To appear in the *Journal of Applied Research in Memory and Cognition 2022*

Hindsight Bias and COVID-19: Hindsight wasn't 20/20 in 2020

Megan E. Giroux[1], Daniel G. Derksen[1], Patricia I. Coburn[2], Daniel M. Bernstein[2]

[1]Simon Fraser University, Canada

[2]Kwantlen Polytechnic University, Canada

This work was supported by funding from Kwantlen Polytechnic University, the Canada Research Chairs Program (950-232078), the Social Sciences and Humanities Research Council of Canada (435-2015-0721), and a Vanier Canadian Graduate Scholarship. We would like to thank Shayna Rusticus, Hartmut Blank, Brent Roberts, and Edgar Erdfelder for their support with the analyses. We would also like to thank Iaren Rai for help with data collection.

Our data and materials are available at https://osf.io/dmfjy/?view_only=1361d36a3ef34397b4d1074432962431

Address correspondence to: Megan E. Giroux, Department of Psychology, RCB 5246, 8888 University Dr, Burnaby, BC V5A 1S6. Email: megiroux@sfu.ca

Total Word Count: 9,497 words

Introduction & Discussion Word Count: 2,887 words

Abstract

Hindsight bias occurs when outcome information distorts people's memories of past beliefs or exaggerates perceptions of outcomes' foreseeability or inevitability. We investigated whether community and university participants in Canada and the U.S. exhibited hindsight bias for COVID-19. In Experiment 1 (N = 175), participants made original judgments about COVID-19 outcomes. Two months later, participants learned outcome information and recalled their original judgments (memory design). They also rated the foreseeability and inevitability of COVID-19. In Experiment 2 (N = 754), we used a hypothetical design. Participants learned outcome information before estimating how naïve peers would have responded two months earlier. Participants exhibited hindsight bias in memory and hypothetical estimations. However, they rated COVID-19 as unforeseeable and avoidable and generally did not exhibit differences in foreseeability and inevitability ratings across the two timepoints. Thus, hindsight bias for COVID-19 differs across memory distortions, foreseeability, and inevitability and extends to hypothetical judgments.

Keywords: hindsight bias; COVID-19

General Audience Summary

Hindsight can cloud the past by biasing people's beliefs about what was known prior to an outcome. Hindsight can also bias people's beliefs about the foreseeability or inevitability of an outcome. We explored hindsight bias for COVID-19 in two experiments. In both experiments, Canadian and U.S. participants made foresight judgments about several COVID-19 outcomes (e.g., case rates in various countries). Participants also judged how foreseeable and inevitable COVID-19 was. Two months later, we recruited two groups of participants: (a) a sample that previously completed the foresight judgments; and (b) a new sample. Both groups received outcome information. In Experiment 1, group (a) had to ignore their current outcome knowledge and recall their original judgments for the COVID-19 outcomes. They also rated their current perceptions of foreseeability and inevitability. In Experiment 2, group (b) had to ignore their current outcome knowledge and estimate how a naïve peer would have responded to the same questions two months prior.

We observed hindsight bias: Outcome information biased people's judgments about what they, or someone else, previously believed about various COVID-19 outcomes (e.g., death rates). Additionally, Canadian and U.S. residents perceived COVID-19 as unforeseeable ("I never could have seen this coming") and avoidable ("this didn't have to happen"). However, they generally did not demonstrate differences in foreseeability and inevitability ratings across the two timepoints. This work has applied significance for the field because it is one of the few studies to investigate hindsight bias for a real-world, evolving event that is negative and self-relevant for everyone. Our results reveal the importance of studying hindsight bias for COVID-19 through various measures to determine how it affects other types of judgments (e.g., evaluations of public

health authorities, vaccination tendencies, etc.); this can inform public health practices aimed at mitigating COVID-19 and future public health crises.

Hindsight Bias and COVID-19: Hindsight wasn't 20/20 in 2020

Think back to January 2020. The first few cases of a novel coronavirus had just been reported in Wuhan, China. With little information about the transmissibility and fatality of the virus, many of us continued to travel, hold large gatherings, and meet in public places without masks. When COVID-19 first emerged, many of us thought that it would settle within a matter of months. However, now that we have been living through this pandemic and its many variants for nearly two years, we may think back to our initial beliefs about when businesses would reopen or how many people would be infected by COVID-19 and misremember these beliefs as being closer to the outcomes than they really were. In other words, as time has passed, our ability to accurately recall our previous beliefs about the course of the pandemic may have become distorted by hindsight.

Hindsight bias occurs when outcome information biases people's judgments about what was known in the past (Christensen-Szalanski & Willham, 1991; Fischhoff, 1975; Guilbault, Bryant, Brockway, & Posavac, 2004). Two experimental paradigms are typically used to study hindsight bias across various domains (Pohl, 2007; Roese & Vohs, 2012). In the within-subject *memory design*, participants judge event outcomes (e.g., "how many Canadians will have been infected by COVID-19 as of March 2021?"). Then, participants learn the outcomes to half the events (experimental items) and do not learn the outcomes to the remaining events (control items). Finally, participants try to ignore their knowledge of the outcome information to recall their original judgments. Participants' judgments are generally closer to the outcomes for experimental items than control items. Hindsight bias occurs when outcome information influences people's memory such that in hindsight, they remember their foresight responses being closer to the outcome information than their foresight responses actually were.

In the between-subject *hypothetical design*, there is a foresight condition and a hindsight condition. In the foresight condition, participants judge event outcomes without learning the outcomes. In the hindsight condition, participants learn and then try to ignore event outcomes while indicating how they would have responded had they not learned the outcomes. Participants in the hindsight condition generally provide judgments closer to the outcomes than participants in the foresight condition. Hindsight bias occurs when outcome information influences people's hindsight judgments, such that they indicate that others' foresight responses would be closer to the outcomes than others' foresight responses actually were.

Hindsight Bias Components

While outcome information can bias people's recall of what was known in the past, it can also bias people's beliefs about the foreseeability and inevitability of outcomes (Blank, Nestler, von Collani, & Ficher, 2008; Nestler, Blank, & Egloff, 2010). Blank et al. (2008) delineate three distinct hindsight bias components. Different variables affect each component, and each component requires different judgments. The first component, *memory distortion*, occurs when outcome information affects people's recollections about what they previously knew or believed. For example, when the pandemic started, an individual named Lisa might have estimated that 20,000 Americans would die from COVID-19. Now that over 900,000 Americans have died, Lisa might say "I always thought that least 100,000 Americans would die." Thus, outcome information biased her recollection of what she believed in the past. General memory processes such as depth of encoding and delays between encoding and retrieval influence the memory distortion component (Erdfelder, Brandt, & Bröder, 2007; Groß & Bayen, 2015).

The second component, *foreseeability*, is a person's subjective beliefs about whether they could have predicted an outcome. People with outcome information often rate outcomes as more

predictable compared to people without outcome information. For example, some may claim that they "knew all along" that COVID-19 would turn into a global pandemic. However, this may depend on the outcome's valence and self-relevance (Nestler et al., 2010). If an outcome is negative and self-relevant (e.g., getting fired from a job), people may claim that they could *not* have foreseen the outcome. Absolving themselves of personal responsibility for the outcome (e.g., lack of performance) may ease the disappointment associated with the outcome (see *defensive processing theory;* Louie, 1999; Pezzo & Pezzo, 2007). Accordingly, they convince themselves that they could not have foreseen the outcome, and thus, there is nothing they could have done differently. Metacognitive processes are thought to primarily influence the foreseeability component, though people's ability to make causal attributions between an event and its antecedents and motivational mechanisms likely play a role as well (Blank et al., 2008; Müller & Stahlberg, 2007).

The third component, *inevitability*, is a person's perceptions about the objective likelihood that an outcome had to occur. Inevitability differs from foreseeability because judging the likelihood of an outcome occurring (i.e., inevitability; "it had to happen") differs from judging one's own ability to have predicted that outcome (i.e., foreseeability; "I knew it was going to happen"). People with outcome information often rate outcomes as more probable compared to people without outcome information. Inevitability is mainly influenced by *sense-making*, or people's ability to draw causal connections between the outcome and its preceding events (Pezzo, 2003). However, motivational mechanisms may override more basic cognitive sense-making processes when an outcome is disappointing and self-relevant (Nestler et al., 2010; Pezzo & Pezzo, 2007). In fact, people may convince themselves that a negative outcome was inevitable, and therefore, ultimately out of their control. This helps individuals cope with

disappointing outcomes: Instead of dwelling on how an outcome could have differed, individuals accept the outcome and move on (see *retroactive pessimism;* Tykocinski, 2001; Tykocinski & Steinberg, 2005). Thus, sense-making, and to some extent, motivational mechanisms influence the inevitability component (Nestler et al., 2010; Wasserman, Lempert, & Hastie, 1991).

Previous work shows that memory distortion, foreseeability, and inevitability can differ in magnitude and direction and are largely uncorrelated (Blank et al., 2008). For example, people may perceive an outcome as inevitable (e.g., given international travel and trade, a global pandemic was inevitable); yet they do not believe that they could have personally foreseen that outcome. These components can also occur in tandem. Foreseeability and inevitability may be especially likely to co-occur because the causal reasoning processes that lead one to judge the objective likelihood of an event occurring (i.e., inevitability) often influence one's beliefs about whether they could have personally predicted an outcome (i.e., foreseeability; Nestler et al., 2010).

The Current Study

We tested how outcome information affects participants' judgments about COVID-19. It is worth studying hindsight bias for COVID-19 because it is a real, ongoing event that has world-wide implications. Additionally, this pandemic is a defining period that will have lasting effects on many aspects of our lives. Thus, studying this event can reveal important information about how people's beliefs and memories of COVID-19 can change across time, as this event evolves. Studying the public's response to COVID-19 may also have implications for how we learn from and respond to public health crises in the future.

Our primary goal was to determine whether participants exhibited hindsight bias across the memory distortion, foreseeability, and inevitability components. Given that COVID-19 is self-relevant and negative for everyone, we believed that differences might emerge across the three components. This is because the foreseeability and inevitability components are especially sensitive to manipulations of self-relevance (Nestler et al., 2010). Comparatively, general memory processes (e.g., delays) govern memory distortions, and thus, the memory distortion component should be less affected by manipulations of self-relevance. We employed a typical memory design to test this.

Our second goal was to test whether the same data pattern would emerge across these components in a hypothetical design. Our final goal was to examine hindsight bias in an ongoing naturalistic event affecting a wide and varied demographic (i.e., Canada and the U.S.). We conducted this study while COVID-19 was ongoing. The number of cases and deaths continued to rise throughout data collection, and more outcome information (e.g., death rates) emerged over time. Though the pandemic's occurrence remained the same throughout data collection, people's memories and perceptions of foreseeability and inevitability may have changed across time as the COVID-19 pandemic continued to evolve. We also believed that there may be differences in individuals' tendencies to exhibit hindsight bias between Canada and the U.S. given the vastly different federal government responses throughout the early stages of the pandemic.

Experiment 1

In Experiment 1, we conducted a within-subject memory design to study hindsight bias for COVID-19 across the three components. We pre-registered our study on Open Science

Framework (OSF) prior to data collection (see https://osf.io/dmfjy/?view_only=1361d36a3ef34397b4d1074432962431).

Method

Design. We conducted a 2 (timepoint: foresight; hindsight) x 2 (country: Canada; U.S.) mixed design with timepoint as the within-subject factor. We refer to our first and second data collection timepoints as the "foresight" and "hindsight" conditions respectively for clarity. However, it should be noted that all judgments about the foreseeability and inevitability of COVID-19 were made after the pandemic began, and thus, there is no true "foresight" condition for foreseeability and inevitability judgments. Rather, compared to Time 1 foreseeability and inevitability judgments, Time 2 judgments were made at a point when more information was available and could conceivably influence participants' perceptions. Participants completed the hindsight condition approximately 8-10 weeks after they completed the foresight condition.

Participants. We recruited participants from Canada and the U.S. through Amazon Mechanical Turk (MTurk), Reddit, and an undergraduate research pool at a Canadian University. across two timepoints (see Supplementary Materials for demographic information). Participants recruited for the second timepoint were sampled from individuals who completed the first timepoint. An a priori power analysis conducted in G*Power 3.1 (Faul, Erdfelder, Buchner, & Lang, 2009) revealed that we needed 52 participants from each country to detect small-to-medium differences between foresight and hindsight using a paired samples t-test within each country (two-tailed; d_z = 0.4; alpha = .05; power = 0.8). Therefore, we required 104 participants in total.

Foresight condition. We recruited 544 participants from MTurk (N = 407), Reddit (N = 18), and a Canadian undergraduate research pool (N = 119) for the foresight condition. On MTurk, participants needed at least a 95% approval rating with 500 or more HITs approved to participate in this study. We excluded 20 participants who withdrew early, 28 participants with missing responses to more than 50% of the questions in any one component category of questions (memory distortion, foreseeability, inevitability), 14 participants who completed the foresight condition twice, and 9 participants who were not Canadian or U.S. residents. Thus, 473 participants (217 Canadian residents and 256 U.S. residents) comprised our foresight condition.

Hindsight condition. Of the participants who completed the foresight condition, we contacted 444 to complete the hindsight condition based on whether they provided their contact information for follow-up (N = 316 from MTurk; N = 9 from Reddit; N = 119 from a Canadian undergraduate research pool). Of those we contacted, 212 completed the hindsight condition (N = 198 from MTurk; N = 2 from Reddit; N = 12 from undergraduate research pool). After excluding participants who withdrew early, failed to respond to more than 50% of questions in any one component category, or provided the wrong code at follow-up preventing us from matching their foresight and hindsight data, we had complete data sets for 185 participants (86 Canadian residents and 99 U.S. residents).

Procedure. Figure 1 provides an overview of the procedure and Table 1 includes a glossary of important terms. We recruited the first group of participants (i.e., foresight participants) between June 17, 2020, and September 14, 2020 from Canada and the U.S. Participants completed the study on Qualtrics. Participants answered several questions regarding the foreseeability of COVID-19 (see Table 2). We adapted these questions from Blank et al.'s (2008) foreseeability scale. Given that we asked the foreseeability questions in both foresight and

hindsight, we didn't include Item 4 from Blank et al.'s original foreseeability scale (i.e., "In the first questioning session, I made a good prediction of the election outcome") because there had not been a "first" testing session at Time 1 (i.e., foresight). We also didn't include Item 3 from Blank et al.'s original foreseeability scale (i.e., "Recently, I was no longer that certain about the election outcome") because we couldn't be sure whether we would be assessing if participants were once certain about COVID-19 outcomes but were no longer or if they were ever certain. We also included a summary item (i.e., "how predictable was the COVID-19 pandemic?"), similar to the summary item included in Nestler et al. (2010).

Participants also answered several questions regarding the inevitability of COVID-19 (see Table 2), which we adapted from Blank et al.'s (2008) inevitability scale. We wanted to keep the task short and streamlined, and therefore, chose to omit some items. Our decision to omit items was based on whether items could easily be applied to the pandemic. For example, we didn't include Item 4 from Blank et al.'s original inevitability scale (i.e., "Because of the traditional voting behavior in Saxonia, the election outcome was already fixed") because we didn't think there was an appropriate adaptation of this question for our study. We also decided to keep Item 2 from Blank et al.'s (2008) original inevitability scale (i.e., "Nothing could have influenced the election outcome") because we were interested in the question of whether participants believed there were factors that could have influenced the pandemic. We also included a summary item (i.e., "how inevitable was the COVID-19 pandemic?"), similar to the summary item included in Nestler et al. (2010).

Participants rated the foreseeability and inevitability of the COVID-19 pandemic's occurrence. Given that COVID-19 had already been declared a pandemic when we started collecting data, we decided that using past tense would be the best way to measure participants'

perceptions of the foreseeability and inevitability of the pandemic. However, we believed that as COVID-19 continued to evolve and outcomes (e.g., case and death rates) changed across time, this could affect people's perceptions of the foreseeability and inevitability of the pandemic.

Participants were also asked to estimate how many people would be infected and the number of deaths that would occur both two months and one year from their participation. They provided separate estimates for the six countries listed: Canada, the U.S., China, Italy, South Korea, and Sweden. In total, they made 24 judgments about case and death rates (see https://osf.io/dmfjy/?view_only=1361d36a3ef34397b4d1074432962431 for materials).[1] They also answered several questions about their government's response to the pandemic as well as their own compliance with the restrictions and guidelines in their country. These additional questions are outside the scope of our specific hypotheses and will not be discussed further. Finally, participants completed demographic information and entered their email addresses if they wanted to be contacted for a follow-up study (the hindsight condition).

Approximately 8-10 weeks after completing the foresight condition (between August 21, 2020 and November 14, 2020), those who left email addresses were contacted to complete the hindsight condition. In hindsight, participants first rated the foreseeability and inevitability of the pandemic by answering the same questions from the first timepoint. They were not asked to recall their original foreseeability or inevitability ratings, but rather, to indicate their current perceptions of foreseeability and inevitability. Then, participants learned the current case and

[1] There were 4 additional items that we initially planned to include in our set of memory distortion items regarding when schools and non-essential stores would re-open, the amount of money participants' Government would pay in relief benefits, and the extent to which participants' country would return to normal post-pandemic. However, due to the way we constructed these questions, the outcome information was either unknown or impossible to quantify, and thus, we did not analyze these items with the rest of the memory distortion data.

death counts in each of the six countries.[2] They were instructed to ignore their current knowledge and recall their foresight responses for these items (i.e., their previous responses from the first timepoint).

We updated the case and death counts for each country bi-weekly based on the numbers reported on the Johns Hopkins Coronavirus Resource Center page (https://coronavirus.jhu.edu/map.html). We presented the outcome information in the following format: "As of [date], there have been [number of cases] reported cases of COVID-19 in [country]. Two months ago, you were asked: 'Please estimate the total number of people who will have been infected by COVID-19 from when it began to two months from now in [country].' Please ignore your knowledge of the current number of cases and type the number of cases you originally estimated there would be in [country] at this point in time. If you can't remember what you originally put, please type your best guess."

Results and Discussion

Data Exclusions. We only included responses for which participants provided exact estimates. Therefore, we excluded general estimates (e.g., "thousands," "7 millions," "9600+," etc.) as well as ambiguous responses where it was not clear what number the participant intended (e.g., "1,20,000"). We also removed responses that were greater than 3 standard deviations above the mean. Of the 8,880 independent responses in Experiment 1, there were 22 (0.25%) excluded and missing (i.e., blank) responses and 80 extreme values (0.90%).

[2] Our pre-registration specified that participants would receive outcome information for only half the memory items and that the other items would be control items. Due to a programming error, participants received outcome information for all memory items. Thus, there were no control items.

Memory Distortions. We first examined participants' foresight and hindsight judgments on the items regarding case and death counts in various countries. To determine whether individuals exhibited hindsight bias, we computed the following for participants' foresight response: |Original Judgment - Correct Judgment|; and hindsight response: |Recalled Judgment - Correct Judgment|.[3] We standardized all item scales (i.e., the case and death rate estimates for each country) by dividing the difference between each participant's judgment and the correct judgment by the standard deviation of all participants' judgments for each item (Pohl's index; Pohl, 2007). Smaller numbers indicate that participants' judgments were closer to the correct judgments (i.e., more hindsight bias).

We excluded ten participants whose standardized foresight memory distortion score exceeded 3 standard deviations above the mean; thus, we analyzed data from 175 participants. We then conducted a 2 (timepoint: foresight; hindsight) x 2 (country: Canada; U.S.) mixed ANOVA with timepoint as the within-subject factor and mean memory distortion as the dependent variable. There was a main effect of timepoint, $F(1, 173) = 62.34$, $p < .001$, $\eta_p^2 = 0.27$. Participants' recalled judgments were closer to the correct judgments in hindsight (i.e., when they had outcome information) than in foresight, $p < .001$, 95% CI [0.19, 0.32], $d_z = 0.60$ (see Table 3).

Foreseeability Impressions. To determine whether our foreseeability scale was unidimensional, we conducted an exploratory factor analysis enforcing a two-factor solution with varimax rotation (Blank et al., 2015). This analysis revealed that Items 1, 3, and 5 on our

[3] This differed from what we pre-registered because, due to a programming error, we did not have separate control and experimental items. Thus, we compared the standardized differences between participants' judgments and the correct judgments for foresight and hindsight to determine whether participants' judgments were systematically closer to the correct judgments in hindsight than in foresight.

foreseeability scale consistently loaded strongly onto Factor 1 while Item 4 loaded more strongly onto Factor 2 across the foresight and hindsight conditions. Therefore, we omitted Item 4 from the foreseeability scale. Alternatively, while Item 2 loaded more strongly onto Factor 1 in the foresight condition, this item loaded more strongly onto Factor 2 in the hindsight condition (see Supplementary Materials for factor analysis output). Additionally, as we report in Experiment 2 below, we found that Item 2 more strongly loaded onto Factor 1 in both the foresight and hindsight conditions. Overall, Item 2 tended to load more strongly onto Factor 1 across conditions and experiments and including this item increased the reliability of our foreseeability scale (see Supplementary Materials for reliability statistics). Thus, in an effort to maintain consistency in our foreseeability measure across Experiments 1 and 2, we included Item 2 on our foreseeability scale. We calculated the mean of the four foreseeability items, which yielded raw scores ranging from 1-8. Then, we subtracted the scale midpoint (4.5) from all raw scores. This yielded values ranging from -3.5 to +3.5, with 0 as the new scale midpoint. Thus, negative values denoted that the pandemic outcomes were unforeseeable and positive values denoted that the outcomes were foreseeable.

We conducted one-sample t-tests using 0 (i.e., the scale midpoint) as the test value to determine whether participants' foreseeability ratings significantly differed from this midpoint. In Canada, participants' foreseeability ratings were significantly below 0, $M = -0.98$, $t(78) = 6.94$, $p < .001$, 95% CI [-1.26, -0.70], $d = 0.78$. Similarly, in the U.S., participants' foreseeability ratings were significantly below 0, $M = -1.09$, $t(95) = 6.66$, $p < .001$, 95% CI [-1.41, -0.76], $d = 0.68$. Thus, Canadian and U.S. participants' foreseeability ratings did not reflect the assertion that they "knew all along" how COVID-19 would turn out. Rather, their responses suggested that they perceived COVID-19 as relatively unforeseeable.

We also compared participants' foresight and hindsight foreseeability impressions to determine whether there was a shift in foreseeability impressions as a function of more outcome information becoming available across time. A 2 (timepoint: foresight; hindsight) x 2 (country: Canada; U.S.) mixed ANOVA revealed a significant difference in participants' foreseeability ratings across the foresight (M = -1.21) and hindsight (M = -1.04) timepoints, $F(1, 173) = 4.26$, $p = .040$, $\eta_p^2 = 0.02$ (see Table 3). Participants rated COVID-19 outcomes as significantly more foreseeable in hindsight than in foresight. There was no interaction between country and timepoint.

Inevitability Impressions. To explore the unidimensionality of our inevitability scale, we once again conducted an exploratory factor analysis enforcing a two-factor solution with varimax rotation (Blank et al., 2015). This analysis revealed that Items 2, 3, and 4 on our inevitability scale consistently loaded strongly onto Factor 1 while Item 1 loaded more strongly onto Factor 2 across the foresight and hindsight conditions (see Supplementary Materials for factor analysis output). Therefore, we omitted Item 1 from the inevitability scale.

We calculated the mean of the three inevitability items which yielded raw scores ranging from 1-8. Then, we subtracted the scale midpoint (4.5) from all raw scores. Values ranged from -3.5 to +3.5 (scale midpoint = 0), with negative values denoting the pandemic outcomes were avoidable and positive values denoting the pandemic outcomes were inevitable.

We conducted one-sample t-tests to determine whether participants' inevitability ratings differed significantly from the scale midpoint (i.e., 0). Canadian participants' inevitability ratings did not differ significantly from 0, M = -0.24, $t(78) = 1.69$, $p = .095$, 95% CI [-0.52, 0.04], d = 0.19. Thus, Canadian participants did not perceive COVID-19 as either inevitable or avoidable. Alternatively, U.S. participants' inevitability ratings were significantly below 0, M = -1.30, t(95)

= 8.15, $p < .001$, 95% CI [-1.62, -0.98], $d = .83$. Thus, U.S. participants perceived COVID-19 to be relatively avoidable.

We also analyzed whether participants' foresight inevitability impressions differed from their hindsight inevitability impressions. A 2 (timepoint: foresight; hindsight) x 2 (country: Canada; U.S.) mixed ANOVA revealed that participants' hindsight inevitability ratings did not differ from their foresight inevitability ratings in either country (see Table 3). Thus, having more outcome information did not change their perceptions of the pandemic's inevitability.

Correlations between Components. Finally, we conducted correlational analyses to determine if there were associations between memory distortion, foreseeability, and inevitability. We first assessed the reliability of each of our component measures. The memory distortion scale consisted of the standardized difference between participants' hindsight and foresight judgments: |Recalled Judgment – Correct Judgment| - |Original Judgment – Correct Judgment|. This scale yielded a Cronbach's α of .73. The foreseeability scale consisted of the difference between participants' foresight and hindsight ratings (1-8) on each of the four foreseeability items and yielded a Cronbach's α of .60. The inevitability scale consisted of the difference between participants' foresight and hindsight ratings (1-8) on each of the three inevitability items and yielded a Cronbach's α of .48.

Memory distortions did not correlate with foreseeability, $r(173) = -.14$, $p = .060$ or inevitability, $r(173) = .05$, $p = .504$. Additionally, there was a positive correlation between foreseeability and inevitability, $r(173) = .18$, $p = .020$.

A limitation of Experiment 1 is that we did not ask participants to recall their original foreseeability and inevitability judgments; thus, we cannot make any claims about whether they

exhibited a traditional "hindsight bias" for these components. We chose to measure foreseeability and inevitability in this manner so that these components were distinct from the memory distortion component. However, we acknowledge that this is not a common way to measure a hindsight effect because it doesn't require participants to suppress their outcome knowledge. We addressed this limitation in Experiment 2 by using a hypothetical design in which we asked participants to ignore their outcome knowledge when judging the foreseeability and inevitability of the COVID-19 pandemic.

Experiment 2

In Experiment 2, we replicated much of our Experiment 1 procedure to explore whether the same data pattern would emerge across the three components using a hypothetical design. Given that memory processes do not govern the foreseeability and inevitability components, a memory design should not be required to observe these components. However, memory distortions cannot be observed in hypothetical designs. Memory distortions occur when outcome information interferes with people's ability to recall their original judgments. Therefore, this multi-component view does not account for hindsight bias that occurs within hypothetical designs—that is, when people with outcome information make judgments closer to the outcome than people without outcome information. Though this hypothetical hindsight judgment lacks memory distortion features (e.g., interference effects), we argue that this judgment differs from foreseeability and inevitability judgments. Thus, we maintained the three categories across both designs, but refer to memory distortions as "hypothetical magnitude estimates" in the hypothetical design. Once again, our pre-registration, materials, and data are available on OSF (https://osf.io/dmfjy/?view_only=1361d36a3ef34397b4d1074432962431).

Method

Design. To measure hindsight bias using the hypothetical design, we conducted a 2 (group: foresight; hindsight) x 2 (country: Canada; U.S.) between-subject design. Once again, we refer to our first and second data collection timepoints as the "foresight" and "hindsight" conditions respectively for clarity but acknowledge that all judgments about the foreseeability and inevitability of COVID-19 were made after the pandemic began. Participants in the hindsight condition completed the study approximately 8-10 weeks after participants in the foresight condition. The foresight condition was comprised of participants who completed the foresight condition in the memory design (Experiment 1).

Participants. As in Experiment 1, we recruited participants across two timepoints (see Supplementary Materials for demographic information). However, unlike Experiment 1, participants in the hindsight condition (i.e., recruited at the second timepoint) were an independent sample of individuals who *did not* complete the foresight condition (i.e., recruited at the first timepoint). We conducted an a priori power analysis in G*Power 3.1 to determine the sample size required to detect a small-to-medium effect. When we originally planned to collect data from four countries, we decided that the sample size required to detect differences in an independent samples t-test within each country was too conservative given that we did not have specific hypotheses regarding the differences between countries. However, after opting to collect data from Canadian and U.S. participants only, we chose the more conservative power analysis that would allow us to detect small-to-medium effects in an independent samples t-test (see pre-registration). Thus, we required 100 participants per timepoint within each country for a total of 400 participants (two-tailed; $d = 0.4$; alpha = .05; power = 0.80).

Foresight Group (First Timepoint). The foresight condition was comprised of the 473 participants who completed the foresight condition at the first timepoint in the memory design.

All participants with complete data sets were included. There were 217 Canadian participants and 256 U.S. participants.

Hindsight Group (Second Timepoint). We recruited 427 new participants between September 28, 2020 and November 24, 2020 for our hindsight group in our hypothetical design. Participants were recruited through MTurk (N = 270) and a Canadian undergraduate research pool (N = 157). Once again, MTurk participants needed at least a 95% approval rating with 500 or more HITs approved to participate in this study. We excluded participants who withdrew early (N = 100), with more than 50% of data missing in any one component category of responses (N = 13), who completed the study at the first timepoint (N = 6), or who did not indicate their country of residence (N = 4). Thus, we included a total of 304 participants in our hindsight group (208 Canadian residents and 96 U.S. residents).

Procedure. Figure 2 provides an overview of the procedure. The foresight group participated in the study between June 17, 2020, and September 14, 2020. They answered the hypothetical magnitude estimates, foreseeability, and inevitability questions without outcome information (see Experiment 1 procedure above). We recruited another group of participants (the hindsight group) from Canada and the U.S. who did not participate in the study during the first timepoint. The hindsight group participated in the study between September 28, 2020 and November 24, 2020. Participants in the hindsight group were instructed to complete the survey as if they were a same-aged peer of the same intelligence who completed the survey two months earlier. Specifically, they received the following instructions: “While completing this survey, we want you to imagine that you are one of your peers who is the same age as you and just as smart as you. We want you to imagine that your peer responded to this survey two months ago. Try to imagine how your peer would have responded to each of the questions below if they had filled

out this survey two months ago, knowing what they knew two months ago rather than what you know today."

Participants first answered the foreseeability and inevitability questions. Then, they received outcome information for the questions about case and death rates. This information appeared in the following format: "As of [date], there have been [number of cases] reported cases of COVID-19 in [country]. Please try to ignore your current knowledge of the number of cases in [country] and imagine how a same-aged peer, who is just as smart as you, would have responded to the following question if they had been asked to respond two months ago: 'Please estimate the total number of people who will have been infected by COVID-19 from when it began to two months from now in [country] (if you don't know, just take your best guess).'" They answered each of these 24 items. Finally, they answered several questions about their government's response to the pandemic, their own compliance with the restrictions and guidelines in their country, and completed demographic information.

Results and Discussion

Data Exclusions. We only included responses for which participants provided exact estimates. We also removed responses that exceeded 3 standard deviations above the mean. Of the 18,648 independent responses in Experiment 2, there were 121 (0.65%) excluded and missing (i.e., blank) responses and 80 extreme values (0.43%).

Hypothetical Magnitude Estimates. To see whether individuals showed hindsight bias for the hypothetical magnitude estimates component, we computed the following for both the foresight and hindsight groups: |Judgment – Correct Judgment|. We standardized all item scales by dividing the difference between each participant's judgment and the correct judgment by the

standard deviation of all participants' judgments for each item. Smaller numbers indicate that participants' judgments were closer to the correct judgments (i.e., more hindsight bias).

We excluded 14 participants from the foresight group and 9 participants from the hindsight group whose standardized hypothetical magnitude estimates were greater than 3 standard deviations above the mean. Thus, we analyzed data from 754 participants in total. Then, we conducted a 2 (group: foresight; hindsight) x 2 (country: Canada; U.S.) between-subject ANOVA with mean magnitude estimate as the dependent variable. There was a main effect of timepoint; however, Levene's Test was significant. Given our unequal sample sizes, we conducted a Welch's t-test and once again found a significant effect of timepoint, with the hindsight group's judgments being closer to the correct judgments (M = 0.15) than the foresight group's judgments (M = 0.25), $t(701.26) = 6.62$, $p < .001$, 95% CI [0.07, 0.13], $d = 0.47$.

Foreseeability Impressions. Consistent with Experiment 1, we conducted an exploratory factor analysis enforcing a two-factor solution with varimax rotation. This analysis revealed that Items 1, 2, 3, and 5 on our foreseeability scale consistently loaded more strongly onto Factor 1 while Item 4 loaded more strongly onto Factor 2 across the foresight and hindsight conditions (see Supplementary Materials). Therefore, we omitted Item 4 from the foreseeability scale. We calculated the mean of the four foreseeability items as indicated above, with negative values indicating that COVID-19 was unforeseeable and positive values indicating perceived foreseeability. Generally, participants rated COVID-19 as unforeseeable (i.e., their foreseeability ratings were significantly below 0), $t(753) = 20.45$, $p < .001$, 95% CI [-1.25, -1.03], $d = 0.75$.[4]

[4] We did not pre-register this analysis.

We conducted a 2 (group: foresight; hindsight) x 2 (country: Canada; U.S.) ANOVA for foreseeability ratings. Specifically, we compared the hindsight group's foreseeability ratings to the foresight group's foreseeability ratings at timepoint 1. None of the effects were significant (see Table 4).

Inevitability Impressions. Consistent with Experiment 1, we conducted an exploratory factor analysis enforcing a two-factor solution with varimax rotation. This analysis revealed that Items 2, 3, and 4 on our inevitability scale consistently loaded strongly onto Factor 1 while Item 1 loaded more strongly onto Factor 2 across the foresight and hindsight conditions (see Supplementary Materials). Therefore, we omitted Item 1 from the inevitability scale. We calculated the mean of the inevitability items as indicated above, with negative values denoting the pandemic outcomes were avoidable and positive values denoting the pandemic outcomes were inevitable. Generally, participants viewed COVID-19 as avoidable (i.e., their mean inevitability ratings fell significantly below 0), $t(753) = 7.37$, $p < .001$, 95% CI [-0.49, -0.28], $d = 0.27$.[5]

We conducted a 2 (group: foresight; hindsight) x 2 (country: Canada; U.S.) ANOVA for inevitability ratings. We compared the hindsight group's inevitability ratings to the foresight group's inevitability ratings at the first timepoint. There was an interaction between group and country; however, Levene's Test was significant. Given our unequal sample sizes, we conducted independent samples t-tests for each country with group as the independent variable.[6] We adjusted for multiple comparisons using the Bonferroni correction, adopting an alpha level of .025. In Canada, there were no differences in inevitability ratings between the foresight group

[5] We did not pre-register this analysis.
[6] Levene's Test was not significant for either t-test ($p = .66$ for Canada; $p = .98$ for U.S.); thus, we report the results of the t-tests with equal variances assumed.

and hindsight group, $t(409) = 1.48$, $p = .139$, 95% CI [-0.06, 0.43], $d = 0.14$. In the U.S., participants in the hindsight group rated COVID-19 as more inevitable than participants in the foresight group, $t(341) = 3.65$, $p < .001$, 95% CI [-1.04, -0.31], $d = 0.44$. Thus, despite U.S. participants in both the foresight and hindsight group rating COVID-19 outcomes as avoidable, those with current outcome information perceived them as more inevitable.

Correlations between Components. Finally, we conducted correlational analyses among the three components. The hypothetical magnitude estimate scale consisted of the standardized absolute difference between participants' judgments and the correct judgments. This scale yielded a Cronbach's α of .67 for the foresight group and a Cronbach's α of .58 for the hindsight group. The foreseeability scale consisted of the foresight group's original (i.e., first timepoint) ratings for the four foreseeability items and the hindsight group's estimations of how a peer without current outcome information would have responded to the foreseeability items two months earlier. This scale yielded a Cronbach's α of .79 for the foresight group and a Cronbach's α of .78 for the hindsight group. The inevitability scale similarly consisted of the foresight group's original ratings for the four inevitability items and the hindsight group's judgments regarding how a naïve peer would have responded to the inevitability items two months prior. This scale yielded a Cronbach's α of .61 for the foresight group and a Cronbach's α of .53 for the hindsight group.

Hypothetical magnitude estimates did not correlate with either foreseeability, $r(752) = .02$, $p = .525$ or inevitability, $r(752) = .00$, $p = .980$. However, foreseeability and inevitability positively correlated with one another, $r(752) = .13$, $p < .001$, though the correlation was weak.

General Discussion

We tested whether Canadian and U.S. residents showed hindsight bias for COVID-19 across the memory distortion, foreseeability, and inevitability components. In Experiment 1, we used a memory design. In foresight, participants estimated the case and death rates for different countries both two months and one year from their participation. Approximately 8-10 weeks later, they learned current case and death rates and attempted to recall their original judgments. We found that participants' recalled judgments were closer to the outcome information than their original judgments were. This shift toward the outcome information exemplifies the memory distortion component of hindsight bias.

Participants also rated their beliefs about the foreseeability and inevitability of COVID-19 at two different timepoints. They rated COVID-19 as relatively unforeseeable and avoidable. Although participants generally rated COVID-19 as avoidable, there were differences across countries: Canadian participants' inevitability ratings did not differ from the scale midpoint. Alternatively, U.S. participants rated COVID-19 as significantly avoidable. This may stem from differences in the government response within these two countries. Indeed, approximately 50% of U.S. participants who responded to an open-ended question about whether the government was doing enough to mitigate COVID-19's spread explicitly stated or alluded to the Trump administration's failed response and tendency to undermine scientists (see Hom, 2022).

Furthermore, participants also perceived COVID-19 outcomes as significantly more foreseeable in hindsight than in foresight (although, their ratings were on the *unforeseeable* side of the scale at both timepoints). As more information became known about COVID-19, participants perceived pandemic outcomes as more predictable. Recall that participants were asked to provide their current perceptions of foreseeability in hindsight rather than recalling their original judgments. This was done to avoid conflating memory distortions with foreseeability

impressions. Thus, using the outcome information to inform their foreseeability judgments served an adaptive function by allowing participants to base their decisions on the most current information in their knowledge base (Hoffrage et al., 2000).

Alternatively, there were no differences in inevitability ratings across the foresight and hindsight conditions. Recall that the inevitability component largely depends on people's ability to draw causal inferences between an outcome and the events preceding the outcome (Nestler et al., 2010; Pezzo, 2003). It may be the case that participants perceived the pandemic as avoidable because they considered what could have been done differently to prevent COVID-19 from reaching pandemic status (e.g., limiting travel, wearing masks). This may have biased their beliefs about the probability that this event had to unfold as it did, leading them to perceive COVID-19 as avoidable at both timepoints. Alternatively, the fact that participants perceived COVID-19 as relatively avoidable at both timepoints might suggest that they believed this event and the associated outcomes were surprising at both timepoints. Surprising outcomes can lead to reduced (or even reverse) hindsight bias (Ofir & Mazursky, 1997; Pezzo, 2003).

We also used a hypothetical design to determine whether similar patterns emerged across hindsight bias components. While foresight group participants judged case and death rates of various countries without outcome information, hindsight group participants attempted to ignore outcome information to estimate how a peer would have responded two months prior. Canadian and U.S. participants with outcome information made judgments that were systematically closer to the outcome than those without outcome information, demonstrating hindsight bias.

Despite this, Canadian and U.S. participants rated COVID-19 outcomes as relatively unforeseeable. There were no differences between countries or timepoints. This differs from Experiment 1's findings, where participants rated COVID-19 outcomes as significantly more

foreseeable in hindsight. In Experiment 1, participants rated their current perceptions of foreseeability in hindsight rather than attempting to recall their original judgments. Alternatively, in Experiment 2, participants in the hindsight group tried to ignore their current knowledge to imagine how a peer would have responded two months earlier with the information they had then. Participants in both countries also perceived the pandemic outcomes as relatively avoidable. Additionally, U.S. hindsight group participants rated COVID-19 as more inevitable than U.S. foresight group participants. Thus, current outcome information increased perceptions of inevitability (i.e., hindsight bias) among U.S. participants.

When we examined correlations among the three hindsight bias components in the memory design (i.e., Experiment 1), we found that memory distortions were unrelated to both foreseeability and inevitability. Conversely, foreseeability and inevitability positively correlated with each other. Similarly, in the hypothetical design, hypothetical magnitude estimates were unrelated to foreseeability and inevitability, while foreseeability and inevitability positively correlated with each other. These results show that perceptions of foreseeability and inevitability are largely distinct from judgments about one's own or others' previous knowledge (i.e., memory distortions and hypothetical magnitude estimates). Additionally, the correlation between foreseeability and inevitability appears to be small. However, it is possible that this stemmed from issues with the relatively poor internal reliability of these scales, which makes it harder to observe correlations (Miller & Lovler, 2019).

Limitations

This work had several limitations. First, we did not include control items for the memory distortion measure in the memory design. Given that studies show that participants rarely exhibit hindsight bias in the memory design control condition, we believe that we can reasonably

conclude that participants showed hindsight bias on this measure (Bayen, Erdfelder, Bearden, & Lozito; 2006; Erdfelder & Buchner, 1998; Groβ & Bayen, 2015). However, we acknowledge that while participants' judgments were closer to the correct judgments at the second timepoint than at the first timepoint, this could have resulted from regression to the mean (Pohl, 2007). Second, although we removed extreme values for participants' judgments regarding the case and death counts in various countries, the resulting standard deviations in the hypothetical design were still larger than those in the memory design. Recall that we standardized these item scales by dividing individual responses by the standard deviation of all responses. Thus, we believe that the observed difference between the foresight and hindsight group in the hypothetical design likely underestimates the true hindsight bias effect in this condition.

Finally, we observed relatively low internal consistencies on our foreseeability and inevitability difference score measures in the memory design. However, there are issues with assessing reliability for difference scores, and thus, difference scores often yield lower reliability coefficients than raw scores (Collins, 1996). Indeed, when we calculated the reliability of the foreseeability and inevitability measures for the foresight and hindsight conditions separately, we observed higher internal consistency for both the foreseeability (foresight Cronbach's $\alpha = .82$; hindsight Cronbach's $\alpha = .78$) and inevitability measures (foresight Cronbach's $\alpha = .57$; hindsight Cronbach's $\alpha = .69$).

Conclusion

Though others have discussed how hindsight bias affects beliefs and judgments about the COVID-19 pandemic (see e.g., Hom, 2022; Lechanoine & Gangi, 2020; Redelmeier & Shafir, 2020), this is the first study to empirically investigate hindsight bias for COVID-19. Our study has theoretical, methodological, and applied implications. We contribute to theory by further

demonstrating differences in the three hindsight components for a self-relevant and negative real-world event. We demonstrated that even as evolving events unfold, people rely on their outcome knowledge when judging what they or others knew prior. Yet, this did not translate to participants' foreseeability and inevitability ratings. Rather, participants often showed no hindsight bias for these components. Additionally, this is the first study to explore all three components using a hypothetical design (though see Blank & Peters, 2010; Nestler & Egloff, 2009 for examples of studies investigating foreseeability and inevitability in hypothetical designs). We observed a similar pattern of results across both the memory and hypothetical designs.

We contribute to methodology by studying hindsight bias for an evolving event. As we write this, we are still battling the pandemic. Circumstances continue to evolve across time. In reality, COVID-19 may never be "over" (Tarr, 2021). However, as things continue to evolve (e.g., case rates, death rates, variants, vaccines, etc.), people's beliefs and memories may change across time. While some work shows hindsight bias for conjectures (von der Beck, Cress, & Oeberst, 2019), we believe that our research demonstrates the potential to take multiple measures of hindsight bias as an event unfolds and provide data on the evolution of hindsight bias. We see this as an important and unique methodological contribution of our work.

Finally, our study has applied implications. There are few events in our lifetime that will have as devastating a global impact as COVID-19. Therefore, this work provides important insight into how hindsight bias affects people's judgments for major life events. Our study illustrates the importance of studying hindsight bias for real-world events, where contextual factors such as knowledge about the government's response or varying beliefs in the legitimacy of COVID-19 may change one's perceived sense of foreseeability and inevitability. This work

also has implications for how people will remember and learn from the COVID-19 pandemic. Generally, it shows that outcome knowledge distorts people's memories of COVID-19 such that they tend to misremember their previous judgments or unfairly assess others' naïve beliefs. People's tendency to overestimate what was known in the past may similarly affect other types of judgments (e.g., perceptions of authorities' responses to the pandemic). Therefore, we need more research investigating hindsight bias for COVID-19 to fully understand its impact on people's memories and beliefs about the pandemic.

References

Bayen, U. J., Erdfelder, E., Bearden, J. N., & Lozito, J. P. (2006). The interplay of memory and judgment processes in effects of aging on hindsight bias. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(5), 1003-1018. https://doi.org/10.1037/0278-7393.32.5.1003

Blank, H., Diedenhofen, B., & Musch, J. (2015). Looking back on the London Olympics: Independent outcome and hindsight effects in decision evaluation. *British Journal of Social Psychology, 54*(4), 798-807. https://doi.org/10.1111/bjso.12116

Blank, H., Nestler, S., von Collani, G., & Fischer, V. (2008). How many hindsight biases are there? *Cognition*, *106*, 1408-1440. https://doi.org/10.1016/j.cognition.2007.07.007

Christensen-Szalanski, J. J., & Willham, C. F. (1991). The hindsight bias: A meta-analysis. *Organizational Behavior and Human Decision Processes*, *48*, 147-168. https://doi.org/10.1016/0749-5978(91)90010-Q

Erdfelder, E., Brandt, M., & Bröder, A. (2007). Recollection biases in hindsight judgments. *Social Cognition*, *25*, 114-131. https://doi.org/10.1521/soco.2007.25.1.114

Erdfelder, E., & Buchner, A. (1998). Decomposing the hindsight bias: A multinomial processing tree model for separating recollection and reconstruction in hindsight. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(2), 387-414. https://doi.org/10.1037/0278-7393.24.2.387

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Fischhoff, B. (1975). Hindsight is not equal to foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance, 1*, 288-299. https://doi.org/10.1037/0096-1523.1.3.288

Giroux, M. E., Derksen, D. G., Coburn, P. I., & Bernstein, D. M. (2020, April 20). *Hindsight Bias for COVID-19*. https://osf.io/dmfjy/?view_only=1361d36a3ef34397b4d1074432962431

Guilbault, R. L., Bryant, F. B., Brockway, J. H., & Posavac, E. J. (2004). A meta-analysis of research on hindsight bias. *Basic and Applied Social Psychology, 26*, 103-117. https://doi.org/10.1080/01973533.2004.9646399

Groß, J., & Bayen, U. J. (2015). Hindsight bias in younger and older adults: The role of access control. *Aging, Neuropsychology, and Cognition*, *22*, 183-200. https://doi.org/10.1080/13825585.2014.901289

Hom, H. L. (2022). Perspective-taking and hindsight bias: When the target is oneself and/or a peer. *Current Psychology*, 1-12. https://doi.org/10.1007/s12144-021-02413-z

Lechanoine, F., & Gangi, K. (2020). COVID-19: Pandemic of cognitive biases impacting human behaviors and decision-making of public health policies. *Frontiers in Public Health, 8,* 613290. https://doi.org/10.3389/fpubh.2020.613290

Louie, T. A. (1999). Decision makers' hindsight bias after receiving favorable and unfavorable feedback. *Journal of Applied Psychology*, *84*, 29-41. https://doi.org/10.1037/0021-9010.84.1.29

Mark, M. M., & Mellor, S. (1991). Effect of self-relevance of an event on hindsight bias: The foreseeability of a layoff. *Journal of Applied Psychology, 76*, 569-577. https://doi.org/10.1037/0021-9010.76.4.569

Miller, L. A., & Lovler, R. L. (2019). *Foundations of psychological testing: A practical approach (6th edition).* Sage.

Müller, P. A., & Stahlberg, D. (2007). The role of surprise in hindsight bias: A metacognitive model of reduced and reversed hindsight bias. *Social Cognition, 25,* 165–184. https://doi.org/10.1521/soco.2007.25.1.165

Nestler, S., Blank, H., & Egloff, B. (2010). Hindsight ≠ hindsight: Experimentally induced dissociations between hindsight components. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*, 1399-1413. https://doi.org/10.1037/a0020449

Ofir, C., & Mazursky, D. (1997). Does a surprising outcome reinforce or reverse the hindsight bias? *Organizational Behavior and Human Decision Processes*, *69*(1), 51-57. https://doi.org/10.1006/obhd.1996.2671

Pezzo, M. (2003). Surprise, defence, or making sense: What removes hindsight bias? *Memory*, *11*, 421-441. https://doi.org/10.1080/09658210244000603

Pezzo, M. V., & Beckstead, J. W. (2008). The effects of disappointment on hindsight bias for real-world outcomes. *Applied Cognitive Psychology*, *22*, 491-506. https://doi.org/10.1002/acp.1377

Pezzo, M. V., & Pezzo, S. P. (2007). Making sense of failure: A motivated model of hindsight bias. *Social Cognition*, *25*, 147-164. https://doi.org/10.1521/soco.2007.25.1.147

Pohl, R. F. (2007). Ways to assess hindsight bias. *Social Cognition*, *25*, 14-31. https://doi.org/10.1521/soco.2007.25.1.14

Redelmeier, D. A., & Shafir, E. (2020). Pitfalls of judgment during the COVID-19 pandemic. *The Lancet Public Health*, *5*, e306-e308. https://doi.org/10.1016/ S2468-2667(20)30096-7

Roese, N. J., & Vohs, K. D. (2012). Hindsight bias. *Perspectives on Psychological Science, 7,* 411-426. https://doi.org/10.1177/1745691612454303

Tarr, J. (2021). *Fauci says COVID-19 might not be eliminated but it can be controlled.* Wisconsin Public Radio. https://www.wpr.org/fauci-says-covid-19-might-not-be-eliminated-it-can-be-controlled

Tykocinski, O. E. (2001). I never had a chance: Using hindsight tactics to mitigate disappointments. *Personality and Social Psychology Bulletin, 27,* 376-382. https://doi.org/10.1177/0146167201273011

Tykocinski, O. E., & Steinberg, N. (2005). Coping with disappointing outcomes: Retroactive pessimism and motivated inhibition of counterfactuals. *Journal of Experimental Social Psychology, 41,* 551-558. https://doi.org/10.1016/j.jesp.2004.12.001

von der Beck, I., Cress, U., & Oeberst, A. (2019). Is there hindsight bias without real hindsight? Conjectures are sufficient to elicit hindsight bias. *Journal of Experimental Psychology: Applied, 25*(1), 88-99. https://doi.org/10.1037/xap0000185

Wasserman, D., Lempert, R. O., & Hastie, R. (1991). Hindsight and causality. *Personality and Social Psychology Bulletin, 17,* 30-35. https://doi.org/10.1177%2F0146167291171005

Table 1

Glossary of Important Terms with Operational Definitions.

Memory Distortion	Outcome information biases people's recollections about what they previously knew or believed
Original Judgment	Participants' first judgment about a case or death rate, prior to learning the correct judgment
Correct Judgment	The correct answer participants learn about a case or death rate
Recalled Judgment	Participants' recall of their original judgment about a case or death rate, after learning the correct judgment
Foresight Judgment	Judgment made before learning outcome information
Hindsight Judgment	Judgment made after learning outcome information
Foreseeability	Subjective beliefs about whether one could have predicted an outcome
Inevitability	Perceptions of the objective likelihood that an outcome had to occur
Hypothetical Magnitude Estimate	Participants' judgment about a case or death rate in the hypothetical design

Table 2

Foreseeability and inevitability items used in Experiments 1 and 2

Foreseeability Items	**Experiment 1**	**Experiment 2**
1. How predictable was the COVID-19 pandemic?	X	X
2. I knew all along that this pandemic would occur.	X	X
3. The outcomes of this pandemic were clearly predictable.	X	X
4. From the course of the spread of this pandemic, one could not expect different outcomes of this pandemic.		
5. It was difficult to predict how the pandemic would turn out.*	X	X
Inevitability Items		
1. How inevitable was the COVID-19 pandemic?		
2. Under the circumstances, no different outcomes of the pandemic could have been expected.	X	X
3. If the pandemic had occurred one year later, it would have turned out exactly the same way.	X	X
4. Nothing could have influenced the outcomes of the pandemic.	X	X

*This item was reverse scored. X's denote that Item was included in Experiment. Reliability coefficients for the foreseeability scale are: (1) Experiment 1 = .60; (2) Experiment 2 Foresight = .79; (3) Experiment 2 Hindsight = .78. Reliability coefficients for the inevitability scale are: (1) Experiment 1 = .48; (2) Experiment 2 Foresight = .61; (3) Experiment 2 Hindsight = .53.

Table 3

Mean Memory Distortions, Foreseeability Ratings, and Inevitability Ratings [and 95% Confidence Intervals] as a Function of Timepoint and Country in Experiment 1.

	Foresight	**Hindsight**
	Memory Distortions	
Canada	0.46 [0.36, 0.56]	0.23 [0.19, 0.26]
U.S.	0.46 [0.38, 0.54]	0.19 [0.16, 0.22]
Overall	0.46 [0.40, 0.52]	0.21 [0.18, 0.23]
	Foreseeability	
Canada	-1.17 [-1.52, -0.82]	-0.98 [-1.26, -0.70]
U.S.	-1.24 [-1.56, -0.93]	-1.09 [-1.41, -0.76]
Overall	-1.21 [-1.44, -0.98]	-1.04 [-1.25, -0.82]
	Inevitability	
Canada	-0.04 [-0.33, 0.25]	-0.24 [-0.52, 0.04]
U.S.	-1.17 [-1.44, -0.90]	-1.30 [-1.62, -0.98]
Overall	-0.66 [-0.87, -0.45]	-0.82 [-1.05, -0.59]

Note. For memory distortions, smaller values in the hindsight column than the foresight column denote hindsight bias. For foreseeability and inevitability ratings, larger values in the hindsight column than the foresight column denote greater foreseeability or inevitability with outcome information. Non-overlapping confidence intervals between foresight and hindsight judgments indicate significant differences under null hypothesis testing.

Table 4

Mean Hypothetical Magnitude Estimates, Foreseeability Ratings, and Inevitability Ratings [and 95% Confidence Intervals] as a Function of Group and Country in Experiment 2.

	Foresight	**Hindsight**
	Hypothetical Magnitude Estimates	
Canada	0.26 [0.22, 0.30]	0.16 [0.14, 0.18]
U.S.	0.24 [0.21, 0.27]	0.13 [0.11, 0.15]
Overall	0.25 [0.22, 0.27]	0.15 [0.14, 0.17]
	Foreseeability	
Canada	-1.18 [-1.38, -0.98]	-1.25 [-1.44, -1.05]
U.S.	-1.07 [-1.27, -0.86]	-1.01 [-1.36, -0.67]
Overall	-1.12 [-1.26, -0.98]	-1.17 [-1.34, -1.00]
	Inevitability	
Canada	-0.01 [-0.17, 0.16]	-0.19 [-0.37, -0.01]
U.S.	-0.92 [-1.10, -0.73]	-0.24 [-0.57, 0.09]
Overall	-0.50 [-0.63, -0.37]	-0.21 [-0.36, -0.05]

Note. For hypothetical magnitude estimates, smaller values in the hindsight column than the foresight column denote hindsight bias. For foreseeability and inevitability ratings, larger values in the hindsight column than the foresight column denote greater foreseeability or inevitability with outcome information. Non-overlapping confidence intervals between foresight and hindsight judgments indicate significant differences under null hypothesis testing.

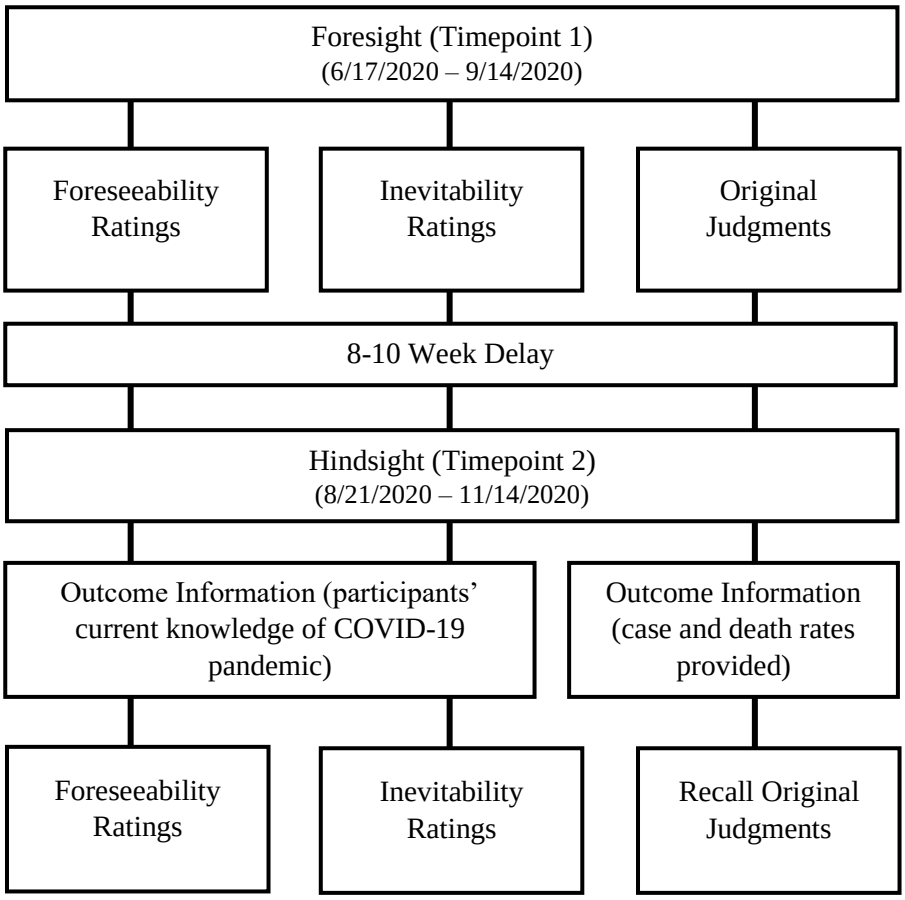

Figure 1. Overview of Experiment 1 procedure.

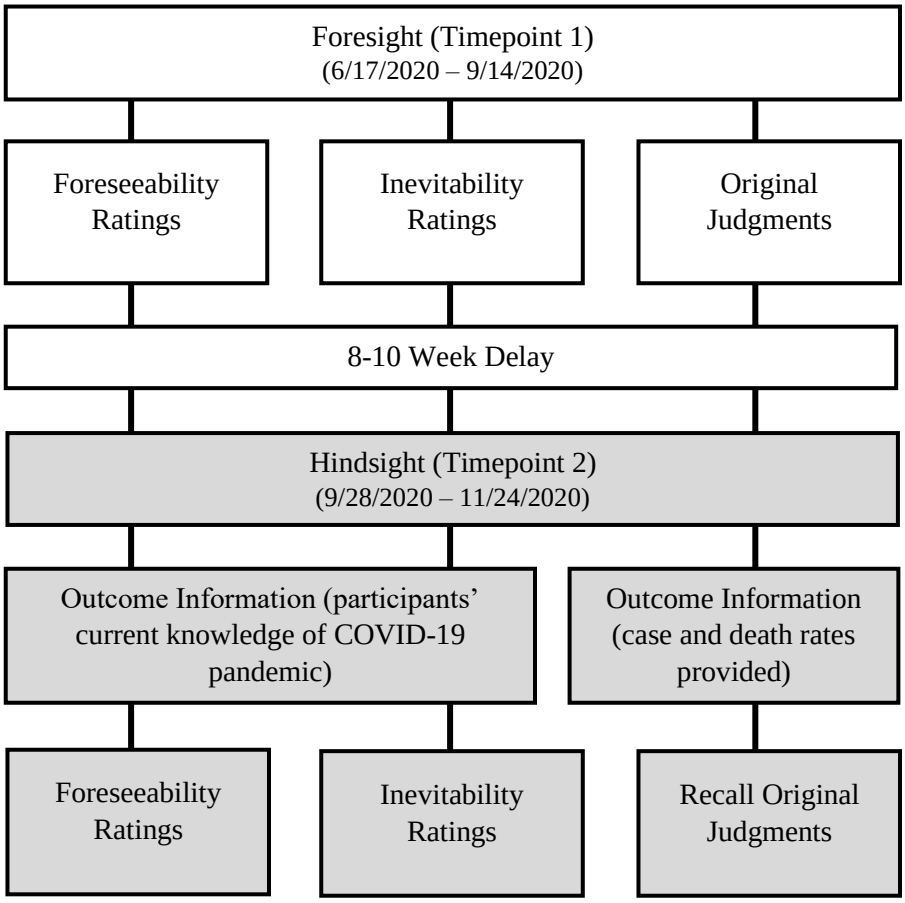

Figure 2. Overview of Experiment 2 procedure. Grey boxes denote an independent sample of participants.