.81 \text{ vs. } M \text{ change} = 2.18, \ SD = 2.18 \). Furthermore, 100% of those who started at the scale floor and remained there showed zero change in belief ratings. This shows that increased belief for individuals who started at the plausibility scale floor was related to suggestion induced increases in plausibility.

**Controlling for pre-manipulation belief.** Because pre- and post-suggestion belief ratings are correlated, it is important to examine if initial belief accounts for post-suggestion belief. Additional regression analyses examined the contribution of baseline belief to the inverse and cubic models (see Figure 3). Sequential multiple regression analyses assessed the relative contributions of baseline personal plausibility and baseline belief to post-
manipulation belief scores (see Table 3). Baseline personal plausibility predicted post-suggestion belief when entered first, and adding baseline belief did not contribute additional variance. While baseline belief alone predicted post-suggestion belief, adding plausibility rendered the belief coefficients non-significant. Hence, personal plausibility uniquely predicted post-suggestion belief after accounting for baseline belief.

**Discussion**

This study provides further support for the event-rejection model. As in Study 1, only plausibility ratings made at the scale floor were associated with lower false belief ratings, whereas belief increased sharply just one point above the scale floor and did not change across the rest of the scale. Some individuals who initially rated plausibility at the scale floor increased their belief after the manipulation. This was due to the suggestion causing increases in plausibility, in line with previous results (Mazzoni et al, 2001).

**General Discussion**

These studies clarify one role that personal plausibility plays in the formation of false beliefs for suggested childhood events. This research shows that the relationship between subjective plausibility and post-suggestion autobiographical belief is non-linear. A critical threshold was identified at the lower end of the plausibility scale. Suggested events were less likely to lead to false beliefs only at the lowest level of plausibility. At any other level, including just one point above the scale floor, false beliefs were notably stronger. This is consistent with an event-rejection model, according to which only events judged to be extremely implausible tend not to be processed further. This implies that false beliefs can be easily created for events that are implausible but are not viewed as impossible. The fact that consistent findings were obtained for three distinct target events indicates the findings are robust; there is no obvious reason to assume that the findings should not generalize to the broader set of events, although this remains an empirical question.
Extending these findings to the debate about childhood sexual abuse (CSA), the results show that only when an event is viewed as extremely implausible is the risk for subsequent false beliefs, and presumably memories, reduced. Rubin and Berntsen (2007) and Pezdek and Blandón-Gitlin (2009) used the plausibility scale mid-point as the threshold for risk for false memories about child sexual abuse. These results suggest that the risk of false belief following the suggestion of CSA events is higher than previously discussed.

It is important to emphasize that event rejection is only one part of the picture. Across the studies, more beliefs developed when plausibility started above the scale floor (64%) than at the scale floor (32%). Hence implausibility reduced but did not completely prevent the creation of beliefs. The suggestion sometimes produced increases in plausibility which then permitted the development of belief. This coincides with research showing that suggestions can produce increases in plausibility for implausible events leading to increased belief (Mazzoni et al., 2001). If plausibility increases, the event rejection concept no longer applies.

These results contribute to a growing body of research that illustrates the role of plausibility in the evaluation of belief and memory for events. Plausibility judgments are particularly important when events are not recalled and the occurrence of events is uncertain. Based on prior findings, events that are viewed to be personally impossible are frequently evaluated rapidly (Mazzoni, 2007), which may occur at the same time that the retrieval cue is processed hence producing rapid rejection of retrieval cues (Reder, 1982). For events that are perceived to be improbable, rather than impossible, the interaction between the cue and self-knowledge becomes important. The more consistent a retrieval cue with self-knowledge, the more likely that a memory search will follow (Conway & Pleydell-Pearce, 2000). This is perhaps why small increases in plausibility can lead to false beliefs. Being told that an unremembered event might have happened provides the minimal knowledge that instantiates systematic evaluation, rather than rapid rejection, of events.
These findings are consistent with prior theoretical statements about plausibility. The Nested Model (Scoboria et al., 2004) proposes under most circumstances belief and memory are restricted to be equal to or lower than the subjective plausibility of an event. Hence if plausibility is low belief and memory will necessarily be low. Other work views plausibility as one of a number of memory-editing mechanisms that aid in preventing the endorsement of non-occurring events as remembered (Lampinen & Odegard, 2006). Blank (2009) argues that remembering is subject to reality constraints set by the possibility of events. These findings are consistent with the reality monitoring framework (Johnson & Raye, 1981; Johnson, Raye, Mitchell, & Ankudowich, 2012), which proposes that plausibility prevents events that do not fit within the limits of semantic knowledge from being attributed to the past. These views converge on the position that plausibility limits the circumstances under which suggestions can lead to false beliefs and memories. Our studies add to these theoretical views by measuring plausibility and empirically assessing the influence of plausibility on autobiographical belief. Whereas the reality monitoring framework emphasizes the attribution of mental representations as memories, our findings support a broader conceptualization of truth attributions in which people decide upon the occurrence of past events, whether or not memory is present (Mazzoni & Kirsch, 2002; Scoboria et al., 2004; see Rubin, Schrauf, & Greenberg, 2003, for a related discussion of memory and belief as distinct metacognitive judgments). This work provides further evidence that people can come to believe that false events occurred in the past without an associated sense of episodic memory.

Fuzzy Trace Theory (FTT; Brainerd & Reyna, 2001) provides another explanation as to why minimal plausibility might promote memories. FTT argues that events are encoded at distinct verbatim and gist levels. Extreme implausibility should prevent evaluation of both components. However, minimal plausibility may serve as a context within which belief and a gist sense can both develop. This is consistent findings in the basic memory research that
critical lures associated with semantically related word lists may be driven by familiarity and not verbatim recollection (Plancher et al., 2008). See Brainerd & Reyna (2002) for further discussion of the relationship between FTT and false memories.

These and similar studies show that people vary widely in their subjective assessments of plausibility, and that this assessment influences the evaluation of events. Greater attention should be dedicated to the measurement of plausibility in autobiographical memory research. We note that these studies were not designed to examine the complete relationship between plausibility and post-suggestion belief. The results of Study 2 suggest a second threshold at the upper end of the plausibility scale at where the probability of developing strong false beliefs is particularly high. This finding is tantalizing, but remains tentative due to the small number of ratings at the upper end of the scale. Future studies will likely require oversampling the upper end of the plausibility scale.

These studies focused on false beliefs. How do the findings relate to false memories? First, belief co-occurs with or precedes the development of memories, hence these studies provide evidence of conditions that facilitate the development of memories. Second, as noted above, autobiographical beliefs can guide behavior in the absence of memory.

One criticism of false memory studies is whether the procedures result in the recovery of true memories. While there is no way of guaranteeing this never occurs, methods such as using events that are unlikely due to the specificity of their details (Study 1) or using impossible events (Study 2) reduce such concerns. All lab based research on suggesting false events can be criticized on the grounds that the events may not be as personally impactful as abuse. This issue is unavoidable, due to ethical concerns. Analogue studies such as these are designed to examine the mechanisms by which beliefs and memories occur. Other methods, such as the survey by Rubin and Berntsen (2007) and case studies (e.g., retractors of recovered memories of abuse, Ost, 2003) can establish the parameters of risk for more serious
events. However, if 90% of individuals develop false beliefs under certain circumstances, it is not unreasonable to assume that some proportion of the population is at risk of believing abuse when encountering much more suggestive situations.

These results do not overlap with studies that suggest objectively impossible events. Studies such as Braun, Ellis, and Loftus (2002) and Braun-LaTour, LaTour, Pickrell, and Loftus (2004), show that false memories result when events are objectively impossible. But objective impossibility does not necessarily correspond to subjective implausibility. Because personal plausibility is not measured in such studies, it is not known if the events were viewed as personally implausible prior to the suggestion. Autobiographical belief and memory are subjective phenomena, so it is the subjective appraisal of plausibility that matters most in decisions about belief and memory for objectively true and objectively false events.

To conclude, these results show that subjective implausibility is often (but not always) sufficient to lead to rejection of false events. Above this level false beliefs can be implanted, with minimal differences among the remaining levels of the scale. The results also show that initial extreme implausibility can be swayed by new information. Hence, the question remains: once plausibility increases, to what extent does the risk of false beliefs and memories at a later time also increase?
Author Note

Correspondence to: Alan Scoboria, Department of Psychology, University of Windsor, 401 Sunset, Windsor, ON, Canada, N9B3P4; scoboria@uwindsor.ca. This research was supported by a Discovery Grant to the first author by the Natural Science and Engineering Research Council of Canada.
References


Rubin, D., & Berntsen, D. (2009). Most people who think that they are likely to enter psychotherapy also think it is plausible that they could have forgotten their own memories of childhood sexual abuse. *Applied Cognitive Psychology, 23*, 170-173. doi:10.1002/acp.1462


IMPLAUSIBILITY INHIBITS FALSE BELIEF…

Applied Cognitive Psychology, 20, 1049–1064. doi:10.1002/acp.1240


doi:10.1002/acp.1062


Table 1

*Study 1, Means and Standard Deviations for Plausibility, Belief and Memory by Event (Target, Control) and Time (Pre-Suggestion, Post-Suggestion)*

<table>
<thead>
<tr>
<th>Event</th>
<th>Suggestion</th>
<th>Pre-suggestion</th>
<th>Post-suggestion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target event</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Personal plausibility</td>
<td>Belief</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>2.50</td>
<td>1.55</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>2.15</td>
<td>1.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>2.80</td>
<td>1.40</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>1.77</td>
<td>0.68</td>
</tr>
<tr>
<td>Control event</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Personal plausibility</td>
<td>Belief</td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>3.14</td>
<td>2.09</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>2.36</td>
<td>1.63</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>4.74</td>
<td>3.05</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>2.33</td>
<td>2.46</td>
</tr>
</tbody>
</table>
### Study 2, Regression Analyses Predicting Post-suggestion Belief from Baseline Personal Plausibility

<table>
<thead>
<tr>
<th>Equation</th>
<th>Adj $R^2$</th>
<th>$F$</th>
<th>df1</th>
<th>df2</th>
<th>Sig.</th>
<th>Constant</th>
<th>$b_1$</th>
<th>$b_2$</th>
<th>$b_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>.19</td>
<td>23.20</td>
<td>1</td>
<td>72</td>
<td>&lt;.01</td>
<td>2.15</td>
<td>.45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quadratic</td>
<td>.24</td>
<td>12.87</td>
<td>2</td>
<td>71</td>
<td>&lt;.01</td>
<td>1.13</td>
<td>1.20</td>
<td>-.09</td>
<td></td>
</tr>
<tr>
<td>Cubic</td>
<td>.33</td>
<td>12.81</td>
<td>3</td>
<td>70</td>
<td>&lt;.01</td>
<td>-.64</td>
<td>3.18</td>
<td>-.61</td>
<td>.04</td>
</tr>
<tr>
<td>Inverse</td>
<td>.33</td>
<td>37.52</td>
<td>1</td>
<td>72</td>
<td>&lt;.01</td>
<td>5.64</td>
<td>-3.79</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3

*Study 2, Sequential Regression Analyses Predicting Post-suggestion Belief from Baseline Personal Plausibility and Belief*

<table>
<thead>
<tr>
<th>Step</th>
<th>$R^2$</th>
<th>B</th>
<th>SE</th>
<th>Beta</th>
<th>t</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Model 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>.19</td>
<td>2.14</td>
<td>.38</td>
<td>5.68</td>
<td>&lt; .01</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Baseline Personal plausibility</td>
<td>.46</td>
<td>.10</td>
<td>.49</td>
<td>4.82</td>
<td>&lt; .01</td>
</tr>
<tr>
<td>2</td>
<td>.27</td>
<td>1.92</td>
<td>.40</td>
<td>4.78</td>
<td>&lt; .01</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Baseline Personal plausibility</td>
<td>.34</td>
<td>.12</td>
<td>.37</td>
<td>2.79</td>
<td>&lt; .01</td>
</tr>
<tr>
<td></td>
<td>Baseline Belief</td>
<td>.35</td>
<td>.23</td>
<td>.20</td>
<td>1.51</td>
<td>.14</td>
</tr>
<tr>
<td></td>
<td><strong>Model 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>.25</td>
<td>2.31</td>
<td>.39</td>
<td>5.87</td>
<td>&lt; .01</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Baseline Belief</td>
<td>.75</td>
<td>.19</td>
<td>.43</td>
<td>4.07</td>
<td>&lt; .01</td>
</tr>
<tr>
<td>2</td>
<td>.27</td>
<td>1.92</td>
<td>.40</td>
<td>4.78</td>
<td>&lt; .01</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Baseline Belief</td>
<td>.35</td>
<td>.23</td>
<td>.20</td>
<td>1.51</td>
<td>.14</td>
</tr>
<tr>
<td></td>
<td>Baseline Personal plausibility</td>
<td>.34</td>
<td>.12</td>
<td>.37</td>
<td>2.79</td>
<td>&lt; .01</td>
</tr>
</tbody>
</table>
Figure 1. Study 1, Plot of baseline personal plausibility predicting post-suggestion belief ratings for participants receiving the suggestion regarding illness due to peach yogurt (inverse model), control participants not receiving the suggestion (linear model), and for responses to the control event (linear model).
Figure 2. Study 1, Average post-suggestion belief and memory ratings with baseline plausibility at the scale floor versus above the scale floor, for individuals who received the suggestion (N = 22; 11 “at scale floor” and 11 “above scale floor”).
Figure 3. Study 2, Plot of baseline personal plausibility predicting post-suggestion belief with and without baseline belief ratings, for inverse and cubic regression models. Also depicted are predicted ratings for control participants who received no suggestion. Respective Ns for each level of the plausibility scale, Suggestion group: 23, 13, 12, 9, 2, 4, 4, 7; Controls: 18, 11, 7, 7, 5, 5, 2, 3.
Figure 4. Study 2, Scatterplot of baseline personal plausibility and belief ratings in the suggestion group. Size of circles indicates the number of data points at each location. The interpolation line shows the average belief rating at each level of plausibility.