NORTHWESTERN UNIVERSITY

Counterfactual Closeness:

How some nonactual events appear closer than others

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL IN
PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Psychology

By

Matthew Myers

EVANSTON, ILLINOIS

2022

CHAPTER 2

PROPENSITY AND MEASURES OF CLOSENESS

2.1 Introduction

According to Kahneman and Varey's (1990) propensity framework, impressions of counterfactual closeness can be explained by two concepts informed by specific event properties. First, *propensity* describes one's belief that the likelihood of a future outcome, based on event cues preceding the outcome's resolution, increases as the event unfolds. If the reasoner believes that an outcome's likelihood is increasing, then they tend to perceive this as evidence of progress or momentum towards that outcome. When momentum diminishes and the outcome fails to actualize, either because later forces directly impede the progress made towards the target outcome or because such progress is eventually eclipsed by comparatively greater gains in momentum from a competing outcome, reasoners tend to interpret this as evidence that the non-actualized outcome had almost occurred. Second, *decisiveness* describes the point in the event sequence at which reasoners are able to infer, with a reasonably high degree of certainty, that the eventual outcome is practically assured. The propensity framework contends that since decisiveness marks the subjective transition between an outcome which is open and unknown to one which is closed and determined, prolonging the onset of the decisive moment, and thereby extending the uncertainty of the outcome later into the event's course, should further contribute to the perceived closeness of the counterfactual alternative. The propensity framework thus grounds the counterfactual plausibility of alternatives in dual considerations of progress (i.e., propensity) and future certainty (i.e., decisiveness).

A number of studies support the view that CCF reasoning processes are responsible for a variety of behavioral effects, such as judgments of luck (Teigen, 1995, 1996; 1998), betting practices (Wahl & Enzl, 2003; Clark, Liu, McKavanagh, Garrett, Dunn, & Aitken, 2013; Wu, Dijk, Li, Aitken, & Clark, 2017), and emotional reactions (Johnson, 1986; Miller & McFarland, 1986; Medvec, Madey, & Gilovich, 1995; Kanten & Teigen, 2015). While the propensity framework seems well positioned to explain such influences, no studies have yet empirically tested the framework's predictions on them, nor has it been shown that its predictions correspond with direct measures of CCF beliefs. As such, the current status of the propensity framework as an account of CCF reasoning rests largely on indirect support.

The primary aim of the following chapter is to provide stronger empirical justification for the propensity framework by testing the account's ability to predict the conditions under which CCF beliefs should arise. The central question is does manipulation of propensity and the onset of the decisive moment independently contribute to people's CCF beliefs. In testing this prediction, I seek to overcome two important methodological difficulties particular to research on counterfactual proximity. One problem is the difficulty associated with creating experimental paradigms that allow for the independent manipulation of propensity and decisiveness, two subjective properties of events that are highly sensitive to small contextual cues. The following experiments seek to mitigate this difficulty by exploiting perceptual properties of physical events, such as spatial proximity and the orientation of objects' movement (i.e., approaching versus departing) relative to the target counterfactual, in order to exercise more precise control over the determinants of propensity and decisiveness. A secondary issue relates to the selection of terms used as measures of counterfactual proximity. As discussed earlier, substantive differences in the semantic and pragmatic properties of different counterfactual proximatives can make it difficult to disambiguate counterfactual proximity from other elements of the proximative's meaning. It is further possible that different proxmatives, as well as non-proximative terms whose meaning have CCF entailments, may in fact emphasize subtlety different event features and are thus nonequivalent CCF measures. The experiments presented in this chapter therefore examine the relationship between common CCF expressions, focusing specifically on *close*, *almost*, *barely*, and the non-proximative *lucky*, in order to assess the degree to which they measure the same psychological construct of counterfactual proximity.

2.1.1 Luck as a Measures of Counterfactual Closeness

What makes some outcome appear luckier than another? Early psychological accounts hypothesized that people ascribe luck to outcomes when they are beneficial for some individual or group and when they result from factors that, either through circumstance or by their nature, are uncontrollable (Heider, 1958; Weiner et al., 1972). Much of the discussion over luck has attempted to explicate the latter notion of controllability. Some theorists have hypothesized that *chance* is a necessary component of luck's meaning. When making luck-based claims, speakers assert that factors that influenced the outcome were governed by unknown and external forces beyond which any person could have reasonably predicted. According to this view, the reason why winning the lottery is lucky is because, while the winner may have won through a process which he himself instigated, the means by which he came to win was determined by factors beyond his personal control and, either due to excessive complexity or limited knowledge, were ill understood. These accounts therefore treated luck as a particular type of chance event, one for which the outcome holds a significant degree of importance for some person or group.

However, subsequent findings have suggested how, rather than luck being a subcategory of chance, people's understanding of these two concepts differ in important respects. Wagenaar and Keren (1988) found that among twelve assessment dimensions associated with luck and chance, including luck and chance themselves, people's judgments of stories along these dimensions could be characterized as broadly fitting two distinct conceptual categories. Results of a factor analysis revealed strong associations among the dimensions of *luck*, *consequential importance*, and *accomplishment* on the one hand and *chance*, *surprise*, and *coincidence* on the other. Moreover, by directly manipulating a story's degree of coincidence (i.e., whether personal choice or uncontrolled environmental factors determined an outcome) they found a significant difference in judgments of chance, but no corresponding difference in judgments of luck. Experimental findings by Teigen (1983) further complicate the relationship between luck and chance. When given two stories about lottery winners who participated in equally sized but probabilistically different lotteries, participants' assessment of the two outcomes showed no differences in luck attributions, even though the likelihood of winning one lottery was twice as likely than the other. The dissociation observed in these and other findings (see also Teigen, 1996) cast serious doubt on the extent to which chance is an intrinsic feature of luck attributions.

More recently, Pritchard and Smith (2004) have argued for an alternative account that reconciles these and other problems with luck by substituting the notion of chance with counterfactual proximity. Similar to earlier accounts, they maintain that the impact of the outcome must be significant for some individual or group; however, rather than appealing to chance to explain the unlikelihood typical of most lucky outcomes, they instead argue that luck should be treated as a similarity evaluation comparing possible worlds to our own. The kind of similarity they propose is one based on counterfactual change: Highly similar worlds will be those for which only a small historical change is sufficient to have made those worlds a reality. By evaluating similarity in this sense, we can establish an ordering of possible worlds corresponding to their counterfactual accessibility (see section 1.2.3 for more on possible worlds). Considering only those worlds that are most similar (and, consequently, most accessible) to our own, the authors contend that reasoners will regard a real-world outcome as lucky when they deem most of the nearest counterfactual outcomes as comparatively worse.

The advantage of this change is that, by characterizing luck as a modal concept, the account can explain why people often consider outcomes as lucky even when the antecedents giving rise to them are perceived as probable (e.g., successfully using an elevator). Generally speaking, outcomes that are probable are made that way by reliable mechanisms that tend to behave invariantly across different contexts and situations. Such reliability tends to give the impression that substantively different outcomes were less than attainable, and hence not lucky, since it would require the countenance of exceptional circumstances to allow for such alternative outcomes. However, these tendencies can be undone when further information about features of the event hint at some means by which the outcome could have been different. For example, riding an elevator and successfully stepping out onto the twentieth floor is a reliable and commonplace affair for many people; yet such a mundane outcome will seem quite lucky if we later learn that half of the carriage's support cables snapped while riding up. The luckiness of an outcome therefore seems to depend on those details of events which lend themselves to establishing some minimal degree of counterfactual revision that would have made the outcome different, such as the snapping of the remaining cables. Thus, the account contends that luck is not *directly* related to notions of improbability and controllability, but rather *indirectly* related via particular event properties which allow people to easily envision alternative outcomes.

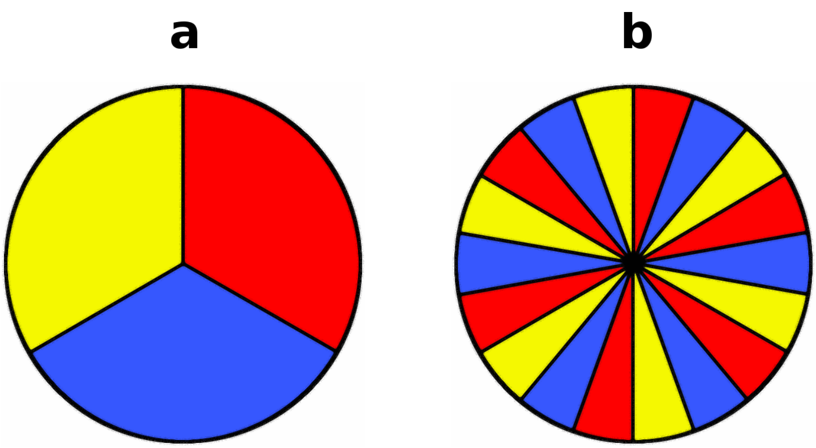


Fig. 2.1 Stimuli from Experiment 1 of Teigen (1996). The left image (a) is a three-sector wheel, and the right image (b) is an eighteen-sector wheel. The probability of landing on any color is 33.3% for both wheels.

If Pritchard and Smith (2004) are correct in their characterization of luck as a modal concept, then CCF intuitions should directly inform people's luck attributions. Indeed, findings by Teigen (1996) support this conclusion. In one experiment examining the influence of outcome probability on luck attributions, Teigen had participants rate the perceived luckiness of outcomes in a wheel-and-spinner game. Two wheels were shown to participants, both of which had three types of colored sectors: Red, blue, and yellow. The distribution of colors were equally proportioned across both wheels; however, the number of sectors differed between them, with one containing three sectors and the other containing eighteen. Figure 2.1 shows the two wheels presented to participants. Players won if a the spinner was described as having landed in the center of a red sector, an outcome that was approximately 33.3% likely to occur for both wheels. Yet despite their equivalent probabilities, Teigen found that participants rated winning outcomes for the eighteen-sector wheels as significantly more lucky than the three-sector wheels. He hypothesized that the reason for the observed difference was because people were engaging in the same CCF reasoning processes outlined in the propensity framework: Because each sector within the eighteen-sector wheel occupied a much smaller area of the wheel, the spinner always landed nearer to the sector's boundary, and therefore gave the impression that it had shown greater propensity either to have stopped within the earlier losing sector or to have continued into the following losing sector.

It is well established that, for physical events, close spatiotemporal proximity is a strong predictor as to whether one will believe some alternative outcome was counterfactually close (Kahneman & Varey, 1990; Roese, 1997; Fischhoff, Gonzalez, Small, & Lerner, 2003; Teigen, 2005). The presumption is that spatiotemproal proximity tends to indicate that there existed some potential for contact or some interaction between entities. If the path of one object either approaches or overlaps the path of another such that neither object makes contact with the other *and* there exists a point in time where the physical distance separating them is perceived to be small, it gives the reasoner an impression that a small change to earlier events could have led to their interaction (although, see Chapter 3 for more on the relationship between spatiotemproal proximity and CCF beliefs).

However, it is possible that the close spatial proximity in Teigen's (1996) experiment may have influenced luck attributions in a way that is unrelated to beliefs about the counterfactual accessibility of alternative outcomes. While it is conceivable that participants visualized an event sequence that induced a feeling of progress (e.g., imagining the spinner gradually approach the losing sector before finally landing on the winning sector), they may have instead made a luck-based generalization from the mere occurrence of close spatial proximity. In other words, they may have formed their luck attributions by interpreting the observed spatial proximity as simply being an event property typical of lucky outcomes and, thus, did not actually consider the accessibility of relevant alternatives. Games of bingo can often elicit these types of generalizations. When a called number fails to complete a winning sequence, the outcome can feel considerably more unlucky when it is directly adjacent to the would-have-been winning number than when it is on the opposite side of the card. However, this judgment is unlikely to be predicated on any assessment of counterfactual feasibility, seeing as the random assignment of numbers on bingo cards make the physical proximity of any two numbers entirely uninformative as to what could have easily occurred. Any inclination to nonetheless label such events as more unlucky would therefore suggest that, at least in some circumstances, spatial proximity can *ipso facto* make certain outcomes appear more or less lucky. The possibility of such spatial generalizations is further supported by evidence of other types of luck-based generalizations. People frequently use categorical knowledge of particular event kinds to structure a set of contrastive outcomes by which they assess the actual outcome's luckiness (Kahneman & Miller, 1986). By contrasting a real outcome with categorical exemplars in this way, it effectively divorces assessments of luck from considerations of counterfactual attainability. For example, the survivor of a car crash may express feeling lucky to have only broke his arm, not because he necessarily believes that the most proximate counterfactuals were worse, but rather because he believes that a broken arm is a better outcome than those typical of the event category *car crashes* (e.g., head-on collisions, vehicle rollovers, etc.).[1] Such possibilities offer an alternative explanation for Teigen's results: If people relied on proximity-based generalizations when assessing wheel-and-spinner outcomes, then it is possible that CCF reasoning processes were not responsible for the observed differences in people's luck attributions.

2.1.2 Overview of Experiments

The following experiments seek to examine whether the propensity framework is capable of predicting people's judgments of luck as well as other measures of CCF beliefs. The experiments' design borrows directly from Teigen's (1996) original wheel-and-spinner paradigm. The rational for using this design is twofold. First, it allows for an approximate replication of Teigen's original findings on luck attributions; and second, the physical properties of wheel-and-spinner games are well suited for testing the broader predictions of the propensity framework, as exploiting the physical location of spinners provide a conve-

[1] It is worth noting that such cross category comparisons are a normative feature of luck attributions. Events can be lucky either because one perceives a counterfactual outcome as close or because one perceives the actual outcome as better or worse than outcomes typical of the event category. Both the bingo and car crash scenarios were examples of genuinely lucky events because they met the latter consideration, even though they did not meet the former. In contrast, only the former consideration is normatively relevant for CCF judgments. Despite this, I suspect that people would in fact accept the claim *he almost won* in the bingo scenario, even though, as explained earlier, spatial proximity does not indicate greater counterfactual accessibility. This may be because people erroneously believe that spatial proximity made the winning outcome more plausible and that, therefore, luck beliefs may not have stemmed from typicality judgments. Nevertheless, the example is merely meant to demonstrate that such luck attributions are appropriate in these circumstances, and for reasons unrelated to counterfactual accessibility.

nient way of manipulating propensity and decisiveness. Experiments 1a and 1b first examine people's judgments of static images of wheel-and-spinners for both losing and winning outcomes. Participants will be tasked with assessing both three- and eighteen-sector wheels. Assessment measurements include luck attributions and the proximative attribution *almost*. I also include the proximative attributions *close* and *barely* in Experiments 1a and 1b, respectively. However, the spatial proximity depicted in static images of wheel-and-spinner outcomes can only provide indirect evidence of propensity and decisiveness. Experiments 2a and 2b therefore examine judgments for animated wheel-and-spinner scenarios in order to assess how temporal cues to propensity, such as the spinner's speed and deceleration over time, influence people's beliefs about luck and counterfactual proximity. These experiments focus exclusively on measures of *almost* and *barely*.

The broad prediction is that people's judgment of wheel-and-spinner outcomes will follow from those CCF determinants outlined in Kahneman & Varey's (1990) propensity framework; namely, people will give higher ratings for luck and proximative attributions in scenarios where there is a high degree of indirect (Experiment 1ab) and direct (Experiment 2ab) evidence of propensity towards counterfactual alternatives *and* when the decisiveness of the outcome (i.e., when the outcome is all but certain for the reasoner) occurs later in the event sequence. Furthermore, like Teigen (1996), I hypothesize that event cues communicating information about propensity and decisiveness, rather than mere proximity-based generalizations, strongly determine people's luck judgments. Consequently, I predict that people's judgments of static outcomes in Experiments 1ab will replicate his original findings and will further correlate with other proximative attributions. Finally, I predict that these same patterns of results will hold when participants assessed directly observable indicators of propensity and decisiveness for the animated scenario in Experiments 2ab.

2.2 Experiment 1a: The unluckiness of losing outcomes

To accommodate the predictions of these experiments, changes were made to Teigen's (1996) original design. Rather than assessing winning outcomes, participants instead assess losing outcomes. They will therefore rate the *unluckiness* of outcomes and the degree to which they believe a player *almost* won and was *close* to winning. This change is partly motivated by research demonstrating how negative outcomes are more likely to elicit consideration of counterfactual alternatives (Kahneman & Miller, 1986; Gavanski & Wells, 1989; Boninger, Gleicher, & Strathman, 1994; Roese, 1997). But more importantly, changing the results to losing outcomes allows for a manipulation that better distinguishes between propensity and decisiveness. In the original stimuli, the close spatial proximity between spinners and sector boundaries for eighteen-sector wheels provided indirect evidence both that spinners showed strong propensity to cross the sector boundary and that there was a high degree of uncertainty about the outcome up until the final moments of the event. However, the resulting uncertainty is more strongly induced from the spinner's proximity to one sector boundary than the other: Until the spinner comes to a complete stop, we remain highly uncertain whether or not it will eke past the approaching sector boundary; however, we are much sooner assured that losing will not occur from landing in the earlier sector, since this possibility is no longer possible as soon as the spinner passes the earlier sector boundary. Teigen's original design was insensitive to this asymmetry since the resulting winning sector was necessarily flanked by two losing sectors, therefore ensuring that the spinner was always approaching a losing sector. But by instead having participants assess losing outcomes, it becomes possible to influence how late uncertainty is preserved into the event by changing which side of the losing sector the winning sector occurs. The location of the winning sector relative to the resulting losing sector therefore serves as the manipulation of decisiveness.

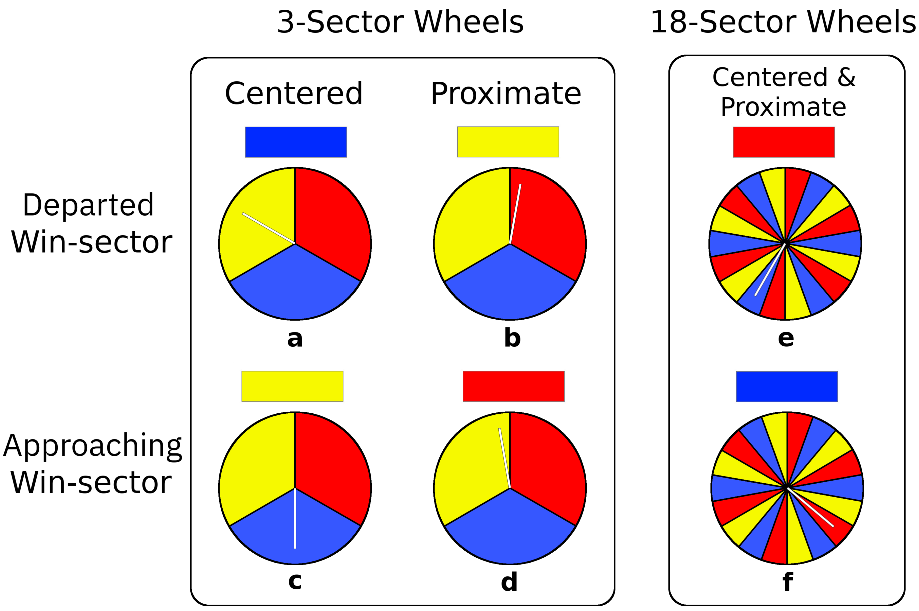


Fig. 2.2 Six wheels used in Experiment 1a. Winning color is depicted above each wheel. Number of sectors differ by three-sector wheels (abcd) and eighteen-sector wheels (ef). Outcome type differs by departed (abe) and approaching (cdf) wheels. Spinner location differs for three-sector wheels by centered (ac) and proximate (bd) wheels, while eighteen-sector wheels are both centered and proximate (ef).

As an example, consider the stimuli for Experiment 1a presented in Figure 2.2, which shows six wheels that vary by the location of the winning sector, the color of which appears as a block above each wheel. Wheels (2.2e) and (2.2f) depict two outcomes where the spinner is equally close to a winning sector boundary. Assuming both spinners rotate clockwise and decelerate at similar rates, the decisive moment for wheel (2.2f) will occur later in the event compared to wheel (2.2e). Since the spinner is approaching the winning sector (i.e., red), the outcome will remain highly uncertain until it comes to a complete stop. However, as soon as the spinner in wheel (2.2e) passes the winning sector (i.e., blue), a losing outcome is all but certain seeing as an entire sector separates it from the next winning sector. The propensity framework therefore predicts that people should perceive the winning outcome as closer when the spinner approaches the winning sector boundary. I refer to these two conditions as *departed* scenarios (i.e., wheels (2.2abe)) and *approaching* scenarios (i.e., wheels (2.2cdf)).

A further change is the inclusion of spatially proximate spinner-boundary configurations among the three-sector wheels. These spinners, depicted in wheels (2.2b) and (2.2d), are the same distance from the winning sector boundary as the spinners in the eighteen-sector wheels. For both three- and eighteen-sector wheels, I refer to these outcomes as *proximate* scenarios. While there is no direct prediction regarding the manipulation of proximity between three- and eighteen-sector wheels, these items were included to see whether the size of a sector's area exerted any additional effect on people's judgments beyond indirect evidence of propensity and decisiveness given by spinner-boundary proximity.

Finally, in addition to luck attributions, participants will rate their degree of agreement with statements describing the player's counterfactual proximity to winning using the proximatives *almost* and *close*. While much of the research on CCF reasoning has focused on *almost*, considerably less attention has been given to the proxmative *close*. This omission is likely a conscious choice on the part of researchers, seeing as *close* is intimately associated with spatial proximity. Much like *almost*, *close* can establish different dimensions of comparison based on the semantic properties of what is being compared: When *close* modifies a past telic event, people will interpret the assessment dimension as counterfactual proximity (see section 1.2.1). Consider the the sentences *the player was close to winning* and *the spinner was close to the winning sector*. *Close* in the latter sentence describes a proximate relationship between two physical entities possessing clearly delineated boundaries, the properties of which invite the reasoner to compare the their physical proximity. In contrast, reference to the telic achievement of winning in the former sentence precludes the possibility of a purely spatial comparison, since space, albeit a relevant consideration for establishing counterfactual proximity, is by itself an inappropriate dimension for measuring event-level proximity. However, the frequent use of *close* as a comparison of spatial proximity introduces the risk that counterfactual evaluations might be confused for simple spatial evaluations. For example, people may simply misinterpret requests for *closeness* judgments as requests for spatial judgments or place greater weight on the spatial properties of events when assessing counterfactual proximity.

To summarize, there are three specific predictions for Experiment 1a. First, despite noteworthy differences between luck and the proximatives *almost* and *close*, a property shared by all three terms is that their interpretation depends, at least partly, on measures of counterfactual proximity. As such, I predict that people's assessment of outcomes described by statements using these terms will show a strong degree of correlation with one another. Second, consistent with Teigen (1996), I predict that the closer spatial proximity for centered eighteen-sector wheels will create impressions both that the spinner had shown greater progress towards the counterfactual alternative (i.e., propensity) and that the outcome was uncertain until the final moments of the event (i.e., decisiveness), therefore contributing to higher overall assessment ratings compared to three-sector wheels. And third, I predict that when controlling for propensity by equating spinner-boundary proximity, decisiveness will have a unique effect on assessment ratings. Specifically, among proximate conditions, outcomes where spinners approach the winning sector will be perceived as having been more uncertain in the final moments of the event and will thus receive overall higher assessment ratings compared to outcomes where spinners departed the winning sector.

2.2.1 Method

Participants. The following analyses include 55 Northwestern University undergraduates. Data from 8 participants were removed because they failed a comprehension check.

Design & Materials. Experiment 1a employs a 2 (wheel: three- and eighteen-sector wheel) $\times$ 2 (sector boundary: Departed and approaching) $\times$ 2 (spinner location: Centered and proximate) $\times$ 3 (assessment: *Almost won*, *close to winning*, and *was unlucky*) within-subjects design. Stimuli were static images of wheel-and-spinner outcomes. Figure 2.2 presents the complete set of target items.

Since outcomes for eighteen-sector wheels did not vary by spinner-boundary proximity (i.e., outcomes were both centered and proximate), it means the results cannot be analyzed using the full factorial design. This was an intentional choice brought about from wanting to replicate Teigen (1996) and also control for spinner-boundary spatial proximity. However, this means that testing specific predictions require analyses that consider only a subset of the data. When examining the influence of spatial proximity, like Teigen, I consider only centered outcomes across three- and eighteen-sector wheels, and therefore omit proximate three-sector wheel from the analysis.[2] Similarly, when analyzing the unique influence of decisiveness, I consider only proximate spinner-boundary configurations across three- and eighteen-sector wheels, thus omitting centered three-sector wheels from the analysis. Therefore, both analyses treat the experiment as a 2 (wheel) $\times$ 2 (sector boundary) $\times$ 3 (assessment) factorial design, with tests of spatial proximity focusing only on centered outcome and tests of decisiveness focusing on proximate outcomes.

In addition to the target stimuli shown in Figure 2.2, two additional eighteen-sector wheels, both depicting winning outcomes, were included in each assessment block. This was done to help give the impression of fairness (i.e., 25% overall winning outcomes). These winning outcomes also served as comprehension checks. Rather than rate their agreement with the particular assessment, participants were instead instructed to select the special option "THIS PLAYER WON" when presented with a winning outcome. Participants' data were removed from analysis if they selected any response other than "THIS PLAYER WON" for any of the six winning outcomes.

The addition of two winning outcomes means that participants saw a total of eight distinct wheel-and-spinner outcomes. All eight outcomes were presented to participants on three separate occasions, once for each of the three assessments (i.e., *was unlucky*, *close to winning*, and *almost won*). Accordingly, a total of twenty-four assessments were made.

[2]For this analysis, the factor *sector boundary* (i.e., the decisiveness manipulation) is unlikely to have a significant influence on participants' assessment ratings. As mentioned before, closer spatial proximity indicates not only greater propensity towards the alternative, but also strongly suggests a later onset of the decisive moment. Having the winning sector either follow or precede the actual outcome is, comparatively, a much more subtle manipulation of decisiveness. Consequently, spatial proximity will likely overshadow any effect of whether the spinner departed or is approaching the winning sector. I therefore expect that differences in which particular sector boundary is most proximate will only meaningfully influence ratings when controlling for spatial proximity.

The three assessments were completed in separate blocks administered sequentially. When beginning a new block, participants received an instructional page informing them of the changes made to the statements they would be rating. Block orders were created using a Latin square design and were randomly assigned to participants using Qualtrics' randomization algorithm.

Procedure. The experiment was administered on PCs using the internet-based program Qualtrics and took approximately ten minutes to complete. The instructions presented participants with information about the two types of wheels (three-sector versus eighteen-sector) and the conditions for winning. Participants were told that they would be assessing the outcome of a player's spin. A picture of the two wheels accompanied this explanation, along with a statement saying that, for both types of wheels, sectors had equal area. They were further told that the spinners would always rotate in a clock-wise direction.

After reading the instructions, participants continued to the assessment phase. Assessment type was presented sequentially based on the block order assigned to participants. All items were presented on separate pages, with participants being unable to revisit and revise past responses. For each item, below the wheel, participants read a statement about the outcome. Depending on the assessment block, the written statement said either "This player was unlucky," "This player was close to winning," or "This player almost won." Participants were then asked to indicate on a numbered 1-10 scale how much they agreed with this statement. Written above the scale's poles were "1–Very much disagree" and "10–Very much agree." To the right of the scale, the option "THIS PLAYER WON" was provided. On the last page, participants were asked to provide a written explanation of which features they considered and how that affected their ratings.

2.2.2 Results and Discussion

Analyses and discussion of the results are organized according to the experiment's three main predictions. These were (1) agreement for statements describing outcomes as *was unlucky*, *close to winning*, and *almost won* will be strong correlated, (2) centered eighteen-sector outcomes will receive higher aggregate assessment ratings compared to centered three-sector outcomes, and (3) proximate outcomes where the spinner is approaching the winning sector will receive higher aggregate assessment ratings compared to proximate outcomes where the spinner had just departed the winning sector. Consistent with the first prediction, participants' ratings of wheel-and-spinner outcomes showed clear correlations among all three assessment, with the relationship between *close to winning* and *almost won* being particularly stronger than *was unlucky* to either *close to winning* or *almost won*. The correspondence between luck and the other more direct CCF assessments supports Teigen's (1996) hypothesis that evidence of propensity inform people's luck attributions. However, subsequent findings revealed that while centered eighteen-sector wheels received higher ratings for *almost won* and *close to winning* assessments than centered three-sector wheels, no difference was observed for *was unlucky* assessments. Despite the overall correlation between luck and the other assessments, this finding suggests that people may be using additional criteria beyond propensity when making luck attributions. Finally, contrary to the third prediction, aggregate assessment ratings did not differ as a function of whether the spinner was approaching or had departed the winning sector. In sum, the findings indicate that spatial proximity was an important consideration for *almost won* and *close to winning* assessments, which suggests that evidence of propensity played a role in informing people's judgments; however, the failure to demonstrate an effect of decisiveness leaves open the possibility that this was instead due to simple spatial proximity biases. Detailed findings are presented in the following sections.

<u>Assessment Correlations</u>. Linear mixed-effects models, conducted with the R package "lme4" (Bates et al., 2020) were used to determine whether there were any meaningful differences between assessment ratings. All analyses utilized random-intercept models that treated assessment type (i.e., *was unlucky*, *almost won*, *close to winning*) as a fixed-effect, and participants and an assessment-by-participants interaction as random effects. [3] An overall analysis of ratings revealed a significant difference in assessment type, $\chi^2(2) = 8.78$, $p < .05$. Subsequent mixed-effects analyses revealed that ratings for average *was unlucky* assessments ($M = 5.86$, $SD = 2.31$) were significantly greater than *almost won* assessments ($M = 5.23$, $SD = 2.48$), $\chi^2(1) = 6.01$, $p < .05$; and marginally greater than *close to winning* assessments ($M = 5.41$, $SD = 2.49$), $\chi^2(1) = 2.98$, $p = .08$. However, there was no difference in ratings between *close to winning* and *almost won* assessments, $\chi^2(1) = 1.30$, $p = .25$.

To examine the relationship between these assessments across items, I further conduct a series of correlations. Due to existence of multiple observations per participant, the fol-

[3] It would be reasonable to argue that random-slope models would have been more appropriate since such models would have further accounted for variation in assessment type by participant. However, preliminary tests that included random slopes resulted in warnings of "singular-fit" by the lme4 package, indicating possible model overfitting due to an overly complex random-effect structure (see Bates, 2010). Since these warning are associated with problems relating to estimates of power and, consequently, the accuracy of inferential tests, I instead report results from the simpler random-intercept models. Nevertheless, all analyses using the singular-fitted random-slope models yielded the same pattern of results.

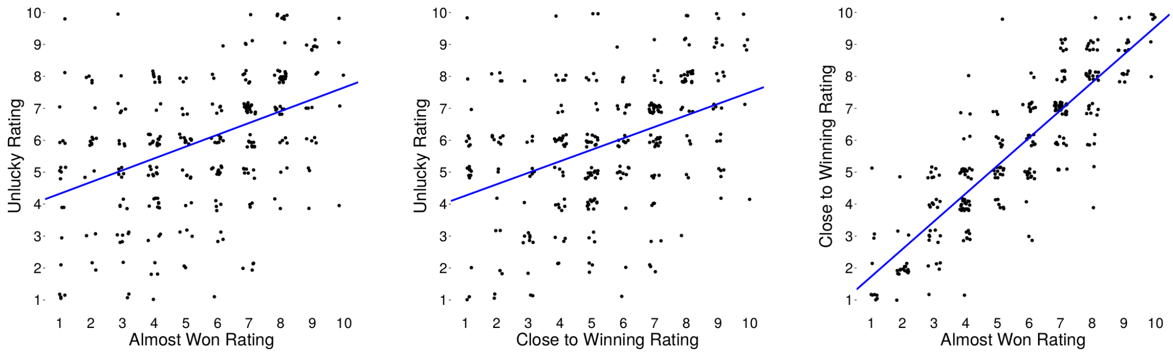


Fig. 2.3 Correlation of assessment ratings in Experiment 1a. Scatterplots depict relationship between *was unlucky* and *almost won* assessments (left), *was unlucky* and *close to winning* assessments (middle), and *close to winning* and *almost won* (right). Regression lines include participants as a random-effect.

lowing analyses were performed using a generalized mixed-effects correlation from the R package "MuMIn" (Barton, 2020). This function calculates a R^2_{GLMM} coefficient that provides a reliable measure of goodness-of-fit for models accounting for variance attributable to random-effects (see Nakagawa & Schielzeth, 2013). Figure 2.3 depicts scatterplots with regression lines for the three correlations examined between *unlucky*, *close to winning*, and *almost won* assessments. While all three models show clear positive correlations between assessment types, the relationship between *close to winning* and *almost won* ratings ($\beta_1 = .87$, $R^2_{GLMM} = .83$) was notably stronger than both the relationship between *was unlucky* and *almost won* ratings ($\beta_1 = .37$, $R^2_{GLMM} = .63$) and *was unlucky* and *close to winning* ratings ($\beta_1 = .36$, $R^2_{GLMM} = .63$).

The slope of the regression line and strong correlation between *almost won* and *close to winning* indicate a close correspondence in the way that these assessments were interpreted. This suggests that modifying telic achievements may be sufficient to elicit counterfactual interpretations of *close*, and that near spatial proximity does not seem to exert any greater effect on ratings compared to *almost*. This was further supported by participants' assessment explanations. Common response included "I regarded the terms *close* and *almost* as practically the same…" and "*Close* and *almost* felt very similar…" I classified explanations as being *almost won/close to winning* equivalent if participants either explicitly stated how these concepts were similar or focused on how luck was somehow distinct. The majority of explanations—approximately 56% (31 out of 55) of the explanations—were *almost won/close to winning* equivalent.

The comparatively higher ratings for *was unlucky* assessments suggest that the concept of luck does not map on directly to properties of the event measured by *almost* and *close*. Comparing intercepts between the three sets of correlations, we see that the regression line's intercept for *unlucky/almost won* ($\beta_0 = 3.95$) and *was unlucky/close to winning* ($\beta_0 = 3.90$) are notably greater than *close to winning/almost won* ($\beta_0 = .85$). The relationship observed between low rated *close to winning/almost won* assessments and *unlucky* assessments may partially explain these results. Examining the scatterplot in Figure 2.3, it appears that *was unlucky* ratings mostly occupy the upper half of the scale, even among items which received low *close to winning/almost won* ratings. Written explanations suggest this may be because of preconceived expectations about what constitutes an unlucky outcome. For example, one explanation read:

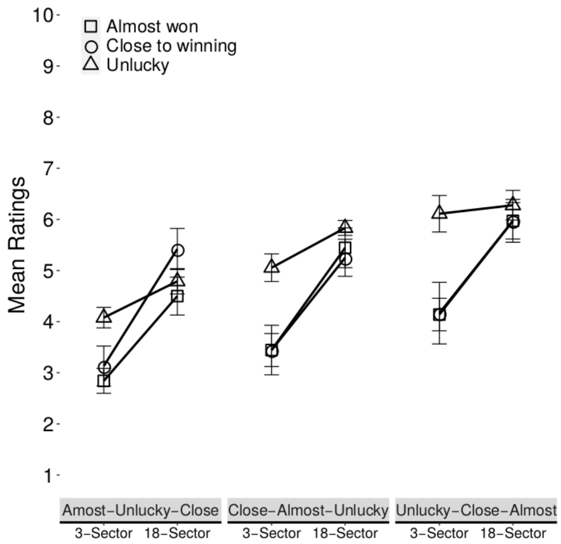
> Close to winning and almost won seem to me to be essentially the same, but unlucky is a little different because probability-wise the three- versus eighteen-part circle were the same, and it's pretty much a binary of being lucky and winning

> or being unlucky and not. Depending on how close the needle was to winning, my answer would change, but in general I felt that if they lost they were unlucky.

The explanation suggests that an important consideration for unluckiness was simply whether the outcome valence was negative. This could explain why higher ratings were observed for *was unlucky* compared to *almost won* and *close to winning* assessments: People perceive losing games of uncertainty as unlucky simply because losing is comparatively unfavorable, regardless of any particular features of the outcome hinting at how things could have happened otherwise. However, as the explanation suggests, other factors, such as spinner-boundary spatial proximity, did make some outcomes appear even *more* unlucky than others. This was reflected in the correlational evidence which showed that approximately 63% of the variance observed in *was unlucky* assessments was attributable to factors which determined *close to winning* and *almost won* assessments. Together, the empirical evidence and descriptive accounts of participants' reasoning provide support in favor of Pritchard and Smith's (2004) proposal that luck is evaluated both according to the outcome's impact on individuals and the perceived attainability of the counterfactual alternatives.

Proximity (Centered Outcomes Only). Teigen (1996) originally observed that luck attributions were greater for eighteen-sector wheels compared to three-sector wheels when the spinner landed in the center of the target sector. To test whether this experiment was able to replicate Teigen's findings, the following analyses only examine centered wheels and, therefore, omit responses for proximate three-sector wheels (i.e., wheels (2.2bd)).

An analysis of assessment blocks revealed that the order in which participants received the different assessments influenced their overall ratings. The left graph in Figure 2.4 shows



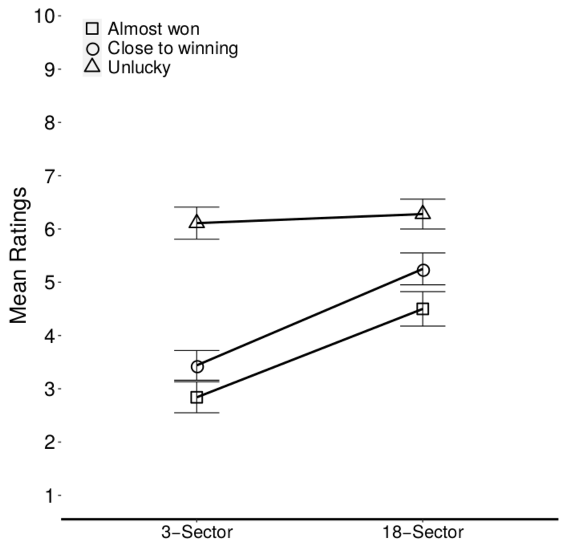


Fig. 2.4 Means assessment ratings for centered outcomes in Experiment 1a. Graphs show mean ratings for *almost won*, *close to winning*, and *was unlucky* assessments for three- and eighteen-sector wheels by assessment block order (left), and by the first assessment within each block (right). Error bars are within-subject standard errors of the mean.

mean assessment ratings for the three- and eighteen-sector wheels by each of the three assessment blocks. A 2 (wheel) × 2 (sector boundary) × 3 (assessment) × 3 (assessment block order) repeated measures ANOVA revealed a significant difference among assessment blocks, $F(2,52) = 3.87$, $\eta_p^2 = .13$, $p < .05$; but with no interactions. Helmert contrasts revealed mean ratings in the *unlucky-close-almost* ordered block ($M = 5.44$, $SD = 1.93$) were significantly greater than mean ratings in the *almost-unlucky-close* ordered block ($M = 4.13$, $SD = 2.45$) and *close-almost-unlucky* ordered block ($M = 4.75$, $SD = 2.33$), $F(1,52) = 6.023$, $\eta_p^2 = .11$, $p < .05$. These findings indicate a possible spillover influence from assessment type. For this reason, the following analyses will only examine data from the first assessment type participants received, thereby treating assessment type as a between-subjects manipulation.

The right graph in Figure 2.4 shows mean ratings for three- and eighteen-sector wheels for the three assessment types. As predicted, a 2 (wheel) × 2 (sector boundary) × 3 (assessment) repeated measures ANOVA revealed that ratings for 18-sector wheels (M = 5.33, SD = 1.85) were significantly greater than three-sector wheels (M = 4.11, SD = 1.86), $F(1,52) = 17.53$, $\eta_p^2 = .25$, $p < .001$. There was also an effect of assessment type, $F(2,52) = 10.67$, $\eta_p^2 = .29$. A post-hoc Tukey HSD test found that, overall, mean *was unlucky* ratings (M = 6.19, SD = 1.91) were greater than mean *almost won* ratings (M = 3.67, SD = 2.38), $t(52) = 4.48$, $p < .001$; and mean *close to winning* ratings (M = 4.35, SD = 2.32), $t(52) = 3.24$, $p < .01$. But there was no difference between *almost won* and *close to winning* ratings, $t(52) = 1.20$, $p = .46$. No effect of sector boundary was observed, $F(1,52) < 1$.

Finally, an interaction was observed between wheel and assessment type, $F(1,52) = 3.20$, $\eta_p^2 = .11$, $p < .05$. The right graph in Figure 2.4 shows that *almost won* and *close to winning* assessments primarily contributed to the effect of wheel, with little change observed for *unlucky* assessments. Indeed, post-hoc simple-effect tests with Bonferroni corrections found that the increase in ratings from three-sector to eighteen-sector wheels was significant for *close to winning* assessments, $t(17) = 3.33$, $d = .81$, $p < .05$; and *almost won* assessments, $t(18) = 2.87$, $d = .69$, $p < .05$; but not *was unlucky* assessments, $t(17) < 1$. Thus, contrary to Teigen's (1996) original findings, the number of sectors did not meaningfully influence luck attributions. Possible explanations of this null finding are addressed in Experiment 1b.

<u>Decisiveness (Proximate Outcomes Only)</u>. The following analyses examine the influence of sector boundary (i.e., whether the spinner departed or is approaching the winning sector). To control for spinner-boundary spatial proximity (i.e., strictly ten-degree spinner-boundary distance), I only examine proximate three-sector wheels and centered eighteen-sector wheels, omitting centered three-sector wheels.

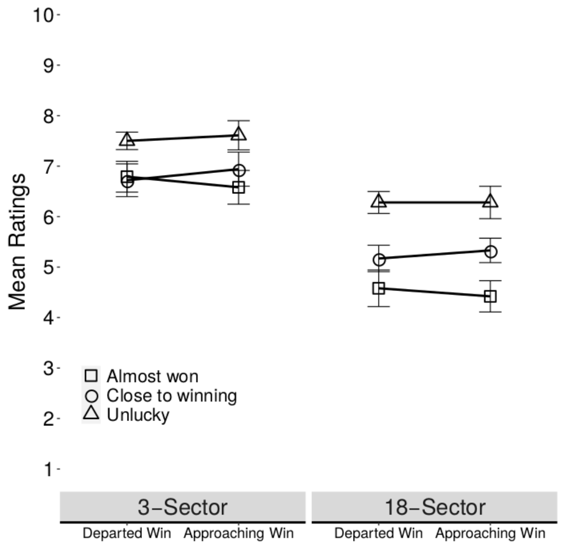


Fig. 2.5 Means assessment ratings for proximate outcomes in Experiment 1a. Graph shows mean ratings for three- and eighteen-sector wheels by outcome type, depicting spinners that departed the winning sector or are approaching the winning sector. Error bars are within-subject standard errors of the mean.

Figure 2.5 shows mean assessment ratings for three- and eighteen-sector wheels outcomes according to the spinner's location relative to the winning sector. Contrary to the prediction that approaching outcomes would elicit greater assessment ratings, a 2 (wheel) × 2 (sector boundary) × 3 (assessment) repeated measures ANOVA found no influence of the sector boundary, $F(1{,}52) < 1$. This suggests later occurring decisiveness (i.e., approaching scenarios) did not contribute to greater assessment ratings. I offer possible explanations of this null effect in the following section.

Lastly, an unexpected difference between wheels emerged, with three-sector wheels receiving greater average assessment ratings ($M = 7.02$, $SD = 1.48$) than eighteen-sector wheels ($M = 5.33$, $SD = 1.60$), $F(1{,}52) = 48.01$, $\eta_p^2 = .48$, $p < .001$. No other effects reached the threshold for significance. That three-sector wheels received greater ratings is noteworthy because it suggests that people do not rely exclusively on the spatial distance separating multiple entities when assessing the counterfactual plausibility of alternative outcomes. While the dissociation of proximity and counterfactual closeness is consistent with the propensity framework, the framework does not explain why participants gave greater closeness assessments for three-sector wheels. Recall that the propensity framework contends that counterfactual closeness beliefs stem from event cues that give the impression of progress towards a particular outcome. However, the only cue that participants may base such beliefs on (i.e., spinner-boundary spatial proximity and the spinner's direction of rotation) cannot explain the difference observed between conditions. An alternative explanation is that counterfactual closeness beliefs further depended on the plausibility of the counterfactual alternative relative to the overall likelihood of the actual outcome. This explanation is further examined in Chapter 4.

2.3 Experiment 1b: The luckiness of winning outcomes and "barely winning"

The findings from Experiment 1a indicate that the spinner's spatial proximity to the winning sector boundary did not appear to influence participants' beliefs about the outcome's unluckiness. While this finding seems to contradict Teigen's (1996) results, methodological distinctions between these experiments might account for the apparent discrepancy. First, Teigen gave participants a forced choice task, where he instructed them to select the luckier outcome given the choice between a three- and eighteen-sector wheel. The ability to directly compare items and the lack of intermediate response options may have contributed to participants focusing on a simple and clear-cut comparative dimension, such as spatial proximity, to help settle outcome luckiness. In contrast, participants would not have needed to settle any such ambiguity in Experiment 1a, making it a potentially less sensitive measure.

Second, it is possible that *unluckiness* and *luckiness* judgments are evaluated differently. The analysis of assessment block order revealed that mean ratings from the luck-first block (see Figure 2.4) were much greater than either of the other blocks. The luck-first block also differed in how participants rated three- and eighteen-sector wheels. While no difference in mean ratings was observed in the luck-first block, post-hoc analyses revealed that ratings for unluckiness were significantly greater for eighteen-sector wheels in the almost-first block ($t(37) = 2.75$, $p < .01$) and close-first block ($t(35) = 2.35$, $p < .05$). The observed effect of order meant these responses were removed from later analyses, but the effect of spatial proximity found within these blocks suggests that counterfactual closeness, while perhaps not a necessary property of unluckiness, is still relevant to its meaning. Consider an example use of *unlucky*. It seems fitting to say *it's unlucky that you were fired* despite not knowing anything about the reason for which the person was fired. If there exists a counterfactual implication in this statement, it is only weakly expressed. However, the counterfactual implication for *lucky* appears much stronger, as illustrated in *it's lucky that you were promoted*. Here the implication is that the speaker understands something about the subject's work circumstances and is, somewhat rudely, claiming they had nearly not received the promotion. While I initially chose unluckiness evaluations because negative events tend to have a stronger association with the activation of counterfactual thought (Roese, 1997), the following experiment eschews such scenarios in favor of luckiness evaluations in order to avoid any possible nuances in *unlucky's* meaning.

Experiment 1b also seeks to address why decisiveness manipulations had no effect on assessments. While it is possible that decisiveness was unimportant for assessment ratings, it may also have been that the manipulation was too weak. Propensity and decisiveness are features of events that are typically experienced over time, as they require the perception of progress and uncertainty. Placing the spinner next to earlier or later sector boundaries, thereby creating the impression of having departed and approaching, was meant to indirectly manipulate beliefs about the spinner's prior movements. However, the sensation of propensity that typically accompanies feelings of counterfactual closeness may strongly depend on the direct observation of progress over the event's course.

Before proceeding to a more direct manipulation of propensity, I introduce two small changes to the prior experimental design that are intended to strengthen the manipulation of decisiveness. While the instructions stated that the spinner always rotated clockwise, four participants expressed confusion about the the spinner's rotation. Therefore, to make Experiment 1b's instructions more clear, arrows pointing in a clockwise direction are placed over the example wheels. The written text accompanying this image is also bolded for emphasis, stating that the spinner always spins clockwise. Furthermore, additional instructions are added to encourage mental visualization, asking participants to "imagine the spinner as it traveled around the wheel to eventually land in the depicted location."

Finally, while assessments for the two proximatives *almost* and *close* appeared to share highly similar interpretations, other commonly used proximatives may elicit a different pattern of responses. Section 1.2.1 addressed how some proximatives, like *barely*, have unique lexical properties which distinguish it from *almost*. First, the polar component of *barely* entails that the object or event it modifies is a true proposition. Thus, *she barely won* means that a win occurred but was close to having not occurred. Second, *barely* is a NPI licenser, which means that it allows for words and phrases which have a distinctly negative evaluative stance (e.g., *I [barely/?almost]* slept a wink) (Horn, 2011). As per Amaral's (2007) analysis, I hypothesize that *barely's* NPI licensing properties are a consequence of perceived insufficiencies in the outcome relating to the proximal component. I hypothesize that this property of *barely* will have consequences for the kinds of events the adverb is seen as appropriately describing. An ideal *barely* event is one for which perceived forces work against the propensity-suggested outcome, but are ultimately unsuccessful in preventing that outcome. If such forces are an important feature of *barely* events, then we should see this reflected in how *barely* assessments are rated in departed and approaching scenarios. *Barely won* should be more appropriate when the spinner has just departed the losing sector, since the force of friction is perceived to work against the spinner's tendency towards winning. However, antagonistic forces are not as easily evoked when the spinner stops short of crossing into a losing sector, as the spinner's own tendency is to continue further along its path into this sector. Since antagonistic forces do in fact counteract the spinner's tendency, *barely* descriptions should be perceived as less appropriate.

2.3.1 Method

Participants. The following analyses include 45 Northwestern University undergraduates. Data from 9 participants were removed because they failed a comprehension check.

Design & Materials. Experiment 1b uses approximately the same 2 (wheel: three- and eighteen-sector wheel) × 2 (sector boundary: Departed and approaching) × 2 (spinner location: Centered and proximate) × 3 (assessment: *Almost lost*, *barely won*, and *lucky*) within-subjects design. The details are identical to Experiment 1a with the following exceptions. First, target items now depict winning outcomes, with assessments replaced with "This player [*almost lost* / *barely won* / *was lucky*]. Correspondingly, participants are given the option of responding "THIS PLAYER LOST."for the filler losing outcomes. Second, while the factor *sector boundary* still varies by departed and approaching scenarios, the change to winning outcomes means that this manipulation is no longer able to distinguish between early and late decisiveness among centered eighteen-sector wheels. Since these outcome will necessarily be flanked by sectors that produce counterfactual results (i.e., losses), it is no longer possible to manipulate decisiveness by changing the relative location of these outcomes. Therefore, sector boundary is only analyzed relative to proximate three-sector wheels. And lastly, assessments for *close* were replaced with *barely*. These items read "This player barely won."

Procedure. The procedure is identical to Experiment 1a.

2.3.2 Results and Discussion

Experiment 1b makes the same three predictions as Experiment 1a, with the addition that *barely won* assessments will receive higher ratings in departed scenarios compared to approaching scenarios. The overall results were broadly consistent with these predictions. Assessment measures *was lucky*, *almost lost*, and *barely won* all correlated with one another, with the relationship between *almost lost* and *barely won* being particularly stronger than either assessment's relationship with *was lucky*. Furthermore, among centered outcomes, an aggregate measure of the three assessments showed that three-sector wheels received higher ratings compared to eighteen-sector wheels. This difference was also observed for luck assessments, thus replicating Teigen's (1996) original finding where Experiment 1a was unable. However, the manipulation of sector boundary among proximate three-sector wheels once again showed no effect across any of the three assessments, thus suggesting decisiveness was unimportant for judgments. In sum, the results of Experiment 1b were similar to those of Experiment 1a: While the difference between losing and winning outcomes was likely responsible for the earlier failure to replicate Teigen, the overall results were again unable to conclusively demonstrate that decisiveness influenced judgments.

Assessment Correlations. Linear mixed-effects models were used to assess differences between assessment types (i.e., *was lucky*, *barely won*, *almost lost*). The analyses used random-intercept models treating assessment type as a fixed-effect, and participants and an assessment-by-participants interaction as random effects. The analysis revealed a significant effect of assessment type, $\chi^2(2) = 74.412$, $p < .001$. Additional comparisons showed that, overall, average *was lucky* ratings ($M = 6.16$, $SD = 2.68$) were greater than both *almost lost* ratings ($M = 4.17$, $SD = 3.25$), $\chi^2(1) = 39.54$, $p < .001$; and *barely won* ratings ($M = 3.91$, $SD = 3.35$), $\chi^2(1) = 48.09$, $p < .001$. However, there was no difference between *almost lost* and *barely won* ratings, $\chi^2(1) = 1.92$, $p = .27$.

Figure 2.6 shows scatterplots with regression lines for each of the three correlations examining assessment relationships. As predicted all three assessments are positively corre-

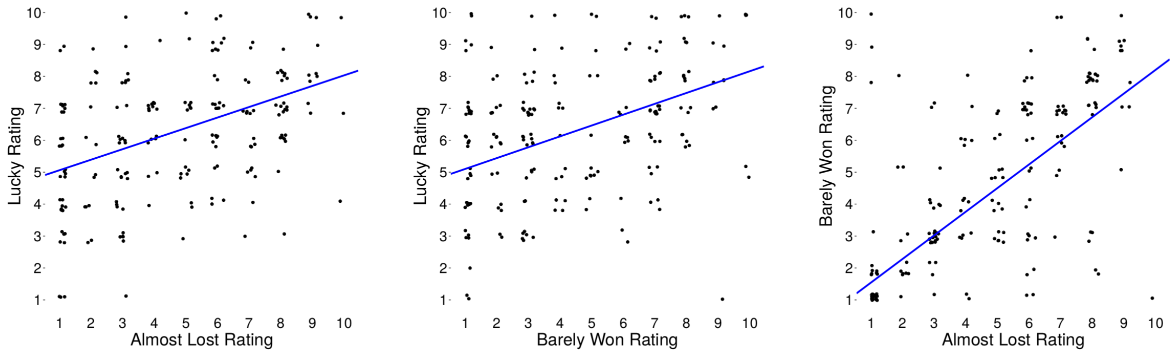


Fig. 2.6 Correlation of assessment ratings in Experiment 1b. Scatterplots depict relationship between *lucky* and *almost lost* assessments (left), *lucky* and *barely won* assessments (middle), and *almost lost* and *barely won* (right). Regression lines include participants as a random-effect.

lated with one another. However, the relationship between *almost lost* and *barely won* ratings (β_1 = .74, R^2_{GLMM} = .58) was stronger than both the relationship between *was lucky* and *almost lost* (β_1 = .34, R^2_{GLMM} = .41) and *was lucky* and *barely won* (β_1 = .33, R^2_{GLMM} = .41).

The close correspondence between *almost lost* and *barely won* ratings suggests these assessments are interpreted similarly. A number of written explanations from participants expressed this similarity. I classified explanations as being *almost lost/barely won* equivalent if participants either mentioned how the concepts were similar or explained how luck was somehow distinct. About half of the responses—approximately 49% of explanations (22 out of 45)—were *almost lost/barely won* equivalent.

Similar to the previous findings, the regression intercepts for *was lucky/almost lost* assessments (β_0 = 4.73) and *was lucky/barely won* assessments (β_0 = 4.87) were notably greater than *almost lost/barely won* assessments (β_0 = .81). The pattern of responses seen in Figure 2.6 suggests a general tendency to rate *was lucky* assessments higher than their *almost* *lost* and *barely won* counterparts. Examining participants' written explanations revealed that, much like the relationship between losing and unluckiness observed in Experiment 1a, participants believed that simply winning was sufficient to consider the outcome lucky. However, properties of the outcome nevertheless influenced participants' ratings, as approximately 41% of the variance observed in *was lucky* assessments was attributable to factors determining *almost lost* and *barely won* assessments. Therefore, consistent with Pritchard and Smith's (2004) account, outcome valence and properties which indicate counterfactual attainability both appeared to influence judgments.

Proximity (Centered Outcomes Only). The analyses in this section examine only centered wheels. I therefore omit responses for proximate three-sector wheels (i.e., wheels (2.2bd)).

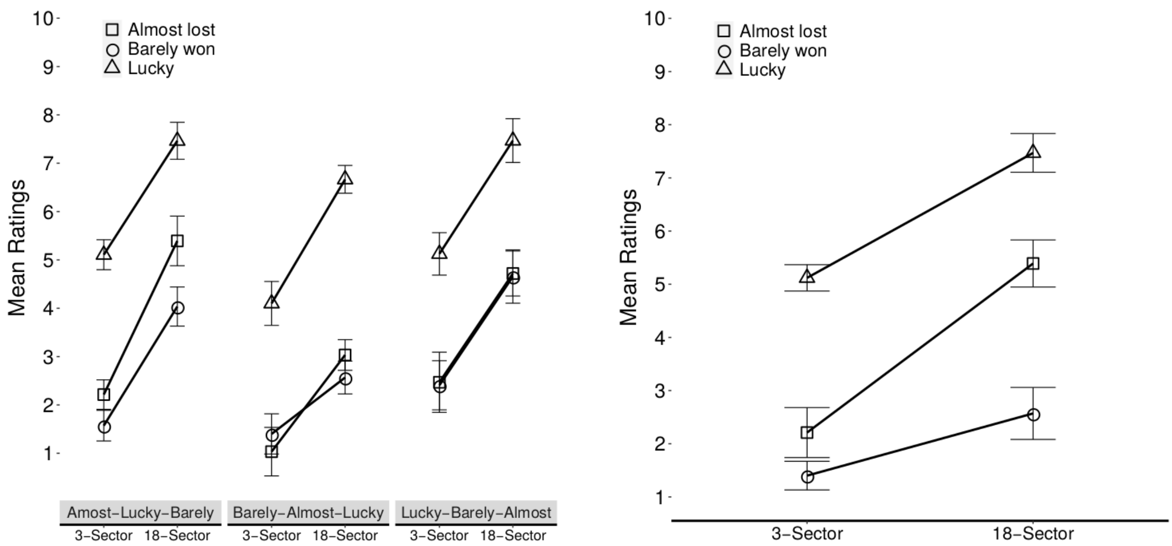


Fig. 2.7 Mean assessment ratings for centered outcomes in Experiment 1b. Graphs show mean ratings for *almost lost*, *barely won*, and *was lucky* assessments for three- and eighteen-sector wheels by assessment block order (left), and by the first assessment within each block (right). Error bars are within-subject standard errors of the mean.

A 2 (wheel) × 2 (sector boundary) × 3 (assessment) × 3 (assessment block order) repeated measures ANOVA revealed a significant effect of block order, $F(2,42) = 7.33$, $\eta_p^2 = .26$, $p < .05$. No interactions with block order were observed. Helmert contrasts revealed that the average mean ratings of the *almost-lucky-barely* ordered block ($M = 4.30$, $SD = 3.39$) and *lucky-barely-almost* ordered block ($M = 4.47$, $SD = 2.76$) were greater than the mean rating for the *barely-almost-lucky* ordered block ($M = 3.31$, $SD = 3.04$), $F(1,42) = 14.44$, $\eta_p^2 = .25$, $p < .001$. The left graph of Figure 2.6 shows mean assessment ratings for the three- and eighteen-sector wheels for each of the three assessment blocks. Due to the influence of earlier assessments on later assessments, the following analyses only examine ratings from the first assessment participants received, thereby treating assessment type as a between-subjects manipulation.

The right graph in Figure 2.7 shows mean ratings for three- and eighteen-sector wheels for the three assessment types. A 2 (wheel) × 2 (sector boundary) × 3 (assessment) repeated measures ANOVA revealed that ratings for 18-sector wheels ($M = 5.19$, $SD = 2.96$) were significantly greater than 3-sector wheels ($M = 2.98$, $SD = 2.17$), $F(1,42) = 44.43$, $\eta_p^2 = .51$, $p < .001$. There was also an effect of assessment type, $F(2,52) = 33.71$, $\eta_p^2 = .62$, $p < .001$. A post-hoc Tukey HSD test further revealed that mean *was lucky* ratings ($M = 6.30$, $SD = .98$) were greater than both mean *almost lost* ratings ($M = 3.80$, $SD = 1.94$), $t(42) = 2.49$, $p < .001$; and mean *barely won* ratings ($M = 1.98$, $SD = 1.41$), $t(42) = 4.31$, $p < .001$. And mean *almost lost* ratings were greater than mean *barely won* ratings, $t(42) = 3.33$, $p < .01$. No effect of sector boundary was observed $F(1,42) < 1$.

Lastly, the prior analysis found a marginal interaction between wheel and assessment type, $F(2{,}42) = 3.00$, $\eta_p^2 = .13$, $p = .06$. This is likely the result of a comparatively smaller increase in mean *barely won* ratings from three- to eighteen-sector wheels. Post-hoc simple-effect tests with Bonferroni corrections found that this difference failed to reach significance $t(14) = 1.61$, $p = .39$. However, there was a significant difference between three- and eighteen-sector wheels for *almost lost* ratings, $t(13) = 6.41$, $d = 1.71$; and, most importantly, *was lucky* ratings, $t(15) = 4.99$, $d = 1.25$ $p < .001$. The observation of greater *lucky* ratings in the eighteen-sector wheels is important because it replicates Teigen's (1996) original finding that greater spatial proximity contribute to greater perceived luck.

Decisiveness (Proximate Outcomes Only). The following analyses examine the influence of sector boundary on attributions, examining scenarios in which the spinner has either just departed from a losing sector or approaches, but does not enter into, a losing sector. As before, I only examine proximate wheels (i.e., spinner-boundary separation of ten-degrees). However, since the losing sector's location for centered eighteen-sector wheels does not allow for the distinction between departed and approaching (i.e., the spinner both departed and is approaching the losing-sector in equal measure), I only examine responses for proximate three-sector wheels. The following analysis therefore omits responses for both three- and eighteen-sector wheels with centered outcomes.

Figure 2.8 shows mean assessment ratings for three-sector wheels according to the spinner's location relative to the losing sector. A 2 (sector boundary) × 3 (assessment) repeated measures ANOVA found no influence of sector boundary, $F(1{,}42) < 1$; no effect of assessment $F(2{,}42) = 2.25$, $p = .12$; and no boundary-by-assessment interaction, $F(2{,}42) = 2.13$, p

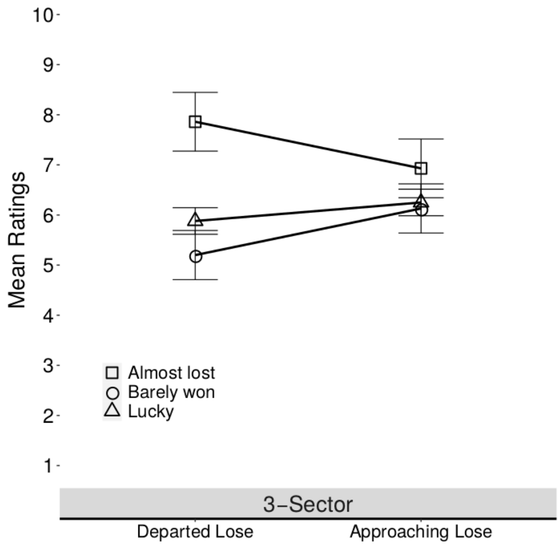


Fig. 2.8 Mean assessment ratings for proximate outcomes in Experiment 1b. Graph shows mean ratings for three-sector wheels by outcome type, depicting spinners that departed the losing sector or are approaching the losing sector. Error bars are within-subject standard errors of the mean.

= .13. This finding runs contrary to the earlier prediction that more decisive outcomes (i.e., departed losing sector scenarios) would contribute to higher assessment ratings. In fact, ratings were trending in the opposite direction, suggesting that participants may have felt approaching scenarios (M = 6.13, SD = 3.14) were better suited for *barely won* descriptions than departed scenarios (M = 5.20, SD = 3.63). However, given the complete absence of an effect for sector boundary manipulations in both Experiments 1a and 1b, a more reasonable explanation may be that static images of wheel-and-spinner outcomes were unable to elicit the types of event sequence considerations that are entailed by such manipulations. As such, Experiments 2a and 2b implement more direct manipulations of propensity and decisiveness by animating the entire wheel-and-spinner event sequence.

Lastly, 3-sector wheels elicited greater assessment ratings than 18-sector wheels. Collapsing across sector boundary, a 2 (wheel) × 3 (assessment) ANOVA revealed once again higher aggregate assessment ratings for three-sector wheels ($M = 6.34$, $SD = 2.35$) compared to eighteen-sector wheels ($M = 5.19$,$SD = 2.96$), $F(1,42) = 7.75$, $\eta_p^2 = .16$, $p < .01$. This same effect was observed in Experiment 1a, with proximate three-sector wheels receiving higher aggregate assessment ratings compared to proximate eighteen-sector wheels. The results again suggest that spatial relativity is influencing participants' ratings: The counterfactual plausibility of some alternative is informed by the overall margin of error that the actual outcome permits. Since each individual sector takes up a larger proportion of the total area of the three-sector wheels, there exists a substantial degree of freedom in where the spinner could have landed within a given sector. The close spatial proximity of the spinner to its neighboring sector therefore creates an even stronger impression of counterfactual proximity since participants recognize that that sector accommodates a wide range of spinner configurations within its bounds. The question of how relativity factors into people's CCF judgments is further examined in Chapter 4.

2.4 Experiment 2a: Animated losing: "Almost won" and "barely lost"

According to the propensity framework, CCF judgments depend on the perception of progress towards a particular outcome (Kahneman & Varey, 1990). If properties of the event create an impression of progress, and that progress is sustained late into the event, then reasoners will have justification for believing that the outcome will eventuate. However, when a reasoner's expectations are violated by the outcome's eventual non-occurrence, the earlier evidence of progress will serve as evidence that the alternative could have easily happened and, therefore, that it almost happened.