Check for updates

aps ASSOCIATION FOR PSYCHOLOGICAL SCIENCE

Current Directions in Psychological Science
2020, Vol. 29(1) 55–62
© The Author(s) 2019
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/0963721419884313
www.psychologicalscience.org/CDPS
SAGE

When Small Signs of Change Add Up: The Psychology of Tipping Points

Ed O'Brien
Booth School of Business, The University of Chicago

Abstract

Things change, but the exact point at which they do is often unknown. After how many loveless nights is a relationship "officially" in trouble? After how many happy days has one's depression "officially" passed? When do recurring patterns in the climate or economy "officially" warrant a response? When is a person's identity "officially" accepted? Everyday fluctuations in oneself and the social world create ambiguities about when people will diagnose lasting, qualitative change (and therefore act). Recent research documents these *tipping points of change* as a psychological process, shaped by individual and situational forces. People judge tipping points asymmetrically across valence and asymmetrically across time. Here, I review discoveries and outline future directions in tipping-points research.

Keywords

tipping points, change perception, self/others over time, evaluative judgment, qualitative and categorical shifts

Things change, but the exact point at which they do is often unknown. After how many loveless nights is a relationship "officially" in trouble? After how many happy days has one's depression "officially" passed? When do recurring patterns in the climate or economy "officially" warrant a response? When is a person's identity "officially" accepted?

Many important changes emerge from the accumulation of smaller fluctuations, but smaller fluctuations create ambiguity about when they reflect lasting signal versus passing noise. Recent research has begun to discover that these *tipping points* are a psychological process: People subjectively diagnose tipping points (as opposed to passively responding to objective reality), a process that is shaped by individual and situational forces. In this article, I highlight tipping points as a topic of broad interest, review recent discoveries, and encourage future research. Because tipping points imply points of action—when people will finally intervene (or give up, hitting "the point of no return")—understanding what affects these diagnoses bears on many persuasive and motivational contexts and, more broadly, on navigating everyday change.

What Are Tipping Points?

Operationalization

Philosophers have long debated the *paradox of the heap* (Fisher, 2000). One grain of sand is not a heap; add a second, still no heap; add a third, still no heap; and repeat, ad infinitum—yet heaps of sand exist. In psychology, the tipping point is "the point at which people begin to perceive noise as signal" (O'Brien & Klein, 2017, p. 161)—the first point across identical observations when people conclude that a pattern is no longer an anomaly. Tipping-points research emphasizes top-down effects in perceiving the self and social world. A few bad grades may be dismissed as a new student adjusts to the course, but the nth bad grade shifts one's attribution from the situation to the person ("This must be a bad student after all"). That n is one's tipping point. The notion that people update impressions on the basis of new information is, of course, foundational to research on learning and decisions under uncertainty. Tipping points highlight an emerging branch from this tradition, focusing on people's thresholds for diagnosing qualitative shifts in assessing themselves and others—when X becomes Y. Tipping-points research asks, "When and why do people cross these thresholds?" And, in particular, "Do they do so on the basis of small signs of change that may compound ambiguously?"

Such questions build on various literatures. Research on perceiving streaks (Ayton & Fischer, 2004; Gilovich,

Corresponding Author:
Ed O'Brien, The University of Chicago Booth School of Business, 5807 South Woodlawn Ave., Chicago, IL 60637
E-mail: eob@chicagobooth.edu

Vallone, & Tversky, 1985) seems relevant, as the "hot-hand fallacy" (assuming a streak will stay) and "gambler's fallacy" (assuming a streak will reverse) both imply tipping points of sorts. These findings center on randomness intuitions: Hot-hand effects tend to emerge when people can devise theories for why a streak could defy randomness (e.g., players gaining confidence); the gambler's fallacy tends to emerge when people cannot (e.g., coin flipping). Tipping-points research tests contexts in which all participants can devise theories about change, as tipping points are inherently subjective. Any differences are therefore especially informative.

Regarding more direct change-perception literatures, one common approach views change perception as a function of attention. People can respond only to changes that they notice, but competing demands on attention make noticing difficult (e.g., the "boiling frog"; Simons & Ambinder, 2005). However, such paradigms do not account for how people perceive change when attention is explicitly held constant or when change occurs in richer social domains (e.g., assessing a full set of dates to diagnose the state of a relationship). Another approach views change perception as a function of lay theories of how things *should* change (Dweck, 2008; O'Brien & Kardas, 2016; Ross, 1989; Wilson & Ross, 2001), but such paradigms assess absolute comparisons to the past or future rather than real-time reactions: Participants judge something as different (e.g., "I'm more outgoing today than last year"), leaving little insight into the dynamic nature of when such changes first emerge and start to matter. In a third literature on impression updating, researchers test how new social information at Time 2 changes Time 1 impressions (e.g., learning that a favorite actor commits bad deeds; Ferguson, Mann, Cone, & Shen, 2019; Fiske, 1980; Reeder & Coovert, 1986; Skowronski & Carlston, 1989). To the extent that this research compares evaluations before and after one big reveal, more research is needed to assess when such changes first emerge and start to matter (such as by assessing the compounding effect of smaller fluctuations). Moreover, this research does not necessarily define "updating" as crossing category thresholds. Tipping-points research helps fill these gaps.

Measurement

My laboratory has documented tipping points in more than 30 experiments testing more than 5,000 participants (Klein & O'Brien, 2016, 2018, 2019; O'Brien & Klein, 2017). The primary paradigm is behavioral: Participants view piecemeal evidence for a pattern and after each piece choose whether to observe the next piece or to end their search. Their task is to stop whenever they have seen enough for their impressions to tip. For example, participants might read about a student, see the grade that student got on the first assignment ("Assignment 1/10: Grade C+"), and choose either to tip at that point ("I've seen enough to conclude this is a bad student") or to see Assignment 2. If they choose the latter, "Assignment 2/10: Grade C+" is displayed, and so on until they choose to tip. In other paradigms, participants are asked to report their tipping points before any evidence unfolds (e.g., "How many consecutive C+s, of the next 10 assignments, would lead you to conclude this is a bad student?") or work backward by assessing responses to a given streak (e.g., participants win two games then bet on winning again).

Across paradigms, some studies test impression formation: Participants begin at a neutral point (knowing nothing about the target) and learn initial evidence ("Assignment #1: C+"), and then this evidence repeats. Other studies test impression change: Participants begin at a known point ("This is a good student") and learn initial evidence ("Assignment #1: C+"), and then this evidence repeats. Both approaches test the same question of interest for tipping-points research (e.g., how participants respond to the same streak of C+s), but there may be other interesting main effects between questions of exploration (e.g., "What kind of student is this?") and questions of verification (e.g., "Is this a good student?").

Discoveries So Far

Generality

Our experiments so far have produced notably few main effects. For example, the speed of diagnosing "official" decline does not depend on measuring change as frequency (e.g., number of bad grades), duration (e.g., calendar time with bad grades), or magnitude (e.g., percentage drop in grades). Nor have we yet found systematic differences across domain, direction (e.g., additive vs. subtractive change), or target (e.g., self vs. other change). We have, however, uncovered two robust differences.

Tipping points are asymmetric across valence

People reach their tipping point more quickly when they are evaluating evidence for possible decline than when they are evaluating evidence for possible improvement, despite identical evidential quality; a handful of poor grades, bad games, and gained pounds leads people to diagnose intellect, athleticism, and health as "officially" changed, yet corresponding positive signs are dismissed as fickle flukes. O'Brien and Klein's (2017) Study 2 offers a vivid demonstration of this effect (see

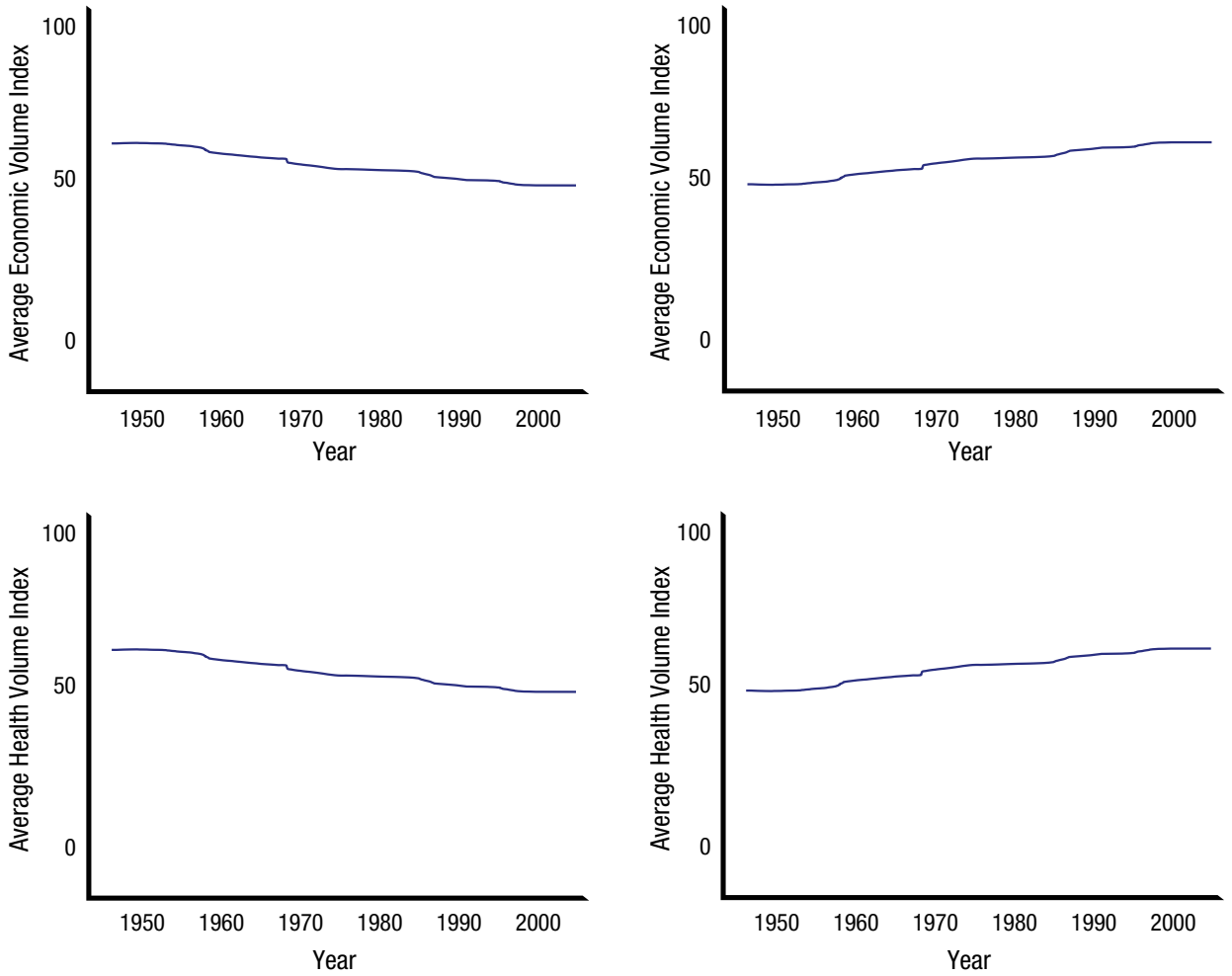

Fig. 1. Stimuli from O'Brien and Klein's (2017) Study 2. Participants viewed one of these charts, each of which showed an ambiguous trajectory of societal change in the average economic volume index or average health volume index. The chart was described as either potentially good (e.g., "Lower values suggest things are getting worse") or potentially bad (e.g., "Lower values suggest things are getting better"). The exact same evidence of change was perceived as more "real" merely when framed as potential decline versus potential improvement.

Fig. 1). Participants viewed a chart showing an ambiguous trajectory of societal change on a fictional metric, which we presented as real so participants did not simply recruit facts for interpretation. They rated how much the chart showed "real" diagnostic change (e.g., "This is showing a clear trend rather than just noise"). The chart was described as potentially good (e.g., "Lower values suggest things are getting worse") or bad (e.g., "Lower values suggest things are getting better"). For generalizability, we also manipulated domain (change in the economy vs. change in public health) and slope (upward vs. downward). One factor mattered: valence. Participants were more convinced of change when charts were framed as possibly indicating decline, despite staring at identical evidence.

In O'Brien and Klein's (2017) Study 3, participants played a card-flipping game and could bet money on a third consecutive loss or a third consecutive win (see Fig. 2). Despite never providing objective chances for either outcome, more participants bet on a third loss (74% placed a bet, citing reasons such as "Even though it is randomized I thought I would go with my gut. I mean . . . it was a loss the first two") than a third win (50% did not place a bet, citing reasons such as "I thought since 2/3 were wins, the next would probably not. Even though I know each time you have 50/50 chance").

People think good luck runs out but bad luck lasts—using the same logic! This contradiction highlights the role of hedging (tipping in response to losses "just in case") from what O'Brien and Klein (2017) call *entropy beliefs*: People may view good entities as truly more capable of losing positive qualities (floors seem universal) than bad entities seem capable of gaining them (ceilings seem selective), reflecting a truth gleaned from everyday experience. Basketball fans know for sure that almost no struggling amateur will make the National

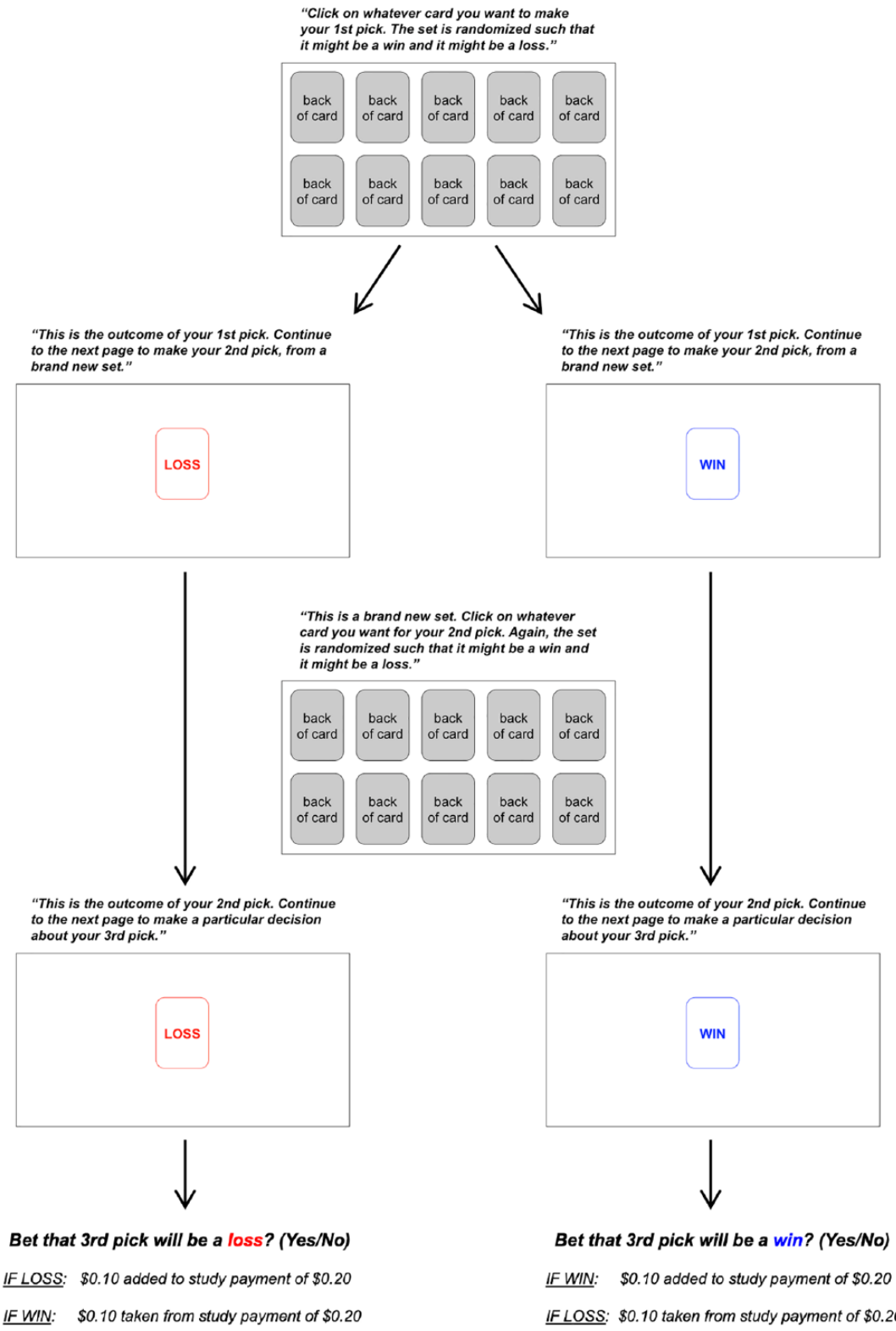

Fig. 2. Design of O'Brien and Klein's (2017) Study 3. Participants flipped three cards one at a time. They always lost twice or won twice on the first two flips, and then they could bet on whether a third flip would continue the streak. More participants who had lost the first two flips took the bet, even though we provided no objective information about differential chances.

Basketball Association (NBA) but that all NBA players will eventually deteriorate. People hold the overgeneralized expectation that, indeed, nothing "gold" can stay, and these entropy beliefs may matter most: The asymmetry flips when people observe initial failure on skills that most people will cultivate, despite failure remaining costly (e.g., professors might give little weight to students' minor stumbles when they know the material will eventually click and instead be very quick to detect change when students first show signs of "getting it").

Tipping points are asymmetric across time

People predict slower tipping points than they express, regardless of valence (Klein & O'Brien, 2018, 2019); people predict that they will patiently observe many pieces of evidence before their impressions "officially" form and change, yet in actuality they make up their minds right away. Figure 3 shows the effect across three of our studies.

In one study, participants evaluated novel artwork. First, participants viewed one painting at random so they generally knew what to imagine. Then, some predicted the total number they would view before making up their minds about liking or disliking the style, whereas others viewed paintings one by one and stopped whenever they made up their minds. Participants predicted they would view more paintings than they viewed in reality. Another study replicated this effect in the number of sips taken to make up one's mind about a novel juice. In another study, students working on their master of business administration (MBA) completed a hypothetical job application in which they wrote bullet-point essays of past management experiences—the exact number they thought would lead hiring managers to hit a tipping point ("I've read enough to know this is a qualified applicant"). The explicit rule of the study was that writing too few *or* too many would cost them the job. Nonetheless, students wrote more than hiring managers read.

Why? Judging tipping points means sorting fact from fiction, so the more convincing the evidence, the quicker people should be to tip. Just as trajectories of decline generally seem more plausible than trajectories of improvement (and thus, people are quicker to tip in response to a decline), evidence experienced in real

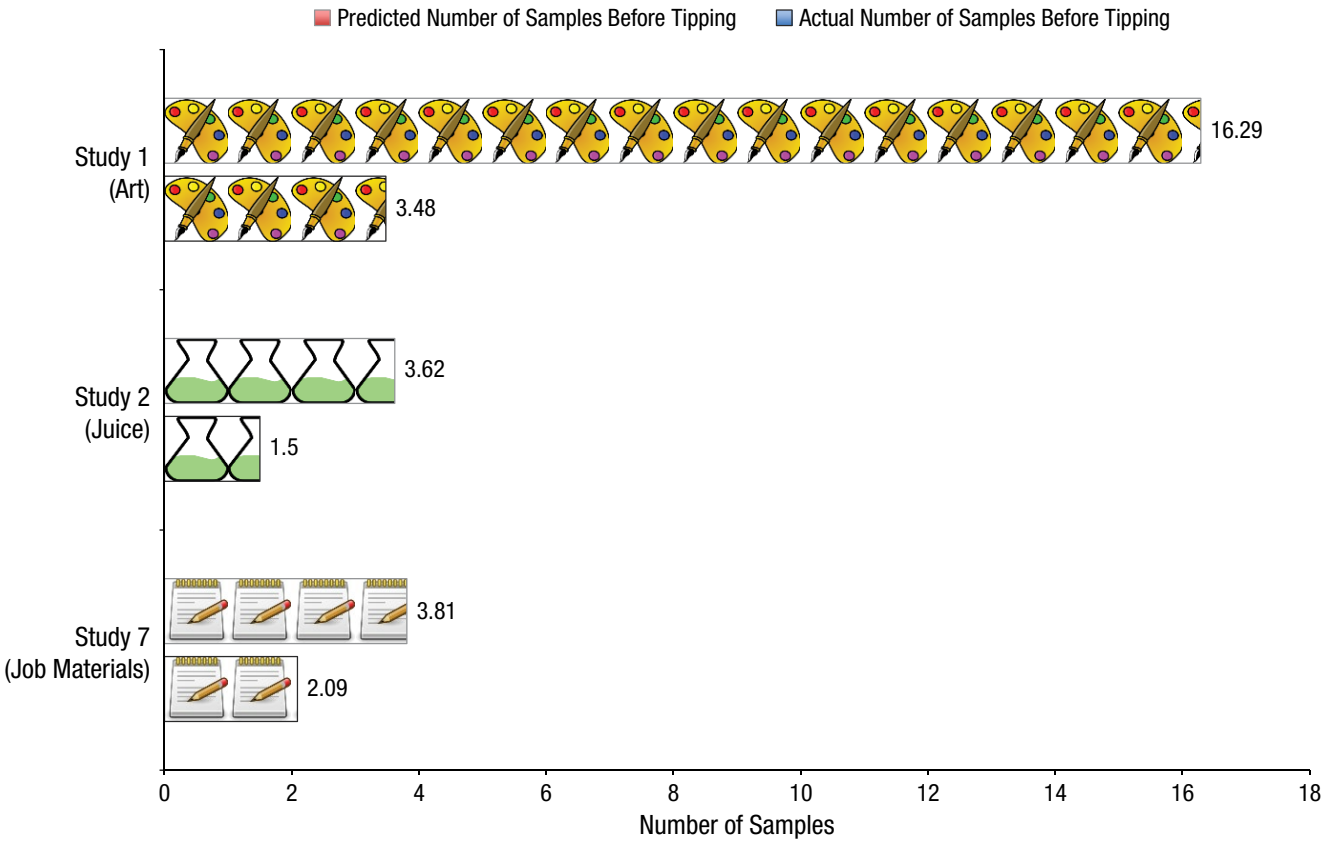

Fig. 3. Mean tipping points in Studies 1, 2, and 7 from Klein and O'Brien (2018). Participants estimated how many samples they or others would need to experience before impressions "officially" tipped. Participants viewed paintings of the same style of art (Study 1), took sips of the same juice (Study 2), or evaluated essays written by the same applicant as part of a job application (Study 7). For all stimuli, impressions tipped on the basis of fewer pieces of evidence than participants had anticipated.

time is simply a more powerful source of information than thinking about evidence in the abstract, since such evidence is replete with emotions, physiological sensations, and other salient forces that grab attention in ways imagination cannot fully capture (O'Brien, 2019; O'Brien & Roney, 2017). In turn, the salience of evidence tends to make evidence more convincing (Tversky & Koehler, 1994). Throughout our studies, participants underappreciated the fact that they and others would react to evidence right from the first piece, rather than play a patient arbiter. Reflecting on real-time reactions hastened participants' predicted tipping points, increasing accuracy.

Road Map for Future Research

Downstream behavior

Tipping points imply points when people become more likely to intervene or surrender. Future research should scale to higher-stakes contexts (e.g., changes in health, climate change action, decisions to change jobs or partners). The valence asymmetry suggests uphill battles for appreciating improvement. The temporal asymmetry suggests conflict between parties who experience evidence from different perspectives (e.g., policymakers may predetermine thresholds for reward or punishment that notoriously prove too high for constituents, who demand action at the first salient strike). Indeed, naive realism in change perceptions may stir conflict over identical evidence (Campbell, O'Brien, Van Boven, Schwarz, & Ubel, 2014). Other research should assess intrapersonal costs (e.g., consumers may overpay for lengthy product trials, assuming they will evaluate more than they actually will before drawing conclusions).

Motivated and nonmotivated mechanisms

If the basic process underlying tipping points is responding to evidence salience, there must be motivated sources of salience that interact with tipping points. Alcoholics may view themselves as more "cured" after their first week of sobriety than friends view them, CEOs may quickly view increases in revenue as signals whereas investors view them as noise, voters may dismiss a few days of poor stock returns or rising unemployment if they support the incumbent administration, and a person who goes on one date with an attractive partner may conclude that he or she is "the one." More research should unpack potential self/other differences, as agents of change likely want to diagnose change. However, this may also reflect nonmotivated differences in accessibility (Klein & O'Brien, 2017; O'Brien, 2013). Only the alcoholic actor knows how effortful that first week felt; he or she actually has a more diagnostic signal. Differences across explicit and implicit change perceptions (Ferguson et al., 2019) may be more informative.

Other boundaries

Beyond self/other differences, testing still other factors that reverse the asymmetries is critical. When do people tip more quickly in response to improvement? Future research should assess additional domain differences (e.g., changes in identity-central features; Strohminger & Nichols, 2014) and individual differences (e.g., trait optimists may flip the valence asymmetry, assuming they reject entropy beliefs). When do people tip more slowly than they think? Extremely emotional events are often rationalized in ways hidden to intuition (Wilson & Gilbert, 2005), and thus may flip the temporal asymmetry; people may assume one horrible fight will forever render a friend a foe, but in reality, friends work to stay friends. For complex stimuli, reacting quickly to initial evidence may itself be mistaken; one may assume that a single reading of a book was enough to form a conclusion, but in reality, rereads may continually reveal new interpretations (Kardas & O'Brien, 2018; O'Brien, 2019). Regardless, the phenomenon appears not easily intuited; future research should assess other ways in which expectations diverge from experiences.

Evidence presentation

Future research should introduce more variance into observations. Variance likely will not affect asymmetries across conditions if it is similarly distributed (e.g., random draws of grades that slowly transition to C+s vs. A+s at equal rates), but extreme draws likely matter; one big shock may disrupt small compounding change. Future research should also integrate the full time course of tipping points. As retrospection and prospection rely on shared lay beliefs (O'Brien, Ellsworth, & Schwarz, 2012; Schwarz, 2012), the temporal asymmetry may stubbornly persist when looking back; people may predict being patient, then quickly make up their minds, yet then later recall being just as patient as imagined. However, other stereotypes about past and future selves (such as past selves seeming emotional and future selves seeming rational: O'Brien, 2015) may interact with tipping-point perceptions over time.

External benchmarks

Some changes are truly instantiated, which can be misperceived because of other attentional demands (Simons & Ambinder, 2005), miscalibrated beliefs

(Davidai & Gilovich, 2015; Ross, 1989), and shifting reference points (Levari et al., 2018). An open question is whether tipping-point thresholds can be objectively quantified. Misperceiving genuine tipping points would bear on many real-world outcomes, from doctors who must anticipate when illnesses will manifest to investors who must anticipate when bear markets will return. One could gain traction on this question by comparing perceptions to other benchmarks, such as normative thresholds (e.g., feverish people may think their temperature has crossed 100.4° F before it does) and mathematical probabilities (e.g., testing how quickly people believe drawn outcomes have shifted from pool A to pool B against Bayesian standards; Massey & Wu, 2005). More research is needed, from all approaches, on categorical change perception in the self and others.

A broad study of tipping points is promising. The point when things change may be fiction, but hopefully this article encourages initial change toward these exciting directions.

Recommended Reading

Bechler, C. J., Tormala, Z. L., & Rucker, D. D. (2019). Perceiving attitude change: How qualitative shifts augment change perception. *Journal of Experimental Social Psychology, 82*, 160–175. A recent example of research that echoes my call for more work on categorical change perception in the self and others outside the tipping-points context.

Klein, N., & O'Brien, E. (2016). (See References). An application of tipping points in the moral domain, highlighting implications of general interest for reward and punishment.

Klein, N., & O'Brien, E. (2018). (See References). Introduces temporal asymmetry and reviews the concept of tipping points over time in greater detail.

O'Brien, E., & Klein, N. (2017). (See References). A recommended starting place that provides a comprehensive, highly accessible introduction to tipping points; reviews how tipping-points research compares and contrasts with other literatures on change perception; and introduces valence asymmetry.

Quoidbach, J., Gilbert, D. T., & Wilson, T. D. (2013). The end of history illusion. *Science, 339*, 96–98. An accessible example of still other change-perception research; conceptually hints at downstream consequences after people hit a tipping point.

Action Editor

Randall W. Engle served as action editor for this article.

ORCID iD

Ed O'Brien https://orcid.org/0000-0002-4481-8408

Acknowledgments

The articles reviewed here (written with Nadav Klein) benefited from helpful discussion with Linda Hagen and the entire University of Chicago Booth School of Business (Chicago Booth) community (especially Nick Epley and George Wu). Chicago Booth's Center for Decision Research collected data. The University of Chicago Institutional Review Board approved all the experiments discussed.

Declaration of Conflicting Interests

The author(s) declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Funding

This work was supported by a Willard Graham Faculty Research Award and a Charles E. Merrill Faculty Research Award, both from The University of Chicago Booth School of Business.

References

Ayton, P., & Fischer, I. (2004). The hot hand fallacy and the gambler's fallacy: Two faces of subjective randomness? *Memory & Cognition*, *32*, 1369–1378.

Campbell, T., O'Brien, E., Van Boven, L., Schwarz, N., & Ubel, P. A. (2014). Too much experience: A desensitization bias in emotional perspective taking. *Journal of Personality and Social Psychology*, *106*, 272–285.

Davidai, S., & Gilovich, T. (2015). Building a more mobile America—One income quintile at a time. *Perspectives on Psychological Science*, *10*, 60–71.

Dweck, C. (2008). Can personality be changed? The role of beliefs in personality and change. *Current Directions in Psychological Science*, *17*, 391–394.

Ferguson, M. J., Mann, T. C., Cone, J., & Shen, X. (2019). When and how implicit first impressions can be updated. *Current Directions in Psychological Science*, *28*, 331–336.

Fisher, P. (2000). Sorites paradox and vague geographies. *Fuzzy Sets and Systems*, *113*, 7–18.

Fiske, S. T. (1980). Attention and weight in person perception: The impact of negative and extreme behavior. *Journal of Personality and Social Psychology*, *38*, 889–906.

Gilovich, T., Vallone, R., & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology*, *17*, 295–314.

Kardas, M., & O'Brien, E. (2018). Easier seen than done: Merely watching others perform can foster an illusion of skill acquisition. *Psychological Science*, *29*, 521–536.

Klein, N., & O'Brien, E. (2016). The tipping point of moral change: When do good and bad acts make good and bad actors? *Social Cognition*, *34*, 149–166.

Klein, N., & O'Brien, E. (2017). The power and limits of personal change: When a bad past does (and does not) inspire in the present. *Journal of Personality and Social Psychology*, *113*, 210–229.

Klein, N., & O'Brien, E. (2018). People use less information than they think to make up their minds. *Proceedings of the National Academy of Sciences, USA*, *115*, 13222–13227.

Klein, N., & O'Brien, E. (2019). *How many good or bad acts make good or bad actors? Moral change occurs sooner than people think*. Manuscript submitted for publication.

Levari, D. E., Gilbert, D. T., Wilson, T. D., Sievers, B., Amodio, D. M., & Wheatley, T. (2018). Prevalence-induced concept change in human judgment. *Science*, *360*, 1465–1467.

Massey, C., & Wu, G. (2005). Detecting regime shifts: The causes of under- and overreaction. *Management Science*, *51*, 932–947.

O'Brien, E. (2013). Easy to retrieve but hard to believe: Metacognitive discounting of the unpleasantly possible. *Psychological Science*, *24*, 844–851.

O'Brien, E. (2015). Mapping out past versus future minds: The perceived trajectory of rationality versus emotionality over time. *Journal of Experimental Psychology: General*, *144*, 624–628.

O'Brien, E. (2019). Enjoy it again: Repeat experiences are less repetitive than people think. *Journal of Personality and Social Psychology*, *116*, 519–540.

O'Brien, E., Ellsworth, P. C., & Schwarz, N. (2012). Today's misery and yesterday's happiness: Differential effects of current life-events on perceptions of past wellbeing. *Journal of Experimental Social Psychology*, *48*, 968–972.

O'Brien, E., & Kardas, M. (2016). The implicit meaning of (my) change. *Journal of Personality and Social Psychology*, *111*, 882–894.

O'Brien, E., & Klein, N. (2017). The tipping point of perceived change: Asymmetric thresholds in diagnosing improvement versus decline. *Journal of Personality and Social Psychology*, *112*, 161–185.

O'Brien, E., & Roney, E. (2017). Worth the wait? Leisure can be just as enjoyable with work left undone. *Psychological Science*, *28*, 1000–1015.

Reeder, G. D., & Coovert, M. D. (1986). Revising an impression of morality. *Social Cognition*, *4*, 1–17.

Ross, M. (1989). Relation of implicit theories to the construction of personal histories. *Psychological Review*, *96*, 341–357.

Schwarz, N. (2012). Why researchers should think "real-time": A cognitive rationale. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 22–42). New York, NY: Guilford Press.

Simons, D. J., & Ambinder, M. S. (2005). Change blindness: Theory and consequences. *Current Directions in Psychological Science*, *14*, 44–48.

Skowronski, J. J., & Carlston, D. E. (1989). Negativity and extremity biases in impression formation: A review of explanations. *Psychological Bulletin*, *105*, 131–142.

Strohminger, N., & Nichols, S. (2014). The essential moral self. *Cognition*, *131*, 159–171.

Tversky, A., & Koehler, D. J. (1994). Support theory: A nonextensional representation of subjective probability. *Psychological Review*, *101*, 547–567.

Wilson, A. E., & Ross, M. (2001). From chump to champ: People's appraisals of their earlier and present selves. *Journal of Personality and Social Psychology*, *80*, 572–584.

Wilson, T. D., & Gilbert, D. T. (2005). Affective forecasting: Knowing what to want. *Current Directions in Psychological Science*, *14*, 131–134.