Journal of Experimental Social Psychology 45 (2009) 421–423

Contents lists available at ScienceDirect

Journal of Experimental Social Psychology

journal homepage: www.elsevier.com/locate/jesp

FlashReport

The implications of imperfect measurement for free-choice carry-over effects: Reply to M. Keith Chen's (2008) "Rationalization and cognitive dissonance: Do choices affect or reflect preferences?" ☆

Brad J. Sagarin *, John J. Skowronski

Northern Illinois University, Department of Psychology, DeKalb, IL 60115, USA

ARTICLE INFO

Article history:
Received 15 May 2008
Revised 15 July 2008
Available online 10 September 2008

Keywords:
Cognitive dissonance
Measurement
Choice

ABSTRACT

Examinations of post-choice decision-making behavior often involve two successive choices. At time 1, participants choose between two equally attractive items. At time 2, participants choose between the unchosen item from time 1 and a new item that is roughly equal in attractiveness to the other two. The option rejected at time 1 will tend to again be rejected at time 2, a tendency often attributed to a psychological carry-over effect. Chen (2008) (Available from: http://cowles.econ.yale.edu/P/cd/d16b/d1669.pdf) discounts these psychological explanations. He argues that, given the time 1 choice, there is, in fact, a baserate probability of 66.7% that the initially unchosen item will be rejected again at time 2. However, Chen's argument rests on the unwarranted assumption that the time 1 choice provides a perfectly reliable measure of subjects' preference for the chosen item over the unchosen item. With more realistic estimates of the association between preference and choice, Chen's statistical explanation cannot fully account for the carry-over effect. Alternative experimental methodologies that eliminate Chen's statistical explanation are discussed.

This article has an unusual mission: refuting an argument presented in an unpublished working paper (Chen, 2008). Chen's paper offers a statistical explanation for carry-over effects in the free-choice paradigm, significantly challenging psychological explanations (e.g., cognitive dissonance) for such effects.

Refutations are not typically presented until the target paper successfully traverses peer review into publication. However, we believe that the circumstances surrounding this paper justify an exception to this guideline. First, Chen's argument, if correct, would have devastating consequences for a decades-long line of research. Second, Chen's argument has already affected the design of current and future experiments (L. Egan, personal communication, April 16, 2008). Third, Chen's argument gained high visibility by virtue of being featured in John Tierney's *New York Times* science column (Tierney, 2008), which included qualified support for the argument from prominent psychologists Daniel Gilbert and Laurie Santos. Fourth, feedback from the public posted to nytimes.com revealed the skepticism that the argument generated towards psychological science:

> I only have one question left: if psychologists refuse to listen to Dr. Chen's points (which, according to me, are perfectly valid and devastating), are they effectively exercising some form of choice rationalization of their imperfect theory of choice rationalization? (WiseGuy, 2008)

Given these considerations, we believe that an immediate refutation is warranted. To this end, we: (a) briefly review post-choice decision making behavior research; (b) summarize Chen's argument; (c) identify the unwarranted assumption central to the argument; (d) derive mathematical implications of likely violations of this assumption; (e) discuss the impact Chen's argument should have on interpreting past experiments; and (f) present two approaches for future research that address Chen's argument.

Psychological research has extensively examined post-choice decision making behavior. Such work spans decades, from Brehm's (1956) early work to Egan, Santos, and Bloom's (2007) recent experiments on decision making in children and capuchins. The typical paradigm is simple. At time 1, a chooser selects one of two options that are roughly equally attractive (e.g., a blue M&M vs. a red M&M). Later, the chooser makes another choice between two equally attractive options, one of which is the option rejected at time 1 (e.g., the rejected blue M&M vs. a green M&M). Data suggest that the initial rejection carries over: The option rejected at time 1 will tend to be again rejected at time 2.

This behavior is notable because the choice options used in such studies are pretested to establish equivalence in initial attractive-

DOI of original article: 10.1016/j.jesp.2008.08.026

☆ We thank Saurabh Asthana, Eric Downes, Joshua Goldberg, Bo Lawler, Bo Morgan, Chris Phelan, Eric Prebys, and Mary Reppy for their helpful comments on an earlier draft of this paper and Amanda Durik and Kimberly Lawler-Sagarin for their insightful discussions on the topic.

* Corresponding author. Fax: +1 815 753 8088.
E-mail address: bsagarin@niu.edu (B.J. Sagarin).

0022-1031/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.jesp.2008.08.024

ness. Hence, there should be no special bias against the rejected first option at the time of the second choice. The fact that there *does* seem to be a bias against the initially rejected option at the time of the second choice has led psychologists to invoke various psychological mechanisms (e.g., cognitive dissonance) to explain why the initial rejection carries over to the second choice.

In his working paper, Chen (2008) argues that explaining this effect does not require the invocation of psychological mechanisms. Instead, Chen argues that the time 2 rejection of the initially disfavored choice is a statistical likelihood.

Chen's argument rests on two linked assumptions. The first is that psychologists' pretesting only ensures that the choices are *approximately* equal in attractiveness, not *exactly* equal. Hence, choosers may still exhibit ordinal preferences. If this is the case, then in terms of the M&M color example there are six possible permutations of preference order:

(1) Red > Blue > Green
(2) Red > Green > Blue
(3) Blue > Red > Green
(4) Blue > Green > Red
(5) Green > Red > Blue
(6) Green > Blue > Red

Chen argues that the first choice made by subjects provides information regarding subjects' ordinal preferences; this constrains subsequent choices. For example, Chen argues that if at time 1 a subject chooses the red M&M over the blue M&M, such a choice eliminates three permutations (3, 4, and 6) because those permutations specify that blue is preferred to red. Importantly, green is preferred to blue in two out of three of the remaining permutations (2 and 5, not 1). According to Chen, this implies that the baseline probability of choosing a green M&M in the second choice is 66.67%, not 50% (and relates this to the Monty Hall effect). Hence, Chen argues that failure to select the blue M&M on the second choice is not a "rejection" of the blue M&M, but merely reflects the baserate odds of choosing green over blue *given the prior choice of red over blue.*

Although ingenious, Chen's explanation may rest on an unwarranted assumption. The argument seems to assume that the first choice provides a perfectly reliable measure of subjects' preference for the chosen option over the unchosen option. We don't believe that this is correct. Given that the three items were pretested to be roughly equal in attractiveness, any preference for one over another probably would be slight. Indeed, research frequently shows that "better" options are *not always* chosen over "poorer" options. Instead, choice behavior tends to be probabilistic. Choice-making often mirrors the perceived psychological "value" of attractive options, such that options that are slightly more valuable are chosen only slightly more often than options perceived as slightly less valuable (Carroll & de Soete, 1991; Estes, 1984; Navarick & Chellsen, 1983; Williams, 1985). Hence, while we agree with Chen that ordinal preferences may exist among choice options, and that a preferred option will tend to be chosen over a less preferred option, we suspect that in choice-making behavior such a preference will only be probabilistic (and slight), and will not take the all-or-none form depicted in Chen's argument.

What are the ramifications of any tendency for choices to stray from the preferred option? As it turns out, they are crucial. If the probability of choosing the preferred item over the less preferred item is less than 100%, then the choice of a red M&M over a blue M&M *reduces but does not eliminate* the probability of permutations 3, 4, and 6. Under these circumstances, calculating the precise statistical odds for choosing green over blue requires the estimation of two probabilities: the probability of choosing the preferred item over the less preferred item (e.g., the probability of choosing red over blue, given an underlying preference for red over blue) and the probability that the chosen item correctly identifies the underlying preference (e.g., the probability of an underlying preference for red over blue, given a choice of red over blue).

If, for simplicity's sake, we assume that red and blue M&Ms are preferred by equivalent proportions of subjects and that the probability of choosing the preferred color over the less preferred color is the same regardless of which particular color is preferred, then by Bayes' Theorem, these two probabilities are equivalent. (The basic argument is identical without these assumptions, but the calculations are a bit more cumbersome.) If we assume, for the sake of argument, a 75% figure for these probabilities (which implies that a subject who prefers red over blue has a 75% probability of choosing red over blue, and that a subject who chooses red over blue has a 75% probability of actually preferring red over blue), the choice of red over blue increases the probability of permutations in which red is preferred to blue (1, 2, and 5) from 16.7% to 25%, and reduces the probability of permutations in which blue is preferred to red (3, 4, and 6) from 16.7% to 8.3%. Given these probabilities, the probability that green is actually preferred to blue is 58.3% (25% + 25% + 8.3%). More generally, given a probability P that the chosen item correctly identifies the underlying preference, the probability that the third item (a green M&M) is actually preferred over the unchosen item (the blue M&M) would be $(P + 1)/3$.

Of course, the second choice is also subject to the same tendency to sometimes choose the less preferred option. Thus, if the probability that the third item is actually preferred to the unchosen item is 58.3% and the probability of choosing the preferred item over the less preferred item is 75%, then the probability of the subject choosing the third item over the unchosen item would be $58.3\% * 75\% + (1 - 58.3\%) * (1 - 75\%) = 54.2\%$. More generally, if P also represents the probability of choosing the preferred item over the less preferred item, the probability of the subject choosing the third item over the unchosen item (in the absence of any psychological carry-over effects) would be $(P^2 - P + 1) * 2/3$ (without the earlier specified assumptions, this probability would be $(2P_1P_2 - P_1 - P_2 + 2)/3$, where P_1 is the probability that the chosen item correctly identifies the underlying preference and P_2 is the probability of choosing the preferred item over the less preferred item). See Fig. 1 for a graphical depiction of this relation.

Note that the logic outlined above provides a statistical "expected value" for choice-making behavior. That is, if $P = 75\%$ (i.e., there is a 75% probability of choosing the preferred item over the less preferred item and a 75% probability that a chosen item correctly identifies the underlying preference) and the subject chooses red over blue at time 1, the expected probability of choosing green over blue at time 2 is only 54.2%. Importantly, this is lower than the probabilities observed in Egan et al.'s (2007) recent studies of children (63%) and capuchins (60%). Statistical tests would be needed, of course, to determine whether these data-based choice rates differ significantly from 54.2%, but they certainly suggest rejection at time 2 of the unchosen time 1 option.

Moreover, these calculations are based on relatively strong associations between preference and choice (75%). It is possible that, given the usual pretest data suggesting that there are very small initial attractiveness differences among the choice options, the actual associations might be substantially lower. If this were true, the actual expected value would be much closer to 50%, and the actual choice data would be even more indicative of a carry-over effect. Looked at another way, we can consider the plausibility of Chen's argument by calculating the value of P necessary for the expected probability of choosing green over blue at time 2 to reach the 60% level observed in the study of capuchins. It turns out that fully accounting for the choice data observed in the study of capuchins would require a value of P of 88.7%—a value that, in the con-

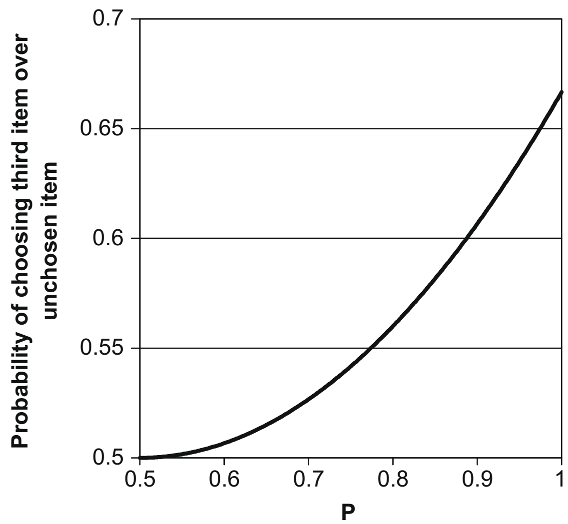

Fig. 1. Calculated probability of selecting new item instead of initially rejected item as a function of the strength of association between preference and choice (*P*).

text of the pretest data showing no measurable preferences, seems highly implausible.

We acknowledge that Chen's reasoning is thought-provoking. It certainly was exciting enough to generate considerable buzz, and to provoke us to analyze his arguments in detail. However, we conclude that Chen's reasoning does not provide a satisfactory explanation of choice-making behavior at time 1 that carries over to choices made at time 2. Indeed, our analysis suggests that Chen's claims about the failure of the choice manipulation to produce psychological carry-over effects are themselves to be viewed with considerable skepticism.

However, this debate is ultimately best settled via additional experimentation that sets up conditions eliminating Chen's baserate explanation. In this sense, Chen has performed a service to the science of psychology by offering an intriguing challenge to established orthodoxy. It is such challenges that provoke additional experimentation, and we are sure that Chen's argument will have exactly that effect. To this end, we offer two approaches.

One approach would be to determine an upper bound of *P*. To do so, researchers could measure initial preferences (latency to retrieve M&Ms, ratings on a scale, etc.) and choose two items that demonstrate just-measurable preference differences (i.e., two items that show a preference order, but of the minimum possible magnitude on the measure). The researchers would then have participants make a choice between the two. The proportion choosing the preferred item will provide an upper bound of *P*. From this upper bound, researchers could then calculate the statistical "expected value" for the time 2 choice and compare actual time 2 choice data against this value.

Another possibility is, in the successive choice paradigm, to have subjects make the initial choice "blind" (essentially making the first color choice a random one). That is, imagine that choice #1 is made when boxes hide the M&Ms, so subjects pick boxes, not colors (assuming colors are randomly assigned to box/position). Such a procedure would control for subjects' initial preferences (all 6 permutations described above would still be in play in such a circumstance). In such a case, rejection of the non-chosen alternative in subsequent choices would be compared appropriately to 50%. Indeed, we have recently learned that the Egan, Bloom, and Santos research team has already carried out a series of studies implementing this type of "blind" choice, with results supporting a psychological carry-over effect (L. Egan, personal communication, April 16, 2008).

While our reasoning suggests that Chen's proposal will not provide an adequate explanation for carry-over effects in choice, we know that we could be wrong, and that the data are the ultimate arbiter. Or, as people in our line of work say, "that's an empirical question." Let the experimentation begin! (Anyone have a few spare bucks for some M&Ms?)

References

Brehm, J. W. (1956). Postdecision changes in the desirability of alternatives. *Journal of Abnormal and Social Psychology, 52*, 384–389.
Carroll, J. D., & de Soete, G. (1991). Toward a new paradigm for the study of multiattribute choice behavior: Spatial and discrete modeling of pairwise preferences. *American Psychologist, 46*, 342–351.
Chen, M. K. (2008). Rationalization and cognitive dissonance: Do choices affect or reflect preferences? (Cowles Foundation Discussion Paper No. 1669). New Haven, CT: Yale University, Cowles Foundation for Research in Economics (2008). Available from: http://cowles.econ.yale.edu/P/cd/d16b/d1669.pdf (Retrieved July 14, 2008).
Egan, L. C., Santos, L. R., & Bloom, P. (2007). The origins of cognitive dissonance: Evidence from children and monkeys. *Psychological Science, 18*, 978–983.
Estes, W. K. (1984). Global and local control of choice behavior by cyclically varying outcome probabilities. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*, 258–270.
Navarick, D. J., & Chellsen, J. (1983). Matching versus undermatching in the choice behavior of humans. *Behaviour Analysis Letters, 3*, 325–335.
Tierney, J. (2008). And behind door no. 1, a fatal flaw. The New York Times. Message posted to: <http://www.nytimes.com/2008/04/08/science/08tier.html> (Retrieved from, May 12).
WiseGuy (2008). Message posted to: <http://tierneylab.blogs.nytimes.com/2008/04/07/monty-hall-meets-cognitive-dissonance/>.
Williams, B. A. (1985). Choice behavior in a discrete-trial concurrent VI-VR: A test of maximizing theories of matching. *Learning and Motivation, 16*, 423–443.