Running head: CHOICE AFFECTS AND REFLECTS PREFERENCES

How Choice Affects and Reflects Preferences: Revisiting the Free-Choice Paradigm

Date Submitted: 04/06/2009

Revision Submitted: 04/22/10

Abstract

After making a choice between two objects, people evaluate their chosen item higher and their rejected item lower (i.e., they "spread" the alternatives). Since Brehm's (1956) initial free-choice experiment, psychologists have interpreted the spreading of alternatives as evidence for choice-induced attitude change. It is widely assumed to occur because choosing creates cognitive dissonance, which is then reduced through rationalization. In this paper, we express concern with this interpretation, noting that the free-choice paradigm (FCP) will produce spreading, even if people's attitudes remain unchanged. Specifically, if people's ratings/rankings are an imperfect measure of their preferences, and their choices are at least partially guided by their preferences, then the FCP will measure spreading, even if people's preferences remain perfectly stable. We show this, first, by proving a mathematical theorem that identifies a set of conditions under which the FCP will measure spreading, even absent attitude change. We then experimentally demonstrate that these conditions appear to hold, and that the FCP measures a spread of alternatives, even when this spreading cannot have been caused by choice. We discuss how the problem we identify applies to the basic FCP paradigm as well as to all variants that examine moderators and mediators of spreading. The results suggest a reassessment of the free-choice paradigm, and perhaps, the conclusions that have been drawn from it.

Keywords: dissonance reduction, free-choice paradigm, revealed preferences

How Choice Affects and Reflects Preferences: Revisiting the Free-Choice Paradigm

For over fifty years, researchers have used variations of Brehm's (1956) seminal free-choice paradigm (FCP) to study the effect that choosing has on people's subsequent preferences. These studies have found, time and time again, that after being asked to make a choice, people's evaluation of their chosen alternative tends to rise, and the evaluation of their rejected alternative tends to fall (see, for example, Brehm, 1956; Festinger, 1964; Gerard & White, 1983; Lieberman, Oschner, Gilbert, & Schacter, 2001). This "spreading of alternatives" has been taken as evidence for choice-induced attitude change, and is cited as an example of the broader phenomena of cognitive dissonance reduction.[1]

Dissonance, the unpleasant motivational state that arises when one's behaviors or cognitions are inconsistent with one another, is reduced by shifting attitudes and beliefs to eliminate the inconsistency (Festinger, 1957, 1964). Thus, if an individual has a choice between two ice cream flavors, art prints, or appliances, the act of choosing one flavor, print, or appliance over another is believed to produce dissonance because any negative thought about the chosen alternative or positive thought about the rejected alternative will be inconsistent with the decision (Brehm, 1956; Festinger, 1957, 1964). To reduce the inconsistency and the experience of dissonance, the chooser can shift his or her preferences to desire the chosen object more and the rejected object less than he or she did before the choice.

In this paper, we argue that, although the spreading of alternatives has been found reliably in FCP experiments, it cannot be taken as evidence of choice-induced attitude change. This is because under a wide set of conditions, the FCP will measure spread, even if people's attitudes remain completely unchanged. Specifically, in any setting in which people's ratings/rankings imperfectly measure their preferences, and their choices are at least partially guided by their preferences (i.e., they are not simply random choices), then a FCP experiment will measure positive spreading among participants' evaluations, even for participants whose attitudes are perfectly stable. Fundamentally, this occurs because all FCP experiments employ a subtle form of non-random assignment; participants are analyzed differently depending on the choices they make in the study. If participants' choices are driven by underlying preferences, then this assignment process will produce positive spreading, even if participants' preferences remain perfectly stable.

The remainder of the paper proceeds as follows. First, we describe how spreading is typically calculated in the FCP. Then, we describe why incorporating choice information into the calculation of spread will produce positive spreading, even when participants' attitudes do not change. We prove a theorem to formalize our argument. The theorem identifies sufficient conditions under which the FCP will measure a spread of alternatives, independent of any attitude change. Next, we clarify the intuition behind our argument with two examples; first, with a study that examines the effect of an initial choice on a subsequent choice and, second, with the traditional paradigm that examines the effect of choice on subsequent preferences. Finally, we report the results of two empirical studies that were designed to disentangle the effects of choice-induced attitude change from the effects predicted by choice information. In our experiments, participants show spreading (as predicted by our theorem) despite the fact that the spread occurs before they make their choice, and therefore could not have been caused by the act of choosing. We conclude by discussing the implications of this work for studying choice-induced attitude change and dissonance reduction more generally.

Measuring Spread in the FCP

The typical setup for the free-choice paradigm includes three stages:

Stage:	One	Two	Three
Participants:	Rank or Rate	Choose	Rank or Rate

In Stages One and Three a participant is asked to rank or rate each item from a large set of goods, and in Stage Two, she is asked to choose which good she would prefer to have, from a two-good subset of the larger set. If the ranking for the item chosen in Stage Two improves between Stages One and Three, and/or the ranking for the item not chosen in Stage Two declines between Stages One and Three, then this is taken as evidence of cognitive dissonance reduction. In other words, spread is calculated by adding the amount the chosen item moves up in the ranking to the amount that the rejected item moves down (i.e., "*chosen spread*"). *Chosen spread* is then compared to zero. If it is positive, researchers conclude that choice-induced attitude change is present.[2]

Note, that the computation of *chosen spread* relies on participants' stage-two choices. Imagine, for example, that participants have been asked to choose between their 7th and 9th ranked items (out of fifteen) from Stage One (this is a common procedure). For participants who choose the item they initially ranked 7, *chosen spread* is calculated by measuring how much the 7th ranked item moves up and adding it to the amount that the 9th moves down (we will refer to this as *7-9 spread*). For participants who choose the item they initially ranked 9, *chosen spread* is calculated in exactly the opposite way; for them, *chosen spread* is how much the 9th ranked item moves up and how much the 7th ranked item moves down (we will refer to this as *9-7 spread*).

Often times, the FCP employs a control group. The control group ranks or rates twice (Stage One and Three), but does not choose between any items in the set in Stage Two (sometimes those participants are given one of the items as a gift and sometimes they are asked to make choices about other items). Because the choices of those in the control group are never learned, it is impossible to compare *chosen spread* across conditions. Thus, when comparing an experimental condition to a control group (e.g., Brehm, 1956), one common treatment is to calculate spread by measuring how much the item that is initially rated or ranked higher moves up and adding it to the amount that the item that is initially rated or ranked lower moves down (i.e., *high-low spread*). Then *high-low spread* is compared across conditions. Critically, if experimental participants choose the lower rated or ranked item (25% of participants show this "reversal," on average),[3] they are excluded from analysis, and the remaining 75% are compared to *everyone* in the control condition. For example, if participants in the experimental condition are given a choice between the items they initially ranked 7 and 9, then *high-low spread* is compared across conditions by calculating *7-9 spread* for only those experimental participants who chose their 7th ranked item, then comparing that to *7-9 spread* for all participants in the control condition. Thus, this measure of spread also relies on each participant's choice in the study because the choice is used as a criterion for including or excluding participants.

Note, critically, that both measures of spread in the traditional FCP paradigm (*chosen spread* and *high-low spread*) incorporate information revealed by participants' choices. The choice information is either used to calculate *chosen spread* or is used as a criterion for including or excluding participants when comparing *high-low spread* across conditions. This non-random treatment of participants is extremely problematic. If participants choose different items because they have different underlying preferences for the two items, then this process will produce spreading, even if participants' underlying preferences remain unchanged.

Interpreting Spread in the FCP

From the perspective of dissonance theory, spreading is predicted in the FCP because the act of choosing one item over another changes participants' underlying preferences for the chosen and unchosen items. A dissonance model of choice is not unique, however, in its prediction of spread in the FCP. The key insight we explore in this paper is that because the calculation of *chosen spread* and the criteria for including or excluding participants is based on each participant's choice, a preference-driven model of choice will also predict positive spreading. In other words, a model that assumes that people's choices are guided by their preferences also predicts spreading, but unlike dissonance models, in preference-driven models the choice process does not *cause* the spreading.

To see why a preference-driven model of choice predicts spreading, imagine that Jane and John are asked to rank fifteen items and then choose between items i and j (where item i is initially ranked 7 and item j is initially ranked 9). Imagine that in Stage Two, Jane chooses item i, which she ranked better than item j. Based on her choice, it seems reasonable to infer that she truly prefers i to j (perhaps to an even greater extent than was implied by her first ranking). Imagine that in Stage Two, John chooses item j. In other words, he ranks item i above item j, but then chooses item j. Based on his choice, it seems reasonable to infer that he does not prefer i to j to the extent that was implied by his first ranking (perhaps it is even reasonable to infer that he truly prefers item j).

If we observe that Jane and John make different choices in the experiment, then it is not unreasonable to suspect that Jane and John had different preferences for items i and j all along. That is, Jane probably always had a greater preference for i over j than John did. And, if we believe that they had different preferences all along, then we should not be surprised to see Jane and John re-rank the items differently. In other words, if we believe that Jane really likes item i more than item j, but we suspect that John may really like item j more than item i, then we should predict that Jane is more likely than John to show positive *7-9 spread* and John is more likely than Jane to show negative *7-9 spread*. Note that this prediction does not require us to assume that the choice process affected their subsequent preferences. Instead, the prediction is based on the assumption that each participant's choice of item i or item j helps inform us about each participant's true preferences for those items.

A preference-driven model of choice also predicts a conditional difference in *7-9 spread* when those who choose the higher ranked item in the choice condition are compared to all participants in the control condition (who rank twice but never choose). Imagine that participants like John (those who choose the item that they ranked 9 over the item they ranked 7 - i.e., item j over item i) are excluded from the choice condition. The choice condition, then, is only made up of participants like Jane-those who rank item i over item j AND choose item i. The control condition, in contrast, is made up of participants like Jane and John-those who would choose item i and those who, on second thought, would choose item j. Even if there is no dissonance reduction or change of preferences, this non-random selection of participants will produce more spreading in the choice condition than in the control condition because those participants who are expected to show negative *7-9 spread* are excluded from the choice condition and included in the control condition.[4]

The description above should make clear why, as an observer, one should expect Jane to show more positive *7-9 spread* than John, and why treating participants differently based on their choice is problematic for the FCP procedure. It is worth considering each of their behaviors from a first-person perspective as well.

Imagine that you are like Jane. You rank i above j and then choose i over j. Do you prefer i or j? It seems quite likely that you truly prefer i to j. It is not absolutely certain, of course. You might not have looked at the items carefully, you might have rushed through the task, or you might have been indifferent between the items and selected "at random". There is always some noise or random variation that accompanies the elicitation of preferences. Nevertheless, it is extremely likely that you truly prefer i to j (at least at the time of the experiment).

Imagine instead that you are like John. You rank i above j and then choose j over i. Do you prefer j or i? There are several possibilities. First, you may truly prefer j and your initial ranking may have been especially influenced by noise or random variation. Second, you may truly prefer i and your choice may have been especially influenced by noise. Third, you may be entirely indifferent between i and j and both the ranking and the choice were influenced by noise. Fourth, you may have changed your mind, such that you truly preferred i when you completed the ranking and truly preferred j moments later when you made your choice. In the paper, we will not explore which of these four cases best captures the underlying psychology of John's "reversal". We assume that if you rank and choose differently, there is some chance that you truly prefer i, there is some chance that you truly prefer j, there is some chance that you are entirely indifferent, and there is some chance that you changed your preference to j before choosing it.[5]

Note, however, that for all four cases, a preference-driven model of choice assumes that, at the time of re-ranking, your preference for i over j is likely to be smaller than the preference of someone who ranked i higher and also chose it. Therefore, even if the choice process does not affect anyone's preferences, we should expect more *7-9 spread* from someone who chose i over j than from someone who chose j over i.

Both a preference-driven model of choice and a dissonance model of choice predict positive spreading in the FCP. Theoretically, the two models can coexist. That is, it is possible for choices to reveal people's preferences and also affect their future preferences. To test this empirically, however, we must modify the FCP. The traditional FCP cannot distinguish the spreading that may result from dissonance processes and the spreading that is predicted because choices reveal information about participants' existing preferences. In order to estimate spreading that is due to the choice process, we have to try to account for the spreading that is predicted from the information revealed by choice.

The experiments presented in this paper were designed so that we could control for the information revealed by choice and isolate the spreading that may be caused by the choice process. Before describing the modified paradigm, however, we will more fully explain the problem with interpreting spread in the traditional FCP as evidence for choice-induced attitude change.

A Formal Approach to the FCP "Problem"

In this section, we describe three basic properties common to most preference-driven models of choice. Then, we provide a mathematical proof that demonstrates that these preference-driven models of choice predict positive spreading even if attitudes remain unchanged. The proof in the text is simplified to only deal with the case of comparing *chosen spread* to 0 for participants who make a choice in the FCP. Then, in the Appendix, we extend the proof to deal with alternative FCP procedures (e.g., comparing *high-low spread* after excluding participants who choose initially lower rated/ranked items, comparing spread for participants who make a choice or do not make a choice, and comparing spread for choices that are made between items initially rated/ranked "close" or "far").

To be most conservative, we confine our analysis to stable-preference models of choice. Note that we are not suggesting that in the typical FCP study, participants' preferences are perfectly stable. But, because our goal is to demonstrate that positive spreading is predicted even if there is no shift in preferences, our analysis explores what will happen in a FCP experiment when people's preferences are perfectly stable across time. By showing that a FCP will measure positive *chosen spread* even for people who have perfectly stable preferences, we demonstrate why "spreading" cannot be taken as a measure of cognitive dissonance reduction.

Assumptions

In preference-driven models of choice, people have preferences and these preferences are meaningfully (though sometimes imperfectly) captured by their ratings/rankings and choices. These models typically have three basic properties. First, people's ratings/rankings are at least partially guided by their preferences. In other words, an outside observer learns something about participants' preferences when participants rate or rank goods. Second, people's choices are at least partially guided by their preferences. In other words, we also learn about participants' preferences when we observe their choices. Note that this does not imply that individuals must always prefer their chosen item. Rather, it assumes that an individual's choice provides enough information about her preferences to predict that she is more likely to prefer the chosen item than the non-chosen item (i.e., people are not choosing completely at random). Finally, people's ratings/rankings are often not a perfect measure of their preferences. In other words, when we observe a person rate or rank good k higher than good l, we think it is more likely than not that she likes good k more than l, but we do not know this with certainty. We formally flesh out these properties below, then test empirically for their presence in two studies. We encourage readers to read the formal statement of the assumptions. Some readers, however, may wish to skip ahead to the section: *An Intuitive Approach to the FCP "Problem"*.

Our Formal Argument

What we will show is that in any setting where peoples' choices have these three properties - if people's preferences guide their evaluations and choices, but their preferences are not perfectly captured by those evaluations - then a FCP experiment will find positive spreading, even if participants' preferences are perfectly stable. To do this, we will now introduce some notation, which will allow us to state precisely under what conditions our critique of the FCP will be valid. Of course, a valid criticism does not imply that choice-induced attitude change does not exist. Our criticism simply implies that when we see spreading in the FCP, we cannot infer that attitude change has occurred. For simplicity we will use notation suited to describe a FCP experiment that asks participants to consider a set of fifteen goods. Formally:

Notation 1: Utility. Denote good $k \in \{1, 2, ..., 15\}$ as $x_k \in X$, where X is the set of all possible goods, and good x_k has an associate "preference level" or "utility" u_k.

Our first assumption relates people's preferences for goods to how they rate/rank those goods. Formally, let:

Notation 2: Ratings / Rankings. Denote by r_k^t the rating/ranking of good k in Stage t.

and:

Notation 3: Chosen Goods. Denote by $c_{\{k,l\}}^2$ the good a person would choose if asked to choose between x_k and x_l in Stage 2.

Assumption 1a (A1a). Ratings/Rankings (Stages One and Three) are partially guided by people's preferences over goods. In other words, if a person likes good k more than good j, they will tend to rate/rank k better than they rate/rank j. Mathematically, this assumes that:

$$u_k > u_l \quad \Rightarrow \quad \Pr[r_k^t > r_l^t] > 1/2$$

Given that we never directly observe how people feel about goods, this assumption may not appear to be very useful. However, note that we can apply Bayes' rule to rewrite this expression in a form which relates how people rate/rank goods to how they feel about those goods, a form we will find much more useful.[6] That is, Bayes' rule implies that we can also write $A1a$ as:

$$r_k^t > r_l^t \quad \Rightarrow \quad \Pr[u_k > u_l] > 1/2 \tag{A1a}$$

That is, when good k is rated/ranked higher than good l, good k has a strictly better than random chance of being liked more than good l.[7] For simplicity, it will be easier to assume a (somewhat stronger) condition.

Assumption 1 (A1). Ratings/Rankings provide a statistically unbiased measure of a participant's feelings about that good.

Written mathematically, this is:

$$E[r_k^t|u_k] = u_k \tag{A1}$$

Our theorem will hold under the more general $A1a$, but $A1$ will dramatically simplify the proof of our theorem (see the mathematical appendix for the proof under $A1a$).[8]

Assumption 2 (A2). Choices (Stage 2) are at least partially guided by preferences. In other words, we learn about participants' preferences by observing their choices. Note that this assumption does not imply that individuals must always prefer the chosen item. Rather, it implies that an individual's choice provides enough information about her preferences to predict that she is more likely to prefer the chosen item than the non-chosen item (i.e., our prediction is strictly greater than > 50%). Stated differently, when participants are given a choice between two goods, at least half of a study's participants should be expected to choose the good they truly prefer. Formally:

$$\Pr[c^2_{\{k,l\}} = x_k|u_k > u_l] > 1/2$$

A more useful (and mathematically equivalent) way to write this assumption is to say:

$$E[u_k - u_l|c^2_{\{k,l\}} = x_k] > 0 \tag{A2}$$

which is to say that if we see a person choose k over l, then we expect that she likes k more than she likes l.

Assumption 3 (A3). People's ratings/rankings are not a perfect measure of their preferences. This is to say, when we observe a person rate/rank good k higher than good l, we think it more likely than not that she likes good k more than l, but we do not know this with certainty. Formally:

$$1 > \Pr(u_k > u_l | r_k^t > r_l^t) > 1/2 \quad (A3)$$

This is the key assumption of our analysis. That is, for our theorem to be a critique of the FCP, it must be that ratings/rankings are a *useful* but *imperfect* measure of preferences.

Our theorem demonstrates that in any setting in which people's ratings/rankings and choices satisfy these assumptions, the FCP will measure a strictly positive amount of rating/ranking spread, even though participant's attitudes are assumed to be perfectly stable.

We can now formally state and provide a proof of our theorem. For a more complete exposition of this proof, (which covers more cases and relaxes several simplifying assumptions), please see the Appendix.

Our Theorem

Theorem 1 *Suppose that people's choices are driven by stable preferences which satisfy assumptions A1, A2 and A3. Then the free-choice paradigm will in expectation measure an increase in "spread" between stages 1 and 3, despite the lack of any attitude change.*

Remark 1 *Readers who are not interested in the exposition of the mathematical proof may wish to skip ahead to the section: An Intuitive Approach to the FCP "Problem".*

Mathematical Proof:

To see why the FCP will measure "spread" even when participants do not display attitude change, begin by supposing that in Stage Two, a participant in a free choice experiment is asked to choose between goods i and j, where good i was the good that was initially rated/ranked (in Stage 1) as better by the person by D rating/ranking points.

In most FCP experiments, i and j are chosen to be close together so that D is small, usually less than one rating-point or two ranks apart. Mathematically that is:

$$r_i^1 = r_j^1 + D \quad \text{where} \quad D > 0 \tag{1}$$

Note that for the purposes of this proof we have labeled the good initially rated/ranked higher as good i, and the good initially rated/ranked lower as j, and we follow a ratings-like convention that more attractive goods have higher evaluations ($r_i^1 > r_j^1$). This means that the initial amount of spread D is positive by construction, hence we have that $D > 0$ without loss of generality. Rewriting the proof to deal with a rankings-based ratings (in which a "1" if often better than a "2") system can be done easily.[9]

Now the free choice paradigm is going to look at what good the person chooses, and ask if the **chosen** good's rating/ranking rises and the **unchosen** good's rating/ranking falls. Mathematically then, what the FCP measures is:

$$\text{“}spread\text{”} = E[\left\{ \begin{array}{ll} (r_i^3 - r_i^1) - (r_j^3 - r_j^1) & \text{if} \quad c_{\{i,j\}}^2 = x_i \\ (r_j^3 - r_j^1) - (r_i^3 - r_i^1) & \text{if} \quad c_{\{i,j\}}^2 = x_j \end{array} \right\}] \tag{2}$$

or equivalently:

$$\text{“}spread\text{”} = E[\left\{ \begin{array}{ll} (r_i^3 - r_j^3) - D & \text{if} \quad c_{\{i,j\}}^2 = x_i \\ (r_j^3 - r_i^3) + D & \text{if} \quad c_{\{i,j\}}^2 = x_j \end{array} \right\}] \tag{3}$$

Free choice paradigm papers would claim that cognitive dissonance is present if the quantity in (3) is strictly greater than 0, that is if averaged among a large number of people, the spread of the chosen good over the unchosen good (which was D), increases.

Remark 2 *We will show that under assumptions* $A1$, $A2$, *and* $A3$ *people will display spreading, even when their attitudes remain completely stationary.*

First note that we can usefully expand (3) to its values in either possible case:

$$\begin{aligned}&\Pr(c^2_{\{i,j\}} = x_i) * E[(r^3_i - r^3_j) - D | c^2_{\{i,j\}} = x_i] + \\ &\Pr(c^2_{\{i,j\}} = x_j) * E[(r^3_j - r^3_i) + D | c^2_{\{i,j\}} = x_j]\end{aligned} \tag{4}$$

This writes expected spread as a probability weighted sum of expected ratings/rankings depending on which good a person chooses. Now note that we can say something about these expected Stage Three ratings. Since good i was rated/ranked D units better than j in Stage One (1) and since we are assuming that participants have stable preferences, you would expect that on average the distance between goods i and j would not change.[10] Mathematically then:

$$A1 \quad \Rightarrow \quad E[(r^3_i - r^3_j)] = D \tag{5}$$

Rewriting this expectation by expanding it out into the two possible choices a person can make, we can write the mathematically equivalent statement:

$$\begin{aligned}&\Pr(c^2_{\{i,j\}} = x_i) * E[(r^3_i - r^3_j) | c^2_{\{i,j\}} = x_i] + \\ &\Pr(c^2_{\{i,j\}} = x_j) * E[(r^3_i - r^3_j) | c^2_{\{i,j\}} = x_j] \\ &= D\end{aligned} \tag{6}$$

Evaluating this expression, note that we can say something about the probabilities that people will chose either object. Since participants initially rated/ranked good i better than good j, then by Assumptions $A1$ and $A2$ we can predict that they are more likely to pick i. That is to say:

$$A1, A2 \,\&\, A3 \quad \Rightarrow \quad 1 > \Pr(c^2_{\{i,j\}} = x_i) \geq \Pr(c^2_{\{i,j\}} = x_j) > 0 \tag{7}$$

where $A3$ was required for the strict inequalities.

Now note that some people both rated/ranked i better than j, and also chose i over j. Other people rated/ranked i better than j, but then on second thought, chose j. Therefore if ratings/rankings are not a perfect measure of preferences and choices tell us something about preferences, then those who chose i will on average like i relatively more than those who chose j. That is:

$$A2 \quad \Rightarrow \quad E[(u_i - u_j)|c^2_{\{i,j\}} = x_i] > E[(u_i - u_j)|c^2_{\{i,j\}} = x_j] \tag{8}$$

Since ratings/rankings are a noisy and unbiased measure of preferences, we can move from a statement about preferences to a statement about expected ratings:

$$A1 \ \& \ A3 \quad \Rightarrow \quad E[(r^3_i - r^3_j)|c^2_{\{i,j\}} = x_i] > E[(r^3_i - r^3_j)|c^2_{\{i,j\}} = x_j] \tag{9}$$

Recall from (6) that the weighted average of these two quantities was equal to D. Therefore, we know that the larger must be bigger than D and the smaller must be less than D. Mathematically:

$$E[(r^3_i - r^3_j)|c^2_{\{i,j\}} = x_i] > D > E[(r^3_i - r^3_j)|c^2_{\{i,j\}} = x_j] \tag{10}$$

Looking at the first part of this inequality and subtracting D from both sides gives us:

$$E[(r^3_i - r^3_j) - D|c^2_{\{i,j\}} = x_i] > 0 \tag{11}$$

Similarly, looking at the second part the inequality gives us:

$$E[(r^3_j - r^3_i) + D|c^2_{\{i,j\}} = x_j] > 0 \tag{12}$$

Combining the inequalities from (11) and (12), tells us that (4) is strictly bigger than 0, because (4) is the weighted sum of two things which are both bigger than 0. That is to say, we should expect the rating/ranking of the chosen good to rise and the rating/ranking of the unchosen good to fall, even if no person's preferences have moved.

Our proof shows that the FCP will measure positive spreading absent any form of attitude change or dissonance reduction. That is, we show that people who rate/rank goods and make choices simply based on their preferences will show positive spreading in the FCP (even if their attitudes remain unchanged). If spreading is predicted even when preferences are assumed to be stable, then spreading cannot be taken as evidence for dissonance reduction.

We believe that these three assumptions are not controversial – almost all models of choice incorporate these assumptions. Nevertheless, we believe it is prudent to empirically test whether choice information predicts spreading (as suggested by our proof). Before describing our experiments, however, we will try to clarify our argument intuitively.

An Intuitive Approach to the FCP "Problem"

As a field, psychologists clearly recognize that choices reveal information about preferences. Take, for example, the choice between two pieces of fruit. If your friend chooses an apple over a banana, it is not unreasonable to assume that she prefers the apple. And if two of your friends rate both an apple and banana as a 7 on a 10-point scale, but only one is given a choice between the two fruits, more is learned about the friend who rates and chooses than about the one who only rates. In other words, there is reason to believe that the chooser actually prefers the apple (assuming that she chose the apple, of course) and there is no reason to believe that of the non-chooser. Because we believe that the chooser actually prefers the apple, we would expect her to rank the apple higher than the banana if she were asked to rank several fruits (because this ranking makes sense given our knowledge of her existing preferences, irrespective of whether she experienced dissonance). We have no prediction for the non-chooser.

Although psychologists know that choices reveal information about people's existing preferences, this has been ignored in the FCP.

Subsequent choice

The clearest illustration of this problem is in a study that examines the effects of choice on subsequent choice behavior. In this methodology, participants rate several items and the experimenter chooses three that the participant has rated equally (**A**, **B**, and **C**). The participant is then given a choice between two of them, (say, **A** and **B**). Next, participants are given the choice between the rejected item (say, **B**) and the third item (**C**). For example, if an individual rates three pieces of fruit equally (**A**pple, **B**anana, and **C**herry), and then chooses an **A**pple over a **B**anana, she is presented with a **B**anana and a **C**herry for the final choice.

Egan, Santos, and Bloom (2007) used this methodology with 4-year olds and capuchin monkeys (using stickers and M&Ms). They predicted that participants who chose **A**pple over **B**anana in their initial choice would then choose **B**anana less often than **C**herry in their subsequent choice because rejecting it the first time made them desire it less than before (and therefore less than they desired **C**herry). In other words, they predicted that in the final choice, **B**anana would be chosen significantly less often than **C**herry (i.e., percent choosing **B**anana $< 50\%$). In contrast, they predicted that participants who did not make an initial choice, but were instead given an **A**pple at random, would be equally likely to choose **B**anana or **C**herry (i.e., percent choosing **B**anana $= 50\%$). They found their predicted effect. Children in the choice condition only chose **B**anana 37% of the time (significantly less than their predicted null of 50%), but those in the control condition chose it 53% of the time. Although, on the surface, it seems as if their hypothesis was supported, we contend that this is exactly the same set of results that you would expect if you knew that participants did not change their attitudes towards any of the fruits. We suggest that their test was incorrect because 50% is the wrong null. It does not account for the fact that information about participants' preferences was revealed by the first choice of **A**pple over **B**anana.

Fifty percent is only the correct null if participants are absolutely indifferent among the three options or if their choices are entirely random. If there are any differences in their initial preferences (i.e., if there is any noise in the initial ratings), and if their choices have any relationship to their preferences whatsoever, then the null for choosing **B** should be strictly smaller than 50%.

If the three items are not exactly equal to start, then there are six possible orders of preference: **A**pple, then **B**anana, then **C**herry (**ABC**); **ACB**; **BAC**; **BCA**; **CAB**; **CBA** (see Table 1). Because participants' choices are not random, we learn something important about their preferences from their initial choice of **A**pple over **B**anana. Three of the possible orders are contradicted by the initial choice (the actual order is unlikely to be **B**anana-**C**herry-**A**pple, **B**anana-**A**pple-**C**herry, or **C**herry-**B**anana-**A**pple), leaving three likely orders.[11] Two of the remaining three orders have **C**herry ranked above **B**anana and only one has **B**anana ranked above **C**herry. Thus, once we include the information provided by the initial choice, we expect that close to 1/3 of participants will choose **B**anana and close to 2/3 will choose **C**herry. Thus, even if participants' preferences remained perfectly stable during the experiment, we'd still expect them to choose **B**anana less than 50% of the time. Because we didn't learn anything about participants' preferences in the control condition, the correct null is 50% – three out of the six possible orders have **B**anana above **C**herry and three have **C**herry above **B**anana. Thus, even though Egan, Santos, & Bloom (2007) found that less than 50% chose **B**anana over **C**herry, we cannot conclude that participants' preferences changed during the experiment because we expect the exact same set of results if participants preferences do not change.[12] This does not mean that kids and monkeys can't display choice-induced dissonance reduction. It just means that these results do not demonstrate that they do.

Subsequent preferences

Egan, Santos, and Bloom (2007) examined subsequent choice because the traditional FCP was an unfeasible method for studying the preferences of young children and monkeys. We contend, however, that both measures of spread in the traditional paradigm (*chosen spread* and *high-low spread*) suffer from the same inherent problem as comparing the percent of participants who choose **B**anana over **C**herry to 50%. The problem is not as obvious as an incorrect null, but again, participants are treated as if their choices were random rather than a function of their existing preferences. By treating participants differently based on their choice (i.e., using the choice information to determine how to calculate spread or whom to include in the analyses), the FCP introduces a measurement bias into the protocol that causes ratings/rankings to spread, even if participants' attitudes have not been affected by their choices.

Let's look at our concern intuitively, using an example (see Table 2). Instead of fruit, we'll examine preferences for ice cream flavors. Imagine that your friend Beth is asked to rank fifteen flavors of ice cream twice. Imagine that her underlying preferences remain unchanged from time one to time two, but that there is some noise in her rankings (i.e., her rankings do not perfectly reflect her true preferences for all fifteen flavors). In the first ranking, Rocky Road is her seventh ranked flavor and Chocolate Chip is her ninth ranked flavor. A few minutes later she is asked to re-rank the ice cream flavors based on how she feels at that moment. Assuming that her second ranking is not identical to the first (which it almost never is), there are a number of possible ways that the ranking could have changed. Imagine that you know her re-ranking shifted one of these two flavors up two ranks and shifted the other flavor down two. Thus, Rocky Road could have moved up 2 in the rankings (to #5) and Chocolate Chip could have moved down 2 (to #11) (call this re-ranking X). Or, Rocky Road could have moved down 2 (to #9) and Chocolate Chip could have moved up 2 (to #7) (call this re-ranking Y). If we only know her first ranking, then these two possible re-rankings (X and Y) are both likely to be her true re-ranking.

Now, imagine that later that day she was going to be offered a choice between Rocky Road and Chocolate Chip ice cream. If you were asked to predict whether re-ranking X or re-ranking Y was the correct re-ranking, wouldn't you want to see into the future and learn what her choice would be? Imagine that you can see into the future and you know that she is going to choose Chocolate Chip. If you were asked to predict whether re-ranking X or re-ranking Y was correct, would you use that information or would you ignore it? We suggest that if you use the information from her upcoming choice (Chocolate Chip), re-ranking Y is more likely to be her second ranking (i.e., it is more likely that Chocolate Chip moved up and Rocky Road moved down than vice versa). If you assume that her first ranking is not a perfect representation of her preferences (i.e., if there is any noise in the ranking), then we learn more about her preferences by learning that she would choose Chocolate Chip over Rocky Road. By incorporating choice information, we have a better sense of her preferences, and we can use that to reasonably predict the direction of her re-ranking. Without choice information, re-ranking X and Y are equally likely. Knowing that Beth would choose Chocolate Chip, however, makes it reasonable to infer that re-ranking Y is more likely to be the true re-ranking than re-ranking X.

Instead, imagine that when you look into the future, you can see that she is going to choose Rocky Road. Now, X seems more likely. Using this information about her choice, we should expect that Rocky Road moved up (rather than down) relative to Chocolate Chip.

As long as one believes that the choice of one flavor or the other reveals anything about how much Beth likes each flavor, then learning which one she would choose affects our expectations for her re-ranking. This is because we now have two pieces of information: her initial ranking and a choice she would make. By incorporating the information revealed by the choice, we should expect the flavor that she would choose to move up (on average) and the flavor that she would reject to move down (on average). Note, the move from her initial ranking to her final ranking cannot be a reaction to choosing. Instead, it is predicted by the information that the choice reveals about her pre-existing preferences. Just as we learned that participants are likely to choose Cherry over Banana almost 66% of the time (if we know that they chose Apple over Banana), we learn that participants are more likely to re-rank an object higher rather than lower on a subsequent ranking if they would choose it over a close alternative, and more likely to re-rank it lower than higher if they would reject it. Thus, if Beth had been asked to choose between Rocky Road and Chocolate Chip before her second ranking (rather than later that day) and we found positive spreading, we would be unable to determine whether this was because Beth changed her preference to be consistent with her choice (i.e., to reduce dissonance) or whether she always had a stronger preference for a particular flavor, which we learned from her choice. To determine whether spreading is a reaction to the choice process, we need to compare spreading following a choice to the proper control group—that is, we need to compare it to a group for whom we have the same choice information.

Note that if all participants in a free-choice paradigm choose the higher ranked item (e.g., 7 over 9), then there would be no information revealed by the choice, and we could safely assume that spreading reflected choice-induced attitude change. This, as we mentioned, is not the case. But, it suggests that one way to fix the problem is to design a study that gets everyone to choose the same item while still having participants believe that they are freely choosing the item. This is the approach that has protected other dissonance paradigms from the problem raised here (see the General Discussion for a complete discussion of the implication of our argument for other dissonance paradigms, and see Risen and Chen, 2010a for a discussion of methods that researchers can use to properly study choice-induced attitude change).

We take a different approach in the current research, however. Rather than getting all participants to make the same choice, we try to control for choice information by finding out what participants in the control condition would have chosen. With this design, we have the opportunity to disentangle the spreading due to the choice process from the spreading due to choice information. In addition, with this design we can empirically examine the assumptions of preference-driven models of choice (i.e., do we find spreading when we only rely on choice information?).

Study One

Study one was designed to disentangle spreading due to choice-induced attitude change from spreading predicted by the extra information we have about choosers' preferences. Participants ranked art prints (Rank 1), chose between art prints (Choice 1), re-ranked them (Rank 2), and chose between more art prints (Choice 2). All participants made a critical choice between the prints they initially ranked 7 and 9. Participants who were randomly assigned to the Rank, Choose, Rank (RCR) condition made the critical choice between the two rankings, (in Choice 1). Participants assigned to the Rank, Rank, Choose (RRC) condition made the critical choice after the second ranking (in Choice 2).

With this method, we can compare the spreading of alternatives from Rank 1 to Rank 2 (i.e., how much the chosen item moved up in the rankings and the rejected item moved down) for participants who made the choice before the second ranking and for participants who made it after. Dissonance theory only predicts spreading in the RCR condition because the spreading of preferences is considered an effect of choosing. A preference-driven choice theory, in contrast, predicts positive spreading for both groups because everyone makes a choice between the items that they initially rank 7 and 9. This choice provides additional information about each participant's preference for these items, which should help predict the direction of their re-ranking.

The RRC condition in our experiment allows us to explicitly test whether information from choice produces positive spreading (as our theorem predicts). In addition, by measuring the change in ranking for participants who have not yet made the critical choice, we can estimate the spreading that cannot be due to the choice process. Then, by comparing chosen spread for RRC and RCR participants, we have the opportunity to estimate the spreading that can be explained by the choice process.

Method

Participants. Eighty University of Chicago undergraduates and graduate students participated in exchange for $4.

Materials and procedure. Upon arrival, participants were given fifteen postcard-sized art prints and were asked to rank them according to their aesthetic preference. They were told to put the print that they liked and would want to own most to the far left, the one they liked and would want to own least to the far right, and arrange the others between them accordingly. The art prints were 4 x 6 inches so that they could easily be handled, but participants were asked to imagine them as full sized posters. They included prints of artists such as Monet, Van Gogh, Kandinsky, Chagall, and Renoir.

The experimenter recorded the rankings, and then told participants that they would continue to evaluate art prints, but this time by choosing between pairs of prints. Participants were told for each pair to choose the print that they liked and would want to own more. The experimenter explained that participants may or may not see some of the same prints from before. In addition, participants were encouraged to make the choice carefully because they would be given one of their choices as a full sized poster at the end of the experiment. Thus, it was clear that their choices determined their prize. Participants in the RCR condition were asked to choose between 5 pairs of novel art prints and between the prints they had previously ranked 7th and 9th. The critical choice was the 2nd choice (out of 6). Participants in the RRC condition chose between 6 pairs of novel prints.

After making their choices, participants were asked to re-rank the initial prints from 1 to 15, as they had before. They were told that it was not a memory test and they should rank the prints according to what they most liked at that moment. The experimenter recorded their re-rankings, and then gave participants a second set of choices (6 more pairs) with the same instructions (choose the one you like/want more, you may or may not have seen these before, and you will be given one of your choices at the end, so choose carefully).

Participants in the RCR condition chose between 6 novel prints. Participants in the RRC condition chose between 5 novel pairs of prints and between the prints they had initially (i.e., on Rank 1) ranked as 7 and 9. The critical choice was the 2nd that they made. All participants were given a choice between Monet's Water Lily Pond and Van Gogh's Bedroom at the Arles as the 4th choice. Participants were given the poster they chose from this pair at the end of the experiment along with $4. Thus, as promised, participants received one of the posters that they chose. After the second set of choices, participants were asked how familiar they were with the prints (on a scale from 0 to 10) and were fully debriefed.

Results and Discussion

Participants in the two conditions were equally likely to choose their 7th ranked item over their 9th ranked item. 70% of participants in the RCR condition ($n = 28$) and 72% of participants in the RRC condition ($n = 29$) chose the item that they initially ranked 7th. And, in line with previous research, 30% of RCR participants ($n = 12$) and 28% of RRC participants ($n = 11$) contradicted their initial ranking by choosing their 9th ranked item instead (see Brehm, 1956; Gerard & White, 1983; Lieberman et al., 2001 for similar findings).

Three calculations were used to examine change in rankings. *Chosen spread* was the sum of the amount that the chosen item moved up in the ranks and the amount that the rejected item moved down in the ranks between Ranking 1 and Ranking 2. *Chosen spread* was also decomposed into *7-9 spread* (the sum of the amount the 7th ranked item moved up and the amount that the 9th ranked item moved down) and *9-7 spread* (the sum of the amount that the 9th ranked item moved up and the amount the 7th ranked item moved down) (see Table 3).

With the employed method, we were able to calculate *chosen spread* for both groups. In other words, we could determine the amount of spreading from Rank 1 to Rank 2 for participants who made their choice between the rankings and for those who made their choice after the second ranking.

We found positive *chosen spread* for participants in the RCR condition ($M = 2.20$, $SD = 3.33$, $n = 40$). In other words, on average, after choosing between the 7-9 pair, the chosen and the rejected items moved further apart in participants' rankings. Typical paradigms would compare this spread to 0 and conclude that there was choice-induced attitude change, $t(39) = 4.18$, $p < .001$.

Critically, however, as predicted by a preference-driven model of choice, we also found positive *chosen spread* for participants in the RRC condition ($M = 1.75$,$SD = 2.66$, $n = 40$). In other words, on average, from Rank Order 1 to Rank Order 2, the item that participants *eventually chose* also moved further apart in participants' rankings. This spreading was also significantly different from 0, $t(39) = 4.16$, $p < .001$. There was no difference in the amount of *chosen spread* for participants in the two conditions ($p > .50$). The spreading for RRC participants cannot be considered choice-induced attitude change because the spread occurred before the choice. Furthermore, it cannot be considered pre-decisional rationalization (see Brownstein, Read, & Simon, 2004) because participants did not know they would choose between these prints.

If the spreading for RRC participants could not have been a reaction to their choice, could it have been a reaction to the first ranking procedure? In other words, could the act of initially ranking 7 over 9 have produced dissonance that participants then reduced in Rank 2? Although this argument is not incompatible with dissonance theory, it is incompatible with our results. If RRC participants wanted to reduce dissonance after having ranked 7 over 9, then they would need to spread 7 up and 9 down (regardless of what decision they make in the future). Participants who later chose 9, however, spread their re-rankings in the opposite direction. Thus, empirically, we can conclude that RRC participants are not reducing dissonance created by the initial ranking. Moreover, why would dissonance be produced specifically for ranking 7 over 9? If dissonance was produced in the first ranking, it would be equally likely to be reduced, for example, by spreading item 5 up and item 7 down or item 9 up and 11 down. Thus, we feel confident concluding that the spread from Rank 1 to Rank 2 for RRC participants was not caused by dissonance reduction.

According to a preference-driven model of choice, regardless of when the choice occurs, choice information provides us with more information about participants' underlying preferences and therefore helps us predict the direction of the re-ranking. If a participant would choose 7 over 9, it is more likely that she prefers 7, and therefore we expect that 7 will move up and 9 will move down (on average) when the participant is asked to rank again. In contrast, if a participant would choose 9 over 7, then it is more likely that she prefers 9, and therefore we expect that 9 will move up and 7 will move down (on average), when the participant is asked to rank again.

Our results fit this interpretation. Looking at *7-9 spread*, we found that participants in both conditions who chose 7 (either before or after the second ranking) tended to rank the 7th ranked item higher the second time they did the ranking and tended to rank the

9th ranked item lower (RRC: $M = 1.14$, $SD = 2.28$, $n = 29$; RCR: $M = 1.36$, $SD = 2.45$, $n = 28$). In contrast, participants in both conditions who chose 9 (either before or after the second ranking) tended to rank the 9th ranked item higher the second time and the 7th ranked item lower (RRC: $M = 3.36$, $SD = 3.23$, $n = 11$; RCR: $M = 4.17$, $SD = 4.30$, $n = 11$). For participants who chose 7, there was no difference in *7-9 spread* between conditions, and for participants who chose 9, there was no difference in *9-7 spread* between conditions ($p > .60$). Thus, the information revealed by the choice helped us predict what direction the re-rankings would move (on average) regardless of when that information was revealed. It appears, then, that choices were revealing underlying preferences.

How can we be sure that the paradigm worked effectively? In other words, perhaps there was no difference between the two conditions because this particular choice did not create dissonance. This is an important concern, to which we will return. It is important to note, however, that if this were a traditional paradigm and we had not known the eventual choice of participants in the RRC condition, we would have concluded that there was choice-induced attitude change. First, as previously mentioned, there was positive *chosen spread* for those in the RCR condition. By comparing this to 0, researchers would (incorrectly) claim to have evidence of choice-induced attitude change. Second, if we compared *7-9 spread* for participants in the RCR condition who chose 7 over 9 ($n = 28$) to all participants in the RRC condition ($n = 40$), we would also see what appears to be dissonance reduction. Following the procedures initially made popular by Brehm (1956), we excluded the participants in the RCR condition who chose 9 over 7, included all participants in the RRC condition (because traditional paradigms could not determine whether they would have chosen 7 or 9), and then compared *7-9 spread* for the two samples. There was a significant difference in *7-9 spread* for these two samples, $t(66) = 2.03$, $p = .046$. In other words, while it may appear that participants became more fond of the higher ranked item after choosing it, in reality, participants were also "more fond" of the higher ranked item before choosing it (so long as they eventually chose it). Thus, the spreading seen here, which would normally be mistaken for evidence of choice-induced attitude change, is better interpreted as evidence for the importance of choice information.

Our paradigm was designed to control for choice information by separating the experience of choice from the information revealed by choice. However, the choice information may not have been identical for the two groups. The RCR group made the choice after one ranking and the RRC group made the choice after two rankings. Thus, the RRC is only an effective control if participants in both conditions made similar choices. To examine this, we compared the choices for participants in the two conditions. As mentioned above, participants in the two conditions were equally likely to choose the item that was initially ranked lower (i.e., Choice and Rank 1 "disagreed" for 30% of RCR participants and 28% of RRC participants). Further, we found that Choice and Rank 2 were also equally likely to "disagree" across conditions (13% of RCR and 10% of RRC). Rank 1 and Rank 2 "disagreed" for 23% of participants in each condition. And, when the two ranks disagreed, participants were equally likely across conditions (20%) to choose the item that was ranked lower in Rank 1. Finally, participants in the two conditions were equally likely to choose the item that was ranked lower in both Ranking 1 and Ranking 2 (RCR = 10% and RRC = 8%). Thus, on all potential measures of choice, we found that participants in the two conditions responded similarly, suggesting that the second ranking did not affect the choice (see Table 4).

The positive spread found for RRC participants supports a preference-driven model of choice. We were also able to test the specific assumptions of the model with our results.

Testing Assumption 1. Our first assumption was that the participants' ratings/rankings are at least partially guided by their preferences for the items they are rating/ranking. While this may seem uncontroversial, a formal test of this hypothesis is possible in the data by using participants from the RRC condition. Recall that these 40 participants were asked to rank fifteen art prints, then after making a set of unrelated choices, were asked to re-rank those same prints. A natural test of our first assumption is that the rank correlation between this first and second ranking is strictly greater than 0; that is, that a participant's first and second rankings are correlated at least a little.

For all 40 RRC participants, the rank correlation between their two rankings was strictly greater than 0. The average Spearman rank correlation for each participant was 0.86; this is significantly greater than 0 in a two-sided t-test at the 0.01% level ($n = 40$, $t = 39.6$, $p < .0001$). Indeed, for each individual participant we can compute the probability that their inter-ranking correlation could have arisen purely by chance. Individually, 39 of our 40 participants were significant at the 1% level. This seems to suggest that $A1$ is warranted.

Testing Assumption 2. Our second assumption was that the participants' choices are at least partially guided by their preferences. While this may also seem uncontroversial, a formal test of this hypothesis is possible in the data from both our RCR and our RRC conditions. Recall that all 80 participants were asked to rank fifteen art prints, then were later asked to make a choice between their 7th and 9th ranked items from their initial rankings. A natural test of our second assumption is whether participants chose their 7th ranked item at levels strictly greater than chance.

We tested this for our 80 participants using a two-tailed binomial test. 57 out of 80 participants chose their higher ranked good; this is significant at the 0.1% level ($p < .001$). This seems to suggest that $A2$ is warranted.

Testing Assumption 3. Our third assumption was that the participants' ratings/rankings contained at least a small amount of noise. A natural test of our third assumption is that the rank correlation between this first and second ranking be strictly less than 1. That is, if $A3$ is true, then when a participant ranks and then re-ranks a set of goods, those ranks should move at least a little. For all 40 RRC participants, the Spearman rank correlation between their two rankings was strictly less than 1 (no participant had perfectly identical rankings). Recall from our test of $A1$ that, on average, the rank correlation between first and second rankings was 0.86. This is significantly less than 1 at the 0% level, since at a rank correlation of 1 it is impossible to observe movements in rank ($n = 40$, $p = 0$).

However, it is of course theoretically possible that the instability between the two rankings of our RRC participants is entirely due to preference instability, not to measurement error in rankings. Put another way, how can we test that the rank correlation of 0.86 is not entirely due to extremely unstable preferences that are perfectly measured? Since we do not directly observe people's preferences, a direct test of extreme preference instability is impossible. However, we can test ancillary predictions of unstable preferences. For example, if rankings perfectly measure preferences, then for RRC participants, after controlling for the second ranking, the first ranking should have no ability to further predict a participant's choices. Intuitively, if preferences are measured perfectly (but move around in an unbiased way) then when predicting future behavior, older measurements of preferences contain no useful information once you control for more recent measurements.

We can test this by asking, does the first ranking help predict when the second ranking and the third-stage choice will disagree? Our 40 RRC participants all made choices in stage three between items ranked 7 and 9 in stage one. In 36 out of 40 participants their second ranking and their third-stage choice agreed; that is the participant chose the good in the third stage that was ranked higher in the second stage. Of those 36 times, 28 were when they chose the good initially ranked 7. By contrast, in only 1 of the 4 times they disagreed did the participant choose the good initially ranked 7.

This difference is significant at the 3% level ($z = 2.24$, $p = .0249$). Hence, even after controlling for more recent rankings, older rankings still have considerable power to predict future choices. This suggests that the imperfect rank correlation cannot be entirely due to unstable preferences, and therefore, that $A3$ is warranted.

Study Two

Study two was designed to give dissonance a "better shot." In study one, participants made twelve choices and only believed that they would receive one of the chosen posters. Thus, although the probability of not receiving a rejected print was 100% (and could prompt dissonance reduction), the probability of receiving the chosen print was only 8%. Perhaps, then, there was no difference in spreading for participants who chose before or after their second ranking because choosers did not need to rationalize a choice when there was a low probability of receiving it. Note, however, that past research has claimed to find dissonance reduction using paradigms that have participants make several choices (e.g., Lieberman et al., 2001) and we too found what appeared to be dissonance reduction even though there was none.

In study two, we used a paradigm in which participants knew that they would receive the chosen item and would not receive the rejected item with 100% certainty. As in study one, we manipulated whether participants made the critical choice before or after the second ranking to disentangle spreading due to dissonance reduction from spreading predicted by the additional information revealed by participants' choices.

Methods

Participants. One hundred University of Chicago undergraduates and graduate students participated in exchange for $2.

Materials and procedure. As in study one, participants were given fifteen postcard-sized art prints and were asked to rank them according to their aesthetic preference. The experimenter recorded the rankings, and then told participants that they would continue to evaluate art prints, but this time by choosing between a pair of prints. Participants were told to choose the print they liked more because they would be given the print at the end of the study. Participants in the RCR condition were asked to choose between the prints they had previously ranked 7th and 9th. Participants in the RRC condition chose between a novel pair (Monet's Water Lily Pond and Van Gogh's Bedroom at the Arles).

After making their choice, participants were asked to re-rank the initial prints from 1 to 15, as they had before. They were told that it was not a memory test and they should rank the prints according to what they most liked at that moment. The experimenter recorded their re-rankings, and then gave participants a second choice with the same instructions (choose the one you like more because you'll receive it at the end of the study). Participants in the RCR condition chose between the novel Monet and Van Gogh pair. Participants in the RRC condition chose between the prints they had initially (i.e., on Rank 1) ranked as 7 and 9. After the choice, participants were asked how familiar they were with the prints (on a scale from 0 to 10) and were fully debriefed. Participants were given the Monet or Van Gogh that they chose and were asked whether they would be willing to take an extra $2 instead of the second poster (we could not stock the fifteen posters from the ranking task). Those who were willing to forgo the poster were paid $4 and given either the Monet or Van Gogh poster. Those who wanted both of their chosen posters, were paid $2, given the Monet or Van Gogh poster, and provided their e-mail so that they could receive the second poster when the study was completed.

Results and Discussion

Participants in the two conditions were equally likely to choose their 7th or 9th ranked item. 22% of participants in the RCR condition ($n = 11$) and 24% of participants in the RRC condition ($n = 12$) "reversed" their initial ranking by choosing the 9th ranked item.

As in study one, we were able to calculate *chosen spread* for both groups (see Table 5). We found positive *chosen spread* for participants in the RCR condition ($M = 1.94$, $SD = 2.51$, $n = 50$, $t(49) = 5.46$, $p < .001$) and for participants in the RRC condition ($M = .94$, $SD = 2.77$, $n = 50$, $t(49) = 2.40$, $p < .02$). In other words, on average, from Rank 1 to Rank 2, the item that participants [eventually] chose moved up and the item that they [eventually] rejected moved down. Thus, replicating study one, we found support for our claim that positive spreading is predicted by the information revealed by choice. Unlike study one, however, there was a marginal difference in the amount of *chosen spread* for participants in the two conditions, $t(98) = 1.89$, $p = .06$.[13]

By decomposing *chosen spread*, we found that participants in both conditions who chose 7 tended to rank the 7th ranked item higher in the second ranking and tended to rank the 9th ranked item lower (RRC: $M = .61$, $SD = 2.28$, $n = 38$; RCR: $M = 1.31$, $SD = 2.04$, $n = 39$). In contrast, participants in both conditions who chose 9 tended to rank the 9th ranked item higher the second time and the 7th ranked item lower (RRC: $M = 2.00$, $SD = 3.86$, $n = 12$; RCR: $M = 4.18$, $SD = 2.82$, $n = 12$). For participants who chose 7, there was a trend for *7-9 spread* to be larger for RCR participants than for RRC participants, and for those who chose 9, there was a trend for *9-7 spread* to be larger for RCR participants than for RRC participants ($p < .16$). Thus, regardless of which art print participants chose, there was suggestive support that participants spread their alternatives more if they just made the choice than if we simply learned what the choice would have been.[14]

The results of both studies support the proof developed for a preference-driven model of choice. Because we see positive spreading for RRC participants, we can be sure that the information from choice is meaningful and that choice information can help predict which direction participants' re-rankings will go. A dissonance model of choice would not predict spreading for participants who made their choice after their second ranking. Note, however, that a simple, stationary preference-driven model of choice would not predict more spreading for participants in the RCR condition than the RRC condition, unless the model included a parameter for dissonance reduction, a parameter for learning through self-perception processes, or did not assume that preferences were stable across time.[15] In study one there was no difference in spreading for participants in the two conditions, but in study two there was a marginal difference. At the moment, then, it is unclear whether a preference-driven model of choice would be improved by including a parameter for dissonance reduction, learning, or drift in preferences. What is clear, however, is that to empirically examine choice-induced attitude change, it is essential to control for the information revealed by choice. Thus, we suggest that the dissonance model of choice would be improved if it formally recognized that choices reveal information about preferences. Guided by this new dissonance model, it would be clear to researchers that to test for choice-induced attitude change, one must control for choice information.

General Discussion

Brehm's (1956) results are often cited as the first experimental test of dissonance theory. Since then, the theory has been tested in several other paradigms. For example, if individuals have to go through a severe initiation to join a group, they come to like the group more than those who engage in a mild initiation (Aronson & Mills, 1959; Gerard & Mathewson, 1966). And, if individuals freely choose to write a counter-attitudinal essay, they come to agree with the position more than those who are forced to write the essay

(Linder, Cooper, & Jones, 1967). In these studies, all participants choose to write the counter-attitudinal essay and are willing to go through the initiation regardless of whether it is mild or severe.[16]

Imagine, however, that this were not the case. Imagine that only three-quarters of the participants in the severe initiation condition had agreed to go through with it. If participants had a real choice (rather than merely a perceived choice), we would have learned something important from their choice. We would have learned about their underlying preference for joining the group (presumably, those who were willing to undergo the severe initiation were especially interested in joining the group). Thus, we would not have been at all surprised that those who chose to go through with the severe initiation liked the group more than other participants. If participants had a real choice, it would be impossible to determine whether a difference in attitudes between conditions was caused by the severe initiation, or whether it should have been expected without regard to dissonance reduction, because the severe initiation condition only included participants with strong pre-existing attitudes (while those with weaker attitudes dropped out of that condition).

Similarly, imagine that only three-quarters of the participants in the free-choice condition had been willing to write the counter-attitudinal essay. If participants had a real choice, we would have learned something about the strength of their oppositional attitude (participants who wrote the essay were amenable enough to the position to write in favor of it). Again, we would not have been at all surprised that those who were willing to write the essay approved of the position more than other participants. And, again, it would be impossible to determine whether a difference in attitudes between conditions was caused by the act of choosing to write the essay, or whether it should have been expected because the free-choice condition only included participants with strong pre-existing attitudes.

To make a case that attitudes develop because of the severe initiation or the process of choosing to write an essay (i.e., to make a case that these preferences are constructed), it is essential that all participants agree to participate. Because there is no actual decision, there is no additional information revealed by their decision. Thus, the ability to convince participants that they have a choice when they don't actually have one is not only experimentally elegant, it is critical for the dissonance claim.

In the current paper we argue that, unlike in other dissonance paradigms, traditional free-choice studies cannot make a case for dissonance reduction. Past studies cannot determine whether spreading is explained by the process of making a choice or by the information revealed by the choice. Because participants are allowed to freely choose one item over another, we need a control group to account for the information that is revealed by choice. By using the proper control group, however, the current paradigm has the potential to disentangle spreading due to choice-induced attitude change from spreading that is expected based on the information revealed by choice.

Our results suggest that our concern about the traditional FCP is warranted. In both studies we find that spreading is predicted by the information provided by the choice even when participants do not make the choice before the re-ranking. This spreading cannot be explained by dissonance theory or self-perception theory. And, because participants do not expect to make this choice, it cannot be explained by a theory of pre-choice rationalization (after all, how would participants know to spread 7 and 9 rather than 3 and 5 or 5 and 7, etc.). Because participants cannot simultaneously spread all items up and down, we can be sure that the spreading is not a reaction to an upcoming choice. Finally, because RRC participants who eventually choose 9 over 7 change their rankings so that 9 moves up and 7 moves down, we can be certain that their re-ranking is not a means of reducing dissonance that was created from the first ranking. In other words, even though it is conceivable that a ranking procedure could produce dissonance, we do not find support for this interpretation. If RRC participants wanted to reduce dissonance for ranking 7 over 9, then they would need to spread 7 up and 9 down. Participants who eventually choose 9 spread their re-rankings in the opposite direction from what would be predicted if there was dissonance from the initial ranking.

Just by knowing whether participants would choose 7 over 9 or 9 over 7, we can predict which way their initial rankings will tend to move when we ask them to rank again. In other words, because there is some noise in the initial ranking and because choice provides additional information about a participant's underlying preference, knowing what they *would* choose can help predict which direction the re-ranking is likely to move. A preference-driven model of choice simply requires that people have preferences and these preferences are meaningfully, but imperfectly captured by their ratings/rankings and their choices. With these assumptions, spreading is predicted from Rank 1 to Rank 2, and no additional psychological mechanism needs to be posited. Dissonance reduction only needs to be invoked to the extent that RCR shows more spread than RRC.

We believe that the reported results validate our initial concern and have important implications for the theory of dissonance and for how it should be studied. First, we suggest that the field revisit recent work examining choice-induced attitude change in unusual groups. Second, we suggest that these implications need to be incorporated into studies designed to explore moderators and mediators of choice-induced dissonance.

Who Shows Choice-Induced Dissonance?

Research in choice-induced attitude change began over fifty years ago, but it is only in the last decade that social psychologists have started testing whether "unusual" participants experience dissonance. Egan, Santos, and Bloom (2007) examined the effect of an initial choice on subsequent choices for monkeys and children, and claimed that both groups displayed choice-induced attitude change. As described in our introduction, however, we contend that the null hypothesis that the results were tested against was incorrect because it failed to consider the information provided by a participant's initial choice. Thus, we do not believe that these experiments provide evidence for monkeys or children showing dissonance.[17]

Lieberman, Oschner, Gilbert, and Schacter (2001) examined the effect of choice on subsequent preferences for amnesiacs and participants under cognitive load. In fact, the methodology of the current study one, in which participants made several choices rather than just one, was based on this paradigm (though they used a within subject design and we used a between subjects design). They had participants rank two sets of art prints, make a choice between pairs of prints from one of the sets (along with several novel choices), and then re-rank both sets of prints. They found positive spread in the critical set. In other words, both amnesiacs and matched controls re-ranked the chosen posters higher and the non-chosen posters lower. However, as we have argued throughout the paper, spreading may or may not indicate choice-induced attitude change. To know whether there is choice-induced attitude change, we need to control for the information provided by the choice. When we controlled for choice information in a similar multi-choice paradigm (study one), we found spreading, but no evidence of dissonance.[18]

In some sense, our participants in the RRC conditions can be considered "amnesiacs." That is, when re-ranking the prints, those participants have no memory for their choice (because they have not made it yet). The positive spread demonstrated by the RRC participants is clearly not a reaction to the choice. It is simply predicted by the information revealed by their eventual choice. At present, we believe that the data for amnesiacs ought to be interpreted the same way.

We do not mean to pick on these particular studies. Egan, Santos, and Bloom (2007), and Lieberman et al. (2001) make the same assumption as do all of the papers using the free-choice paradigm – they assume that there is no information provided by the choice. We note these examples in particular because of the strong influence these results have for researchers trying to understand dissonance theory. For example, if it is taken as given that amnesiacs, monkeys, and children all show choice-induced dissonance then dissonance theory needs to be modified so that explicit memory and a developed sense of self are no longer part of the process. We believe that those modifications are premature, given the evidence. To truly understand what causes choice-induced attitude change, we need to be able to carefully rule in and rule out any unusual group.

Studying Dissonance

Research using the free-choice paradigm has not only tried to demonstrate the presence of choice-induced attitude change, but has also tried to examine the moderators and mediators of this effect. For example, Brehm (1956) compared “high” and “low” dissonance conditions by having participants choose between items initially rated close together (difficult decision ⇒ high dissonance) or far apart (easy decision ⇒ low dissonance). More spreading was found following a high dissonance choice than a low one (Brehm, 1956). In addition, cross-cultural comparisons have found that participants from Eastern cultures (e.g., Japan) show less spreading following personal choice than participants from Western cultures (Heine & Lehman, 1997; Hoshino-Browne et al., 2005; Kitayama, Snibbe, Markus & Suzuki, 2004). And, research has demonstrated less spreading for participants who have been self-affirmed than for those who have not (Steele, 1988; Steele, & Liu, 1983; Steele, Spencer, & Lynch, 1993). Most recently, new experiments have used neural correlates of choice rather than direct preference measures and found that these neural correlates also “spread” in a FCP setting (Sharot, De Martino, & Dolan, 2009).

While interesting, the problem we identify still holds in all of these experiments; all of these results are possible even if participant’s attitudes remain completely stable. Because of this we believe that when interpreting past results and, even more importantly, when designing future studies, researchers need to be careful of whether the moderator or mediator in question is actually affecting dissonance processes or whether the effects on spreading can be explained by a natural interaction with the information revealed by choice.

For example, take the comparison of high and low dissonance. Brehm (1956) argued that a choice between close alternatives creates more dissonance than one between far alternatives, which, in turn, prompts more choice-induced attitude change. But, a preference-driven model of choice also predicts more spreading when the alternatives are close because more information is revealed by the choice. In fact, it is a first-order prediction of a preference-driven model that there should be more spreading for a close choice than for a far choice, and more spreading for a far choice than for no choice (see the Appendix for a formal proof of this). Intuitively, when given a choice between alternatives that are initially far apart, almost everyone will choose the higher ranked item (e.g., 4 over 12), and therefore the choice provides little additional information about their preferences. To test whether there is actually more choice-induced attitude change in the "high" dissonance condition, one could have participants make the high or low dissonance choice before or after the second ranking. A theory of dissonance would only predict the close-far difference in spreading if the choice is made before re-ranking. A preference-driven model of choice predicts the same close-far difference for participants who make the choice after both rankings.

In contrast to the high vs. low dissonance case, we concede that a preference-driven model of choice does not, on its surface, predict less spreading for Eastern or affirmed participants. These moderators were chosen for examination based on dissonance theory, and there is evidence that is particularly difficult to interpret through a preference-driven perspective (e.g., affirmation reduces dissonance in forced compliance studies and Japanese participants show spreading if interpersonal dissonance is prompted with a face prime). Nevertheless, because it is possible that culture, affirmation, or other potential moderators can affect the amount of information revealed by choice, we believe that if one wants to use the FCP to draw definitive conclusions about moderators of choice-induced attitude change, these studies must control for the information revealed by choice.

How, for example, might information from choice generate cultural or affirmation-based effects on spreading? If Japanese or affirmed participants produce more consistent rankings across time, then there is less room for choice information to matter. That is, if for cultural reasons Japanese participants try harder to correctly represent their preferences, or affirmed participants have better insight into their preferences, then the free-choice paradigm would yield differences that might (mistakenly) be interpreted as differences in dissonance reduction. Or, if Japanese or affirmed participants' choices reveal less about their preferences (i.e., there is less stability between ranking and choice) then knowing what they would choose will be less helpful in predicting the movement from Rank 1 to Rank 2. In other words, if Japanese participants' choices reflect something other than their personal preferences, their choices would not be informative for predicting the direction of spreading. In fact, recent work by Savani, Markus, and Conner (2008) found that Indian participants were less likely to choose based on their personal preferences compared to North American participants. If this same cultural difference exists for Japanese and North American participants, this could potentially account for the cultural differences found in spreading.

Thus, even if on the surface it seems unlikely that a moderator would affect spreading by affecting the information revealed by choice, it is experimentally prudent to control for choice information. Only then, can we be sure that the moderator has affected dissonance processes, rather than moderating the degree to which ratings, rankings, and choices reflect preferences.[19] Note that inclusion of the proper control condition does not necessarily mean doubling the number of conditions or the sample. The proper control can be created using a within-subjects RCRC design (see Risen & Chen, 2010b).

These concerns are equally important as future research begins to look at individual differences in the tendency to demonstrate choice-induced attitude change, and the underlying neural mechanisms for choice-induced attitude change (e.g., Harmon-Jones, Harmon-Jones, Fearn, Sigelman, & Johnson, 2008; Sharot, De Martino, & Dolan, 2009). If certain people (e.g., maximizers or those who are high in need for closure) show more spreading, we can only conclude that the trait is associated with choice-induced attitude change if we also control for the information revealed by the choice.

Similarly, we must have the proper control group to be sure that a difference in neural activation is due to choice-induced attitude change rather than the information that tends to be revealed by a choice. For example, Sharot, De Martino, and Dolan (2009) found that activation in the caudate nucleus tracked participants' spreading of ratings after a choice. In other words, the difference in activation for the selected and rejected alternatives was greater after the decision than before the decision. The authors interpret this as physiological evidence of choice-induced attitude change. Note, again, however, that spreading in choice-correlated neural activation is a natural prediction of any preference-driven model of choice. This is precisely because, just like ratings and rankings, these measures may be a useful but imperfect measure of preferences.[20]

To determine whether neural results truly support dissonance theory, one must control for the information revealed by choice (one possibility is to employ the methodology used in the current studies – i.e. compare the neural activation in a RRC condition to the activation in a RCR condition). Without a proper control, however, we cannot draw conclusions about the mediators of choice-induced attitude change. Thus, we reiterate how important it is for dissonance researchers to design studies that control for the information revealed by choice.

Conclusion

Theoretically, it is easy to reconcile the perspective that choice affects preferences with the perspective that choice reflects preferences. The problem is that, as a field, we have overlooked the fact that choice reflects preferences within the FCP. Thus, the empirical study of choice-induced attitude change has not been able to control for the information revealed by choice. By controlling for revealed preferences, however, we then have the right tools to study the behavior-induced construction of preferences.

References

Aronson, E. & Mills, J. (1959). The effect of severity of initiation on liking for a group. *Journal of Abnormal and Social Psychology*, 59, 177-181.

Bem, D. J. (1967). Self-perception: An alternative interpretation of cognitive dissonance phenomena. *Psychological Review*, 74, 183-200.

Brock, T. C. (1963). Effects of prior dishonesty on post-decision dissonance. *Journal of Abnormal and Social Psychology*, 66, 325-331.

Brehm, J. W. (1956). Post-decision changes in desirability of alternatives. *Journal of Abnormal and Social Psychology*, 52, 384-389.

Brehm, J. W., Cohen, A. R. (1959). Re-evaluation of choice alternatives as a function of their number and qualitative similarity. *Journal of Abnormal Social Psychology*, 58, 373-78.

Brownstein, A. L., Read, S. J., Simon, D. (2004). Bias at the racetrack: Effects of individual expertise and task importance on predecision reevaluation of alternatives. *Personality and Social Psychology Bulletin*, 30, 891-904.

Chen, K. (2008). Rationalization and Cognitive Dissonance: Do Choices Affect or Reflect Preferences? *Cowles Foundation for Research in Economics*, Yale University, discussion paper no. 1669. http://cowles.econ.yale.edu/P/cd/d16b/d1669.pdf

Chen, M. K., & Risen, J. L. (2009). Is choice a reliable predictor of choice? A comment on Sagarin and Skowronski. *Journal of Experimental Social Psychology*, 45, 425-427.

Chapanis, N. P., & Chapanis, A. (1964). Cognitive dissonance: Five years later. *Psychological Bulletin*, 61, 1-22.

East, R. (1973). A methodological suggestion for the study of re-evaluation of alternatives. *European Journal of Social Psychology*, 3, 329-332.

Egan, L.C., Bloom, P., & Santos, L.R. (2010). Choice-induced preferences in the absence of choice: Evidence from a blind two choice paradigm with young children and capuchin monkeys. *Journal of Experimental Social Psychology*, 46: 204-207.

Egan, L. C., Santos, L. R., & Bloom, P. (2007). The origins of cognitive dissonance: Evidence from children and monkeys. *Psychological Science*, 18, 978-983.

Festinger, L. (1957). *A theory of cognitive dissonance.* Stanford, CA: Stanford University Press.

Festinger, L. (1964). *Conflict, decision, and dissonance.* Stanford, CA: Stanford University Press.

Galton, F. (1886). Regression towards mediocrity in hereditary stature. *Nature* 15, 246–263

Gerard, H. B., & Mathewson, G. C. (1966). The effects of severity of initiation on liking for a group: A replication. *Journal of Experimental Social Psychology*, 2, 278-287.

Gerard, H. B., & White, G. L. (1983). Post-decisional reevaluation of choice alternatives. *Personality and Social Psychology Bulletin*, 9, 365-369.

Greenwald, H. J. Dissonance and relative versus absolute attractiveness of decision alternatives. (1969). *Journal of Personality and Social Psychology*, 11(4), 328-333

Harmon-Jones, E., Harmon-Jones, C., Fearn, M., Sigelman, J. D., &Johnson, P. (2008). Left frontal cortical activation and spreading of alternatives: Tests of the action-based model of dissonance. *Journal of Personality and Social Psychology*, 94, 1-15.

Heine, S. J., & Lehman, D. R. (1997). Culture, dissonance, and self-affirmation. *Personality and Social Psychology Bulletin*, 23, 389-400.

Hoshino-Browne, E., Zanna, M. P., Spencer, S. J., Zanna, A. S., Kitayama, S., & Lackenbauer, S. (2005). On the cultural guises of cognitive dissonance: The case of easterners and westerners, *Journal of Personality and Social Psychology*, 89, 294-310.

Kitayama, S., Snibbe, A. C., Markus, H. R., & Suzuki, T. (2004). Is there any "free" choice? Self and dissonance in two cultures. *Psychological Science*, 15, 527-533.

Lieberman, M., Oschner, K., Gilbert, D., & Schacter, D. (2001). Do amnesiacs exhibit cognitive dissonance reduction? The role of explicit memory and attention in attitude change. *Psychological Science*, 12, 135-140.

Linder, D. E., Cooper, J., & Jones, E. E. (1967). Decision freedom as a determinant of the role of incentive magnitude in attitude change. *Journal of Personality and Social Psychology*, 6, 245-254.

Oshikawa, S. (1971). Dissonance reduction or artifact? *Journal of Marketing Research*, 81, 514-515.

Risen, J. L., & Chen, M. K. (2010a). *How to study choice-induced attitude change: Strategies for fixing the free-choice paradigm.* Manuscript in preparation.

Risen, J. L. & Chen, M. K. (2010b). *On the measurement of cognitive dissonance from choice*. Manuscript in preparation.

Sagarin, B. J., & Skowronski, J. J. (2009). 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?" *Journal of Experimental Social Psychology*, 45, 421-423.

Samuels, M.L. (1991). Statistical reversion toward the mean: More universal than regression toward the mean. *The American Statistician,* 45, 344–346.

Savani, K., Markus, H. R., & Conner, A. L. (2008). Let your preference be your guide? Choices and preferences are more tightly linked for North Americans than for Indians. *Journal of Personality and Social Psychology*, 95, 861-876.

Sharot, T., De Martino, B., Dolan, R. J. (2009). How choice reveals and shapes expected hedonic outcome. *The Journal of Neuroscience*, 29, 3760-3765.

Shultz, T. R., Léveillé, E., & Lepper, M. (1999). Free choice and cognitive dissonance revisited: choosing "lesser evils" versus "greater goods". *Personality and Social Psychology Bulletin*, 25, 40.

Shultz, T. R., & Lepper, M. R. (1992). A constraint satisfaction model of cognitive dissonance phenomena. *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society* (pp. 462-467). Hillsdale, NJ: Edbaum.

Shultz, T. R. & Lepper, M. R. (1996). Cognitive dissonance reduction as constraint satisfaction, *Psychological Review*, 103, 219–240.

Steele, C. M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 21, pp. 261 – 346). San Diego, CA: Academic Press.

Steele, C. M., & Liu, T. J. (1983). Dissonance processes as self-affirmation. *Journal of Personality and Social Psychology*, 45, 5-19.

Steele, C. M., Spencer, S. J., & Lynch, M. (1993). Self-image resilience and dissonance: The role of affirmational resources. *Journal of Personality and Social Psychology*, 64, 885-896.

Tversky, A. & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124-1131.

Tables

Table 1: Likelihood of preferring **B**anana or **C**herry based on the choice of **A**pple over **B**anana

	Possible Preference Orders					
	1	2	3	4	5	6
Best:	**A**pple	**A**pple	**B**anana	**B**anana	**C**herry	**C**herry
Middle:	**B**anana	**C**herry	**A**pple	**C**herry	**A**pple	**B**anana
Worst:	**C**herry	**B**anana	**C**herry	**A**pple	**B**anana	**A**pple
Prefer: (**B** or **C**)	**B**	**C**			**C**	

Note. Of the 6 possible orders of preferences, only Orders 1, 2, and 5 have **A**pple ranked above **B**anana. Of those three possible orders, two have **C**herry ranked above **B**anana and one has **B**anana ranked above **C**herry. Thus, once we learn that participants prefer **A**pple to **B**anana, we should expect that they will prefer **C**herry to **B**anana close to 2/3rds of the time.

Table 2: Possible Rankings and Re-rankings for Ice Cream Flavors

Rank:	Initial Ranking	Re-Ranking X	Re-Ranking Y
1			
2			
3			
4			
5		Rocky Road	
6			
7	Rocky Road		Chocolate Chip
8			
9	Chocolate Chip		Rocky Road
10			
11		Chocolate Chip	
12			
13			
14			
15			

Note. On the initial ranking, Rocky Road is ranked 7 and Chocolate Chip is ranked 9. If you know that Beth would choose Rocky Road over Chocolate Chip, then re-ranking X is more likely that re-ranking Y. If you know that Beth would choose Chocolate Chip over Rocky Road, then re-ranking Y is more likely than re-ranking X. Thus we can predict the direction of spread even if Beth does not make a choice between ranking and re-ranking. In other words, positive spread is predicted by the information revealed by the choice.

Table 3: Mean Spreading for the RCR and RCR Conditions in Study One

	Condition	
Measures of Spread (Participants included)	Rank-Choose-Rank	Rank-Rank-Choose
Chosen Spread:	$^{a}2.20_{a,b}$	1.75_{a}
(All Participants)	$(3.33, 40)$	$(2.66, 40)$
7-9 Spread:	$^{b}1.36_{a}$	1.14_{a}
(If Chose 7)	$(2.45, 28)$	$(2.28, 29)$
9-7 Spread:	4.17_{b}	3.36_{b}
(If Chose 9)	$(4.30, 12)$	$(3.23, 11)$
7-9 Spread:	$-.30_{c}$	$^{b}-.10_{c}$
(All Participants)	$(3.99, 40)$	$(3.19, 40)$

Note. Standard deviations and number of participants are presented in parentheses (SD, n). Across conditions, means that share a subscript letter do not differ significantly at $p < .05$.

a) To calculate dissonance reduction, traditional paradigms would compare chosen spread in the RCR condition to 0.

b) Or, traditional paradigms would compare 7-9 spread for those who chose 7 in the RCR condition to all participants in the RRC condition. Both traditional measurements find "dissonance reduction." However, the correct comparisons (calculation of spread across the two conditions) suggest that there is no difference in spread for those who choose before or after their second ranking.

Table 4: Number of Participants who Made Particular Choices in Study One

	Condition	
	Rank-Choose-Rank	Rank-Rank-Choose
Total n:	40	40
R1 and Choice disagree:	12	11
R1 and R2 disagree:	9	9
When R1 and R2 disagree, choose 9:	8	8
R2 and Choice disagree:	5	4
Both R1 and R2 disagree with Choice:	4	3

Note. R1 and R2 denote Ranking 1 and Ranking 2, respectively. Across conditions, on every measure of choice, participants in the two conditions behaved similarly.

Table 5: Mean Spreading for the RCR and RCR Conditions in Study Two

Measures of Spread	Condition	
(Participants included)	Rank-Choose-Rank	Rank-Rank-Choose
Chosen Spread:	$^{a}1.94_{a,b}$	$.94_{a}$
(All Participants)	$(2.51, 50)$	$(2.77, 50)$
7-9 Spread:	$^{b}1.31_{a,b}$	$.61_{a}$
(If Chose 7)	$(2.04, 39)$	$(2.28, 38)$
9-7 Spread:	4.18_{c}	$2.00_{b,c}$
(If Chose 9)	$(2.82, 11)$	$(3.86, 12)$
7-9 Spread:	$-.10_{a,d}$	$^{b} - .02_{d}$
(All Participants)	$(3.18, 50)$	$(2.92, 50)$

Note. Standard deviations and number of participants are presented in parentheses (SD, n). Across conditions, means that share a subscript letter are not significantly different from one another at $p < .05$.

a) To calculate dissonance reduction, traditional paradigms would compare chosen spread in the RCR condition to 0.

b) Or, traditional paradigms would compare 7-9 spread for those who chose 7 in the RCR condition to all participants in the RRC condition. Both traditional measurements find "dissonance reduction." The correct comparisons (calculation of spread across the two conditions) provide marginal support for the notion that choice-induced attitude change exists above and beyond the spreading that is predicted from the choice information ($p = .06$ for chosen spread and $p = .16$ for 7-9 and 9-7 spread).

Appendix: Statement of Theorem and Full Proof

In this section, we extend and expand the simplified proof we presented in the text to accommodate several more complex variants of the FCP. The core structure of our argument remains unchanged; but here we show that the basic logic holds across all variants of the FCP. This includes:

1. basic control groups in which subjects move straight from Stage One to Three,
2. control groups in which subjects are not allowed to make choices but are given a randomly chosen good,
3. FCP studies that drop participants from the experimental group, and
4. FCP studies that examine choices when the goods are initially rated/ranked "close" or "far".

Also, in this expanded section we can relax two simplifying assumptions that we made in our simplified proof. First we can show that our proof holds under not just $A1$, but also the much more general $A1a$. Second, we can relax the simplifying assumption we made in equation (5), which asked that ratings/rankings did not experience what statisticians call *regression to the mean*. Here, we prove our theorem is still true even when neither of these simplifying assumptions holds; as a statistical statement, our proof holds for all but degenerate choice processes.

Restating our theorem then:

Theorem 2 *Suppose that people's choices are driven by stable preferences which satisfy* $A1a$, $A2$ *and* $A3$*. Then the free-choice paradigm will in expectation measure a greater increase in "spread" between stages* 1 *and* 3*, among experimental participants than among control participants, despite the lack of any attitude change in either group.*

Full Proof:

Just as in our simplified treatment, begin by supposing that in Stage Two of a FCP experiment a participant is asked to choose between goods i and j, where good i was the good that a person initially rated/ranked better by D rating/ranking points. Written mathematically, this is to say:

$$r_i^1 = r_j^1 + D$$

where since i is the good initially rated/ranked better, we have $D > 0$ (without loss of generality). The FCP is going to look at what good the person chooses, and ask if the **chosen** good's rating/ranking rises and the **unchosen** good's rating/ranking falls. To aid in the exposition of our full proof, we define the random variable which is of primary interest, the "*Chosen Spread*" between goods i and j as $R_{\{i,j\}}$, where:

Notation Four: Chosen Spread.

$$R_{\{i,j\}} = \left\{ \begin{array}{lll} (r_i^3 - r_i^1) - (r_j^3 - r_j^1) & \text{if} & c_{\{i,j\}}^2 = x_i \\ (r_j^3 - r_j^1) - (r_i^3 - r_i^1) & \text{if} & c_{\{i,j\}}^2 = x_j \end{array} \right\} \quad (13)$$

or equivalently:

$$R_{\{i,j\}} = \left\{ \begin{array}{lll} (r_i^3 - r_j^3) - D & \text{if} & c_{\{i,j\}}^2 = x_i \\ (r_j^3 - r_i^3) + D & \text{if} & c_{\{i,j\}}^2 = x_j \end{array} \right\}$$

Note that this is just the amount good i goes up plus the amount good j goes down (if the subject chooses good i), and vice-versa if the subject chooses good j.

Mathematically then, what the FCP measures for participants in the experimental condition is:

$$E[R_{\{i,j\}}] \quad (14)$$

Some FCP papers compare the quantity in (14) to 0, testing whether the spread of the chosen good over the unchosen good (which was D), increases for participants who make choices in Stage Two. That is, these papers test the simple hypothesis that $E[R_{\{i,j\}}] > 0$. This is the simplest version of a FCP, which we deal with in the text.

Other FCP experiments include a control group, and compare the change in spread for participants who make a choice in Stage Two to the rating/ranking changes of participants in a control group who were not asked to make a choice. We will now extend our proof to cover these control groups.

The FCP with a Basic (no-choice, no-good) Control Condition

There are several possible ways to calculate a measure of spread for "control" participants. Most commonly, this is done by computing the change in spread of goods i and j. In this case the FCP is comparing the quantity in (14), or "*chosen spread*" to:

$$E[(r_i^3 - r_j^3) - D] \tag{15}$$

Which is just the change in "spread" between goods i and j for participants in the control group.

Remark 3 *The use of a control group in the full proof allows us to relax two simplifying assumptions we made in our simpler proof.*

Recall that in our initial treatment of this proof, we made two simplifying assumptions.

First, we assumed that ratings/rankings were an unbiased measure of participants' preferences. $A1$ is now relaxed to allow for any strictly monotonic relationship between ratings/rankings and preferences, as expressed in $A1a$.

Second, we assumed that the amount of "spread" that the experimenter should expect goods i and j to display in Stage Three is equal to the amount of spread they displayed in Stage One. Specifically, in equation (5) we asserted that:

$$A1 \quad \Rightarrow \quad E[(r_i^3 - r_j^3)] = D \tag{16}$$

This equality is a simplifying assumption, and will not hold strictly, even for people with perfectly stable preferences. In general, we should actually expect:

$$A1 \quad \Rightarrow \quad 0 \leq E[(r_i^3 - r_j^3)] < D \tag{17}$$

This is because if people's preferences are stable but their rankings express these preferences imperfectly ($A3$), then the statistical phenomena known as *regression to the mean* would suggest that the rating or ranking distance between any two goods should be expected to contract between stages One and Three. Because much of the extra complexity of this full proof owes to this fact, a brief discussion of it here may be warranted.

Regression to the Mean. Regression to the mean has been understood to be an important consideration in both experimental design and data analysis since it was first noted by Galton (Galton 1886). More recently, Tversky and Kahneman identify regression as one of the most fundamental statistical phenomena about which people do not develop correct intuitions (Tversky & Kahneman 1974). Put simply, for any two distributions X_1 and X_2 with identical marginal distributions, if one selects observations from X_1 which deviate from the mean of X_1 by k units, then on average these observations will deviate from the mean of X_2 by less than k units. In general, regression to the mean will be present in any comparison of related measures; the heights of parents and their children, in the performance of students on successive tests, and (most importantly for our purposes) in the repeated measurement of preferences. Intuitively, since goods i and j were initially rated/ranked respectively above and below the mean, the subsequent rating/ranking of i should be expected to drop while the subsequent rating/ranking of j should be expected to rise, even if people's preferences remain perfectly stable. This implies the strict inequality in equation (17).[21]

In practice, most FCP experiments have tried to choose goods that were very close together; i.e. for which D is very close to zero, and for which our simple proof is sufficient. Many FCP, experiments however, compare chosen spread to the spread observed in a control group. This requires our proof to formally account for the regression to the mean that will be observed in both groups; i.e. equation (17) will hold for both the control and experimental groups.

Comparing Experimental and Control Participants. To take this regression into account in our proof, call the reduction in "spread" we expect to see between goods i and j in Stage Three, $R > 0$. That is, let:

$$R = D - E[(r_i^3 - r_j^3)]$$

so that:

$$E[(r_i^3 - r_j^3)] = D - R \tag{18}$$

Remark 4 *We will show that under assumptions* A1a, A2, *and* A3 *people who choose in Stage 2 will display spreading of ratings/rankings greater than $D - R$, even when their attitudes remain completely stationary.*

Remember from (14) that spread for experimental participants is calculated by how much i goes up and j goes down for participants who choose i and how much j goes up and i goes down for participants who choose j. Note that we can usefully expand (14) to either of its values, depending on what a participant chooses. That is, we can rewrite what the FCP will measure for participants in the experimental group as a probability weighted sum of expected ratings:

$$\begin{aligned} &E[R] = \\ &\Pr(c^2_{\{i,j\}} = x_i) * E[(r_i^3 - r_j^3) - D | c^2_{\{i,j\}} = x_i] + \\ &\Pr(c^2_{\{i,j\}} = x_j) * E[(r_j^3 - r_i^3) + D | c^2_{\{i,j\}} = x_j] \end{aligned} \tag{19}$$

Now, to facilitate comparison we would also like to write out the expected spread for participants in the control condition as a probability weighted sum of expected ratings/rankings depending on which choice they would make. Unlike for participants in the experimental condition who make a choice, spread for participants in the control condition is always calculated based on how much i goes up and j goes down. Thus, we can re-write the left side of (18) by expanding it out into the two possible choices a person would make if asked (remember, though, that control participants do not actually make this choice). Therefore, for participants in the control condition, we have that:

$$\begin{aligned}&E[(r_i^3 - r_j^3)] = \\ &\Pr(c^2_{\{i,j\}} = x_i) * E[(r_i^3 - r_j^3)|c^2_{\{i,j\}} = x_i] + \\ &\Pr(c^2_{\{i,j\}} = x_j) * E[(r_i^3 - r_j^3)|c^2_{\{i,j\}} = x_j] \\ &= D - R\end{aligned} \tag{20}$$

Note that (20) is both the spread an FCP experiment will measure for control participants, and a mathematically correct statement about expected spread between goods i and j, for participants in either condition.

To compare these two spreads (equations 19 and 20), note that we can say something about the probability that participants in either the experimental or the control condition will (or would) choose either object. That is, since experimental and control participants are ex-ante identical, we can write one set of equations for both sets of participants, conditioning on the choices that experimental participants will make, and on the choices control participants would make. Since they initially rated/ranked i better than j,

$$A1a \ \& \ A2 \quad \Rightarrow \quad 1 > \Pr(c^2_{\{i,j\}} = x_i) > \Pr(c^2_{\{i,j\}} = x_j) > 0 \tag{21}$$

Next, we would like to be able to say something about expected spreads. To do this, note that some people will (would) choose i over j, while other people will (would) choose j. Therefore if choices tell us something about preferences ($A2$), then those who will (would) choose i will on average like i more than those who will (would) choose j, and vice versa for j. That is:

$$A2 \quad \Rightarrow \quad E[u_i|c^2_{\{i,j\}} = x_i] > E[u_i|c^2_{\{i,j\}} = x_j] \tag{22}$$

and:

$$A2 \quad \Rightarrow \quad E[u_j|c^2_{\{i,j\}} = x_i] < E[u_j|c^2_{\{i,j\}} = x_j] \tag{23}$$

Now, as long as ratings/rankings are monotonically related to preferences and ratings/rankings are a noisy and unbiased measure of preferences, we can move from statements about preferences (equations 22 and 23) to statements about expected ratings:

$$A1a \;\&\; A3 \quad \Rightarrow \quad E[r^3_i|c^2_{\{i,j\}} = x_i] > E[r^3_i|c^2_{\{i,j\}} = x_j] \tag{24}$$

and:

$$A1a \;\&\; A3 \quad \Rightarrow \quad E[r^3_j|c^2_{\{i,j\}} = x_i] < E[r^3_j|c^2_{\{i,j\}} = x_j] \tag{25}$$

Combining these last two expressions, we can pin down the relative size of the two conditional expected ratings/rankings spreads in (20):

$$A1a \;\&\; A3 \quad \Rightarrow \quad E[(r^3_i - r^3_j)|c^2_{\{i,j\}} = x_i] > E[(r^3_i - r^3_j)|c^2_{\{i,j\}} = x_j] \tag{26}$$

Recall from (20) that the weighted average of these two spreads was equal to $D - R$. Therefore, we know that the larger must be bigger than $D - R$ and the smaller must be less than $D - R$. Mathematically:

$$E[(r^3_i - r^3_j)|c^2_{\{i,j\}} = x_i] > D - R > E[(r^3_i - r^3_j)|c^2_{\{i,j\}} = x_j] \tag{27}$$

Looking at the first part of this inequality and subtracting D from both sides gives us:

$$E[(r^3_i - r^3_j) - D|c^2_{\{i,j\}} = x_i] > -R \tag{28}$$

Similarly, looking at the second part the inequality gives us:

$$E[(r^3_j - r^3_i) + D|c^2_{\{i,j\}} = x_j] > R > -R \tag{29}$$

Combining the inequalities from (28) and (29) and using (21), tells us that the spread the FCP will measure for experimental participants (equation 19) is strictly bigger than $-R$, because it is the weighted sum of two things which are both bigger than $-R$. But recall that $-R$ was how much increase in spread the control group was going to display (18). Therefore, we have shown that the experimental group will show more spread than the control group. That is to say, we should expect the ratings/rankings of the chosen good to rise and the rating/ranking of the unchosen good to fall more than they do in the control group, even if by assumption no person's preferences have moved.

FCP Studies with an Expanded (No-Choice, Random-Good) Control Conditions

This proof can be modified to accommodate other control groups found in the literature. For example, some FCP studies employ an expanded control-group. These participants are not asked to make choices in Stage Two, but ARE given goods in Stage Two. In contrast to experimental participants who are allowed to choose, these control participants are randomly given either good i or good j. To accommodate this we introduce notation for what good a particular participant gets assigned:

Notation 5: Randomly Assigned Goods. Denote by $\mathrm{rand}^2_{\{i,j\}}$ the good a person will randomly given from the set $\{x_i, x_j\}$ in Stage 2.

FCP studies with a random-good control condition begin by computing "chosen spread" for experimental participants. This is just $E[R_{\{i,j\}}]$, as computed in equation (13). Instead of being compared to zero however, these FCP studies compare $E[R_{\{i,j\}}]$ to:

$$E[\left\{\begin{array}{ll} (r_i^3 - r_i^1) - (r_j^3 - r_j^1) & \text{if} \quad \mathrm{rand}^2_{\{i,j\}} = x_i \\ (r_j^3 - r_j^1) - (r_i^3 - r_i^1) & \text{if} \quad \mathrm{rand}^2_{\{i,j\}} = x_j \end{array}\right\}] = -R' \tag{30}$$

This control condition (though far less common) not only accounts for regression to the mean in ratings/rankings (like in equation 15) but also controls for any change in spread arising from a mere ownership or exposure (for example, the endowment effect).

Extending our proof to this case is straightforward. For FCP which employ this type of control group we simply need to compute an expected spread for each of the two conditions. Specifically, we would let:

$$E[\ (r_i^3 - r_i^1) - (r_j^3 - r_j^1) \quad \text{if} \quad \text{rand}^2_{\{i,j\}} = x_i\ ] = -R^1 \tag{31}$$

and:

$$E[\ (r_j^3 - r_j^1) - (r_i^3 - r_i^1) \quad \text{if} \quad \text{rand}^2_{\{i,j\}} = x_j\ ] = -R^2 \tag{32}$$

Then we would repeat the argument we made for the basic control group (equation 18 through equation 29), weighting the two means arrived at in the two sub-groups (control participants who received x_i vs. those who received x_j) by the shares of experimental participants who chose x_i vs. x_j. For each of these groups, the proof then goes through unchanged, and we would arrive at an expression like (29) for each group.

FCP Studies which Drop Participants

The oldest FCP variant (and the one used in the original Brehm, 1956) drops participants who choose the good initially rated/ranked lower, then compares the spread among the remaining participants to the spread of i over j for participants in the control condition. This is also problematic. Mathematically, what this original FCP variant asks is if for the experimental group:

$$E[\ (r_i^3 - r_j^3) \quad \text{if} \quad c^2_{\{i,j\}} = x_i\ ] > D - R \tag{33}$$

This is exactly what we derived in (28). Intuitively, this is going to be true whenever those participants who chose good i will on average like good i more than those who choose good j.

High and Low Dissonance Treatments

In several FCP experiments, an ancillary prediction of dissonance theory is tested by looking at choices between goods that are “close” in rating/ranking and goods that are "far", with dissonance theory suggesting that dissonance will only be aroused in the former. For example, the original Brehm study looks at choices made between items around one point apart on an eight-point scale (what Brehm calls a high-dissonance condition), two points apart (medium-dissonance) and three points apart (low-dissonance).

Brehm (1956) and many studies that have followed have tended to find that "chosen spread" is larger when goods are close, and smaller when goods are far (see for example Brehm 1956, Brehm & Cohen 1959, Brock 1963, Greenwald 1969, Oshikawa 1971, Shultz & Lepper 1992, and Shultz, Léveillé, & Lepper 1999). Indeed, this finding of greater spread in "high" vs. "low" dissonance conditions has motivated subsequent formal modeling of dissonance processes (see for example the "consonance model" of constraint satisfaction, Shultz & Lepper 1992, 1996). While this is an interesting ancillary prediction of dissonance theory, note that this difference is also expected to arise in FCP experiments absent any attitude change, for two principle reasons.

First, when goods are initially far apart in ratings/rankings, our earlier discussion of regression to the mean suggests these goods will display less "spread", even if preferences are perfectly stable (since regression produces negative spread). This problem has been discussed in the literature since at least Oshikawa (1971).

As a consequence of our theorem however, a second, more fundamental problem arises. Since choices made between goods initially rated/ranked "far" goods are less likely to overturn people's initial rating/ranking, they are less subject to the non-random selection problem our theorem identifies.

In other words, because almost all participants will choose the item initially rated/ranked higher when the goods are initially far apart, there is little to no information provided by choice. As this selection problem disappears, so does the "spread" it produces; hence "far" goods should be expected to show less spread, even if preferences are perfectly stable.

To see this mathematically, assume that two pairs of goods are chosen such that in stage 1, goods i and j are "close", and goods k and l are "far", so that:

$$r_l^1 < r_j^1 < r_i^1 < r_k^1 \tag{34}$$

Now, a high-low dissonance comparison will ask if the high-dissonance pair (goods i and j) will show more "spread" than the low-dissonance pair (goods k and l). Mathematically then, what this comparison tests is if:

$$E[R_{\{k,l\}}] < E[R_{\{i,j\}}] \tag{35}$$

Remark 5 *What we will show is that for sufficiently "far" apart goods k and l, equation (35) will hold, even if preferences are perfectly stable.*

Following the same logic as our broader analysis, we can re-write (35) as:

$$\begin{gathered}
\Pr(c_{\{k,l\}}^2 = x_k) * E[(r_k^3 - r_l^3) - D^{kl} | c_{\{k,l\}}^2 = x_k] + \\
\Pr(c_{\{k,l\}}^2 = x_l) * E[(r_l^3 - r_k^3) + D^{kl} | c_{\{k,l\}}^2 = x_l] \\
< \\
\Pr(c_{\{i,j\}}^2 = x_i) * E[(r_i^3 - r_j^3) - D^{ij} | c_{\{i,j\}}^2 = x_i] + \\
\Pr(c_{\{i,j\}}^2 = x_j) * E[(r_j^3 - r_i^3) + D^{ij} | c_{\{i,j\}}^2 = x_j]
\end{gathered} \tag{36}$$

where we augment our earlier notation so that:

$$r_i^1 - r_j^1 = D^{ij} \quad \text{and} \quad r_k^1 - r_l^1 = D^{kl}$$

As in our previous analysis, for both pairs of goods we expect a reduction in "spread" since the initial ratings/rankings of the higher ranked goods i and k should be expected to fall, and the initial ratings/rankings of j and l should be expected to rise (see our earlier discussion of regression to the mean). We would also expect that ratings/rankings that started further apart (goods k and l) would regress further than initially closer goods (i

and j). Augmenting the notation we introduced in (18) to denote the amounts these pairs will regress, let:

$$E[(r_i^3 - r_j^3) - D^{ij}] = -R^{ij} \quad \text{and} \quad E[(r_k^3 - r_l^3) - D^{kl}] = -R^{kl} \tag{37}$$

which allows us to write our expectation of differential regression as:

$$0 < R^{ij} < R^{kl} \tag{38}$$

We can now combine both of these insights to see why far choices will produce more spreading than close choices, even if preferences remain stable.

The simplest way to see this is to think about what will happen if goods k and l are chosen far enough apart so that the selection problem we identified no longer applies; that is we choose k and l such that virtually all subjects choose the initially higher rated/ranked good k when offered the choice. That is, if k and l are chosen to be "far" enough apart, we should see that:

$$\Pr(c^2_{\{k,l\}} = x_l) \to 0 \quad \text{and} \quad \Pr(c^2_{\{k,l\}} = x_k) \to 1 \tag{39}$$

Note that in this limit, the left hand side of equation (36) becomes much simpler:

$$\begin{gathered} \Pr(c^2_{\{k,l\}} = x_k) * E[(r_k^3 - r_l^3) - D^{kl} | c^2_{\{k,l\}} = x_k] + \\ \Pr(c^2_{\{k,l\}} = x_l) * E[(r_l^3 - r_k^3) + D^{kl} | c^2_{\{k,l\}} = x_l] \\ = \\ E[(r_k^3 - r_l^3) - D^{kl} | c^2_{\{k,l\}} = x_k] \\ = \\ E[(r_k^3 - r_l^3) - D^{kl}] \end{gathered} \tag{40}$$

with this last equality holding because conditioning on an event carries no information as the probability of that event goes to 1.

Similarly, we can also say something about the right hand side of equation (36), by applying our main theorem. That is, by our main theorem we should see positive "spread" between our “close” items i and j; mathematically this is to say:

$$\begin{gathered} E[(r_i^3 - r_j^3) - D^{ij}] \\ < \\ \Pr(c_{\{i,j\}}^2 = x_i) * E[(r_i^3 - r_j^3) - D^{ij} | c_{\{i,j\}}^2 = x_i] + \\ \Pr(c_{\{i,j\}}^2 = x_j) * E[(r_j^3 - r_i^3) + D^{ij} | c_{\{i,j\}}^2 = x_j] \end{gathered} \tag{41}$$

But note that combining expressions (38), (40), and (41) leads us to (36), proving our proposition. Intuitively, if participants are given a choice between alternatives that are far enough apart, virtually all participants will choose the higher rated/ranked item; therefore there will be no selection-induced spreading. In contrast, if participants are given a choice between alternatives that are close together, most will choose the higher rated/ranked item but many will choose the lower rated/ranked item. We have shown that this induces non-random selection and spreading. Therefore FCP experiments will find more spread among “close” than among “far" choices, even if neither close nor far choices induces any attitude change.

Author Note

The two authors contributed equally to the manuscript and the authorship order was determined by a coin flip. Chen was responsible for developing the formal proof. Risen was responsible for the experimental design. We thank Kyle Schmidt and Jackson Lai for their help with data collection. In addition, we thank Judy Chevalier, Tom Gilovich, Barry Nalebuff, David Nussbaum, Benjamin Polak, and Dennis Regan for providing feedback on an earlier draft.

Notes

[1]Although Brehm interpreted his results from the perspective of dissonance theory, to our knowledge, dissonance theory (Festinger, 1957; 1964) and self-perception theory (Bem, 1967) have never been pitted against each other in the free-choice paradigm. From the self-perception perspective, a chosen alternative may become more desired and the rejected alternative less desired because people learn about their own preferences by observing their choice (Bem, 1967). Note that both dissonance theory and self-perception theory predict that the mere act of choosing induces attitude change. In other words, for both theories, spreading is a reaction to the choice process.

[2]Brehm's original (1956) experiment examined the spread of ratings following choice, but many subsequent studies have used a simpler ranking procedure. For the purposes of our analysis the two forms of evaluation are equivalent, and chosen spread has been documented in FCP studies that use both rating and ranking procedures. Thus, throughout the paper we use the terms "spread," "chosen spread," "high-low spread," or "spreading of alternatives," when we are describing the change of evaluation for chosen and rejected items in rating or ranking.

[3]For example, 21% of participants in Brehm's (1956) study, 21% of participants in Gerard and White's (1983) study, and 36% of participants in Lieberman et al.'s (2001) study chose the lower ranked item.

[4]We are not the first to object to the procedure of excluding participants who choose the lower ranked item. Previous complaints have been mostly statistical in nature (see Chapanis & Chapanis, 1964; East, 1973). We believe that the statistical problem arises because choice is treated as if it were random. We contend that the problems associated with throwing out participants and comparing spread to 0 arise from the common mistake of assuming that there is no important information revealed by the choice. And, we hope that by providing data to support our objection, ours will have more staying power than

those made in the past.

[5]The first three cases assume that measurement error accompanies the elicitation of preferences. Our formal proof deals directly with these first three cases. According to this perspective, if you choose j after ranking i higher than j, your preference for i over j has likely always been smaller than for someone who chose i. Instead, if reversals are due entirely to highly unstable preferences, then we would assume that you liked i more than j initially (as much as anyone else who ranked i above j), but that you changed your mind before the choice. Although these assumptions are not developed in the formal proof in the paper, it should be clear why this set of assumptions would also predict spreading that is not due to choice-induced attitude change. If reversals occur because people change their mind before the choice, then we would predict more *7-9 spread* for those who did not change their mind and more *9-7 spread* for those who did change their mind. Note that the underlying change of preference has happened before the choice and is not a reaction to the choice.

[6]To see that our two formulations are mathematically equivalent, note that our first formulation is equivalent to $\Pr[r_k^1 > r_l^1 | u_k > u_l] > 1/2$. Bayes' rule lets us rewrite the left hand side of this inequality as $\Pr[u_k > u_l | r_k^1 > r_l^1] * \Pr[r_k^1 > r_l^1] / \Pr[u_k > u_l]$. But since goods k and l refer to any two goods, we know what our two prior probabilities must be: $\Pr[r_k^1 > r_l^1] = \Pr[u_k > u_l] = 1/2$. Therefore $\Pr[u_k > u_l | r_k^1 > r_l^1] > 1/2$, which is our second formulation of assumption 1a.

[7]Note that we have chosen to follow a ratings-like convention where more attractive goods have higher evaluations, rather than a ranking convention (in which a "1" if often better than a "2").

[8]Formally, our proof treats rankings, ratings, choices, and preferences as random variables, and focuses on the mathematical expectation of these (and combinations of these) variables. Intuitively, these mathematical expectations can be thought of as the

best prediction of what an experiment would find in a random set of participants. In other words, imagine that for any FCP experiment, there is a population of possible participants and that participants are randomly drawn from this population. Then, our mathematical expectation for the rating of a good can be thought of as the mean rating of that good in the whole population, and therefore as, on average, the mean rating of that good in any particular FCP experiment.

[9]Importantly, most ranking systems assign more attractive goods lower rank numbers. Therefore, if we were to rewrite our proof to replace ratings statements such as $r_k^1 > r_l^1$ ("good k is rated better than good l") with rankings statements such as $r_k^1 \succ r_l^1$ ("good k is ranked better than good l"), then our assumption $A1$ would need to be inverted, to $E[r_k^t|u_k] = -u_k$ since by convention higher utilities refer to preferred items.

[10]This will not be exactly true if ratings display regression to the mean; our analysis still holds if this is the case but the proof becomes more complex (see the Appendix of this paper for a proof which relaxes this assumption). This simplification makes sense because in most FCP experiments goods i and j are chosen to be average-rated goods, for which regression will be minimal.

[11]Although orders CBA, BAC, and BCA are contradicted by the initial choice of Apple over Banana, if choice and underlying preference are not perfectly related, then those orders are not entirely eliminated. Instead, those orders should be thought of as far less likely to reflect actual preferences compared to the other three orders. Also note that the six orders could be expanded to include partial orders (e.g., A is preferred to B and C, which are equally liked). If we expand to include all partial orders, our argument still holds. For a more complete discussion of this issue, see Chen and Risen (2009) and Sagarin and Skowronski (2009).

[12]The belief that the null for choosing B in the choice condition is 50% rather than 33% is analogous to the confusion often seen for the Monty Hall problem. In Monty Hall,

people fail to recognize that switching doors will allow them to win 2/3 of the time because they fail to recognize that Monty is not revealing doors at random. Instead, he is selecting the door based on the fact that it does not have the grand prize behind it. Similarly, people are not randomly choosing A over B – they are selecting it based on their preference.

[13]A critical element for interpreting any differences between a RCR and RRC condition is understanding what information is revealed by the choice in either condition, and how this information is related to the information revealed by the first and second ranks. In the text, we provided evidence for RCR and RRC participants making similar choices. To test whether RRC serves as the right control, however, we need to be somewhat more precise. So, for example, if in a RCR condition the first and second ranks have approximately equal predictive power in predicting a participant's choice, then for the RRC condition to be an adequate control condition, choices must also be equally well explained by the first and second ranks. Intuitively, this kind of "time invariance" would allow us to conclude that choices were revealing similar information in either design, and would allow a cleaner interpretation of any differences in spread that remained. Tests of relative informativeness require more detailed modeling of the choice process, and lie outside the purview of this paper; they are straightforward in principle though.

[14]For a brief discussion of comparing relative spreading, please see the previous footnote.

[15]Specifically, a preference-driven theory of behavior would not predict a conditional difference in spread for RRC and RCR participants if the model assumed that preferences were perfectly stable and assumed that "errors" in ratings and choice (i.e., the difference between participants' ratings/choices and their underlying preferences) were stationary and i.i.d. across time. A preference-driven theory of behavior could be modified to predict a conditional difference if either of these assumptions were relaxed. For example, a

conditional difference would be expected by a preference driven-model if it incorporated either dissonance reduction or self-perception (i.e., learning from previous rankings, ratings, and choices).

[16]In Linder, Cooper, and Jones (1967), all participants who learned about the essay agreed to write it, and in Gerard and Mathewson (1966) all participants agreed to the group initiation. One participant in one of Aronson and Mills' (1959) initiation conditions did not agree to participate.

[17]Reacting to the concern initially described by Chen (2008) and developed more fully here, Egan, Santos, and Bloom adapted their paradigm so that all choices are made while goods are stuffed in stockings (see Egan, Bloom, & Santos, 2010). Although we still have concerns with some of the details of the methodology, this paper represents an important step forward from earlier work. See Risen and Chen (2010a) for a discussion of experimental methods that researchers can use to properly study choice-induced attitude change.

[18]Admirably, rather than simply comparing spread to 0, the authors compared spread in the critical choice set to spread in the control set. However, because participants were never asked to choose between prints in the control set, there was no choice information for the control. Instead, if participants chose the higher ranked item in the critical set, they were assumed to choose the higher ranked item in the control set. In that case, chosen spread was calculated as the amount the higher ranked item went up and the lower ranked item went down. If participants reversed their choice in the critical choice (choosing the lower ranked item), they were assumed to have reversed the other choice as well, and chosen spread was calculated as the amount the lower ranked item went up and the amount the higher ranked item went down. But, this analysis relies on the assumption that reversing is systematic in people (e.g., if Jon reverses in one set and Julie does not, it means that he is an ambivalent person – he will reverse in another set and Julie will not).

Rather, we believe that reversing is systematic of people's preferences for those particular items (If Jon reverses in one set and Julie does not, it means that he is more ambivalent about these items than Julie is – he is no more or less likely to reverse a different choice than Julie is).

[19]It is possible to address some of our concerns using existing data. For example, for a study that compared affirmed and non-affirmed participants in a Rank-Choose-Rank design and found less spreading in the affirmed condition, there are certain analyses that would help rule out an explanation based on a preference-driven model of choice and help support a dissonance explanation. Intuitively, if participants in different experimental conditions do not differ in either how stable their preferences are, nor how reliably those preferences are expressed by their rankings, rating, and choices, then a preference-driven model of choice would not predict differences across conditions. So for example, a preference-driven explanation would be less convincing as an alternative if affirmed and non-affirmed participants were equally likely to choose their 7th or 9th ranked item. A second test that could help determine if a preference-driven account could explain the results would be to see if the rankings and re-rankings were correlated to an equal extent for both conditions. To test the correlations, one would need to remove the 2 items in the choice set from Ranking 1 and Ranking 2 and adjust the ranks of the remaining items accordingly (i.e., from 1 to 13). If participants in the two conditions were equally likely to choose the higher or lower ranked item and if the average Spearman rank correlation for the 13 items was similar across conditions, this would make a preference-driven account less plausible. Of course, a cleaner way to rule out the preference-driven alternative is to attempt to control for the choice information (as we did in the current paper).

[20]To see this formally, note that if some pattern of neural activation predicts which goods a subject will subsequently choose, but does not predict choices perfectly, then our assumptions 1 and 3 apply to this neural activation just as they would to any more

conventional method of measuring attitudes or preferences. Only if a pattern of neural activation *never* disagreed with the choices a subject subsequently makes, would it not be subject to revealed-preference confounds.

[21]To see this formally, consider the generalized "reversion to the mean" theorem in Samuels (1991). There, Samuels shows that if X_1 and X_3 are random variables with mean μ and the same marginal distribution, then for any constant C, $\mu \leq E[X_3|X_1 > C] < E[X_1|X_1 > C]$, subject to the nondegeneracy condition $\Pr[X_3 > C|X_1 > C] < 1$. To see how this theorem applies, for any ordered pair of goods (i, j), let X_1 and X_3 be the distribution of $r_i^1 - r_j^1$ and $r_i^3 - r_j^3$, respectively. If preferences are stable then X_1 and X_3 are likely to have the same marginal distribution (indeed, in ranking experiments X_1 and X_3 are forced to have the same marginal distribution). Nondegeneracy follows from assumption $A3$. Therefore the Samuels result applies, and equation (17) follows immediately.