RATIONALIZATION AND COGNITIVE DISSONANCE: DO CHOICES AFFECT OR REFLECT PREFERENCES?

By

M. Keith Chen

July 2008

COWLES FOUNDATION DISCUSSION PAPER NO. 1669

Rationalization and Cognitive Dissonance: Do Choices Affect or Reflect Preferences?

M. Keith Chen*

First draft: *November, 2007*
This draft: *January, 2008*

One Sentence Summary:
A methodological flaw may have led one of the central literatures in social psychology to spuriously conclude that people rationalize past choices, by failing to appreciate that those choices reflect people's preferences.

Abstract:
Cognitive dissonance is one of the most influential theories in social psychology, and its oldest experiential realization is choice-induced dissonance. Since 1956, dissonance theorists have claimed that people rationalize past choices by devaluing rejected alternatives and upgrading chosen ones, an effect known as the spreading of preferences. Here, I show that every study which has tested this suffers from a fundamental methodological flaw. Specifically, these studies (and the free-choice methodology they employ) implicitly assume that before choices are made, a subject's preferences can be measured perfectly, i.e. with infinite precision, and under-appreciate that a subject's choices reflect their preferences. Because of this, existing methods will mistakenly identify cognitive dissonance when there is none. This problem survives all controls present in the literature, including control groups, high and low dissonance conditions, and comparisons of dissonance across cultures or affirmation levels. The bias this problem produces can be fixed, and correctly interpreted several prominent studies actually *reject* the presence of choice-induced dissonance in their subjects. This suggests that mere choice may not be enough to induce rationalization, a reversal that may significantly change the way we think about cognitive dissonance as a whole.

Introduction:

Cognitive dissonance is the theory that people are motivated to reduce the dissonance caused by conflicting cognitions, often by modifying their attitudes (Festinger 1957). It is one of the most influential theories in social psychology, and having produced over fifteen-hundred papers, is often cited as a central pillar of the field (Jones 1985, Ross & Nisbett 1991, Harmon-Jones & Mills 1999). The oldest experimental demonstration of cognitive dissonance is Brehm's seminal 1956 study, which claimed that the mere act of choosing between goods is enough to induce dissonance (Brehm 1956). Brehm suggested that subjects rationalize choices they make by shifting their preferences, subsequently viewing chosen alternatives as more desirable than before, and rejected alternatives as less desirable. Brehm's experimental method came to be known as the free-choice paradigm, and began an extensive literature on choice-induced cognitive dissonance (for a good review see Harmon-Jones & Mills 1999).

Here, I show that this literature operates under a subtle yet powerful assumption: that subject's preferences can be measured perfectly, i.e. with infinite precision. If this fails, then these papers test the wrong null hypothesis and will conclude that cognitive dissonance is present even when it is not. None of the controls present in the literature address this concern, including comparisons with control groups, comparisons between easy and hard decisions, or comparisons across cultures or affirmation conditions. Indeed, upon reexamination several of the most prominent papers in this literature can be taken to reject the presence of any significant amount of cognitive dissonance in their subjects, as their results are very close to the correct null hypothesis of their tests. Therefore, I show that there is as of yet, no conclusive evidence that mere choice is enough to induce cognitive dissonance.

The intuition behind the problem I identify is simple. In Brehm's classic free-choice paradigm (hereafter FCP) and its modern variants, subjects are asked to choose between two options. The object of study is their preferences for those options before and after this choice. These papers claim that any increase in the measured preference for chosen items over unchosen items is evidence of cognitive dissonance (hereafter CD). They fail to account, however, for the fact that if preferences cannot be measured perfectly before a choice is made, then a subject's choices contain additional information about their preferences. Specify, when a subject chooses an object we learn that they prefer it to what they rejected, and we should expect that subsequent measures of their preferences will reflect this. Every study in this literature has failed to consider this possibility, and has relied critically on the assumption that after initially surveying subjects, a subject's choices teach us nothing. I show that this is both highly unlikely, and rejected by several aspects of these data; further the failure to account for this can explain all findings of cognitive dissonance in the literature both qualitatively and quantitatively.

This failure is significant for several reasons. First, if mere-choice does not induce CD, it would significantly change how we think about the scope and strength of dissonance processes. Second, this failure also bears on the causes and origins of CD. CD has typically been understood either as an ego-defensive bias (Steele et.al., 1983,

1993, Aronson, Cohen & Nail 1999), or the result of people inferring their preferences from their own actions (Brehm 1956). Recently, several prominent authors have used FCP studies to argue that these explanations are incomplete, claiming that CD exists in subjects unlikely to have highly developed egos (capuchin monkeys; Egan, Santos & Bloom 2007) or to remember past choices (amnesiacs; Lieberman, Ochsner, Gilbert & Schacter 2001). These authors have argued that CD must arise from primitive, more automatic processes. The ability of the FCP to find CD where there is none suggests an alternative explanation for these findings.

This paper proceeds as follows. I begin by discussing the two forms of FCP papers in the literature. The first, (less common) form looks at shifts in choices that subject's make. I describe the procedure they use, and explain why they have found cognitive dissonance even when their data do not support that claim. Somewhat counter-intuitively, strong (false) findings of cognitive dissonance persist in these papers, disappearing (discontinuously) only when preferences are measured *perfectly*. The second form of FCP paper looks at changes in how subjects rate or rank items before and after making choices. While less obviously problematic than the first type, I show that these papers suffer from a similar problem. Finally, I discuss the implications of these findings for the literature, and suggest corrected procedures that can be used to test for mere-choice-induced cognitive dissonance.

Free-Choice Paradigms that Examine Shifts in Choices:

In one of the simplest forms of a FCP, the object of study is the shifts in a subject's choices. In a recent FCP of this type (Egan, Santos & Bloom 2007), the experiment begins with subjects rating a number of objects on a five-point scale. Then, three objects that are rated equally (say rated 4) are chosen for use in a second stage of the experiment. Note, importantly, that the discreteness of the scale leaves open the possibility that these items might not be perfectly equivalent; for example, a subject may "truly" rate one of the items 4.1, one 4.26, and one 4.3.

In a second stage then, a subject is asked to choose between a *randomly* chosen two of these items, say A and B. Calling the object which the subject chooses A, the subject is then asked to choose between B (the initially rejected item), and C (the third item that was rated 4). If subjects are more likely to choose C than B in this choice, they are said to suffer from CD.

I argue that this was to be expected in subjects with no CD. In fact, subjects should be *expected* to choose good C 66% of the time. To see this, consider every possible (strict) preference ordering between three goods, listed in Table One.

Table One: All possible preference orderings

	Case 1:	Case 2:	Case 3:	Case 4:	Case 5:	Case 6:
Best:	A	A	B	B	C	C
Middle:	B	C	A	C	A	B
Worst:	C	B	C	A	B	A
Cases where A is preferred to B are shaded. In these cases:					
B vs. C:	B	C			C	

Note that if every possible preference profile is equally likely (which must be true if A, B, and C are chosen randomly from goods that rate 4), then after a subject has revealed themselves to prefer A over B, we should in fact predict that they would choose C over B two-thirds of the time. Therefore, any paper which tests cognitive dissonance by testing that C is chosen more than B will spuriously find cognitive dissonance. Indeed, the belief that C and B should be equally chosen in the third round is mathematically equivalent to a well-known logical fallacy, popularly known in economics as the three-door (or Monty-Hall) problem (for an excellent summary of the problem, see Nalebuff 1987). Fundamentally, even though subjects who are asked to make choices were randomly chosen, *which good they choose* is not random, and induces a bias in comparisons with the unchosen alternative.

In a recent prominent and widely heralded study, Egan, Santos & Bloom use this technique to claim that both four-year old children and capuchin monkeys strongly display CD. Instead, their results actually *reject* that either children or capuchin monkeys display any significant amount of choice-induced CD. In their study, children choose C 63% of the time and capuchins choose C 60% of the time; CD would suggest numbers significantly higher than 66% (which their confidence intervals do not permit).

Scope and Generality of this Problem for Form One:

Note that the analysis above assumes that subjects are never completely indifferent between two options; that if pressed they can always decide which of two options they prefer. Economic theory suggests that this is by far the most likely case, but even if subjects can be indifferent (which a discrete rating can never show) the above analysis does not change; the computation simply becomes more difficult.

This problem with discrete measures is a very general point; in practice *all* ratings systems are discrete. For example, in his original 1956 paper Brehm asked subjects to signal their preferences by marking a continuous line, but then measured their marks to the nearest tenth of an inch.

Note also that the experiment could have initially used a ten-point scale, a hundred-point scale, or a thousand-point scale, and C would still be expected to choose C over B exactly two-thirds of the time. To see this, assume objects were initially rated on

a 100-point scale rather than a five-point scale. If the experimenter then takes three items that ranked 36, their "true" ranking could be 36.2, 36.35, and 36.4, and the exact same argument applies. This problem does not rely on preferences between A, B, and C being strong, it relies only on the fact that subjects are not completely indifferent. The bias in subjects' choices of C over B should be two-thirds *no matter how fine the grid becomes*.

Free-Choice Paradigms that Examine Shifts in Rankings or Ratings:

Most FCP papers do not rely on shifts in choices, and do not fall under the exact objection I raise above. All other FCP papers examine shifts in rankings or ratings rather than choices. This includes most of the prominent papers in this literature, including the seminal studies of Brehm (1956), Gerard & White (1983), Steele (1988), Lyubomirsky & Ross (1999), Shultz, Léveillé & Lepper (1999), Stone (1999), Lieberman, Ochsner, Gilbert & Schacter (2001), and Kitayama, Snibbe, Markus & Suzuki (2004). These papers are less obviously problematic than those of form one, but fundamentally the same problem still holds.

In a typical FCP of this second type, the experiment begins by asking subjects to evaluate a number of goods by ranking them. For example, subjects may have to rank fifteen goods, with 1 being best and 15 being worst. Then in a second stage, subjects are asked to do one of three things:

1. Experimental subjects are asked to make a "hard" decision between two goods which were closely ranked, say 7 and 9.
2. Another set of subjects are asked to an "easier" decision between two goods which were ranked further apart, say 4 and 12.
3. Another set of subjects are not asked to make any choices.

Finally, in a third stage all groups are then asked to re-rank all fifteen items. If experimental, or high-dissonance subjects (group 1) subsequently rate *the good they chose* higher than subjects who either did not have to make a choice (condition 3) or who had an easier, low-dissonance choice to make (condition 2), these papers claim that CD has occurred. All FCP papers not of the first form have tested this "spreading of alternatives" for either rankings or ratings.

Tests of this form still suffer from the same fundamental problem as the first form. Even absent CD, a subject's choices among objects will change what we (the experimenter) should expect subsequent rankings to be. This is because if the initial ranking is an imperfect measure of preferences, then a subject's choices teach us something new about their preferences; choices reflect how subjects feel about the goods they are choosing between. Put another way, while subjects may be randomly assigned to conditions, once they have been allowed to make a choice, the *good they choose* is not random, and comparisons which use that good must take this into account. No study in the choice-induced cognitive dissonance literature does this; indeed many studies in this literature (including the original Brehm 1956 paper) explicitly drop from consideration any subjects who choose B.

To see how this problem corrupts a "spreading" analysis, consider the following intuitive analysis of the FCP (in Appendix One I provide a numeric example that makes

this same analysis much more precisely). Call the good that an experimental subject initially ranked 7 good A, and the good they initially ranked 9 good B. For those subjects who choose A (many subjects choose B) the FCP looks for evidence that their rankings have spread; that the ranking of A has increased and the ranking of B has decreased. Put in terms of probabilities, what the FCP tests for is:

Equation One:
$$P(A \text{ rises and } B \text{ falls} \mid \text{choosing } A) > P(B \text{ rises and } A \text{ falls} \mid \text{choosing } A)$$

That is, the FCP claims that subjects who choose A should be more likely to raise the ranking of A and lower the ranking of B that they are to do the opposite, and vice versa for subjects who choose B. The FCP is only a valid test of cognitive dissonance if unbiased subjects should not be expected to show this pattern.

To think about this another way, we can flip this statement around from one about the probability of rankings spreading into one about the likelihood of choosing A. Recall that Bayes' Rule says that for any events X and Y:

$$P(X|Y) = P(Y|X) * P(X) / P(Y)$$

This lets us re-frame what the FCP tests into a *mathematically equivalent* statement about how likely subjects are to choose A or B. Note that if subject's first ratings are unbiased, then P(A rises and B falls) = P(B rises and A falls). Therefore, applying Bayes' Rule to both sides of equation one tells us that what the FCP tests is also equivalent to:

Equation Two:
$$P(\text{choosing } A \mid A \text{ rises and } B \text{ falls}) > P(\text{choosing } A \mid B \text{ rises and } A \text{ falls})$$

What equation two gives us another way of conceptualizing what the FCP tests when it claims to test for cognitive dissonance. As a thought experiment, imagine that we saw how two subjects ranked goods A and B in the first and third stage, but not how they had chosen in the second stage. Suppose subject one switched their rankings of goods A and B from 7 and 9 to 9 and 7, while subject two did just the opposite, widened their rankings to 5 and 11 (see Table Two). Put another way, imagine subject one *decreased* their "spread" by four places, and subject two *increased* their spread by four places.

Table Two: Two Hypothetical FCP Subjects

First Stage	Second Stage	Third Stage	
		Subject 1	Subject 2
A: 7	subjects chose	A: 9	A: 5
B: 9	A or B	B: 7	B: 11

Bayes' Rules implies that for the FCP to be a valid test of cognitive dissonance, it must be the case that if subjects one and two did not suffer from cognitive dissonance, then they are *equally likely* to have chosen A in the second stage. Put another way, the FCP essentially tests whether the third-stage rankings of subjects have anything to do with how they chose between those two goods, and assumes that only for subjects who suffer from cognitive dissonance can the two be related.

Reframed as a selection bias, while subjects are randomly assigned to conditions in a FCP experiment, which good they choose if given the chance is not random, and this induces a selection bias when comparing *chosen* goods across conditions. Subjects who choose A are those who were likely to spread their preferences in favor of A, and those who choose B are those who are likely to more their rankings in the opposite direction. Quantitatively, in Appendix One I go through in detail a numeric example which translates the simple thought experiment above into the exact "spreading" argument that FCP experiments make; I find that the selection bias I identify can easily explain at least twice the spreading than the typical FCP experiment finds.

My selection argument also easily explains why we see less spreading in "low-dissonance" condition subjects who choose between goods they initially ranked further apart (say 4 and 12), and even less spreading in control subjects who do not make choices. Note that the selection bias I identify is weaker in subjects that make easy choices (because fewer choose against their initial rankings), and absent in subjects who are not asked to choose. From the point of view of the bias I identify, if a subject's choices always conformed to their initial ratings, their choices would contain no new information, and reveal nothing new about their preferences. Therefore, selection should induce considerable "spreading" in experimental subjects, less in easy-choice subjects, and none in control subjects. In other words, what all FCP papers of this second form test when looking for CD is to be expected any time subjects' ratings are imperfect.

Here I have chosen to work through the example in which subjects rank items; indeed the exact same argument holds if subjects rate items. In no study in this literature do subjects rate and re-rate items consistently, even if they haven't been asked to make any choices. Therefore when subjects are observed to choose A over B, just as above it teaches us something about how their ratings of A over B are likely to change, absent any cognitive biases. Another way to see the problem when subjects rate items: simply take those ratings and turn them into a ranking. The above argument tells us that subjects should be expected to move their ranking of their chosen alternative up by several spots; naturally this would also lead us to expect that that item's rating must improve.

Scope and Generality of this Problem for Form Two:

The selection problem I identify would not be a problem for FCP papers if subjects not asked to make choices always ranked items in the exact same way; that is if rankings appeared to measure preferences with no error. But if these rankings do move, then subjects' choices should be expected to predict whether these movements are likely to be positive or negative. Indeed in none of the studies I cite do no-choice subjects' rankings stay constant, and they often move around by large amounts (often around 3 points on a 15 point scale, though the amount does not ultimately matter).

This problem would also not exist if no experimental subjects chose the good they initially ranked lower, but the number that do so is always considerable. For example, in the original Brehm study 21% of subjects choose objects that they had initially rated inferior, and in the Lieberman, Ochsner, Gilbert & Schacter study 36% of subjects choose objects they had ranked lower.

To help think about these numbers, the simple numeric computation I go through in Appendix One would suggest that if 25% of subjects who initially rank A over B then choose B, then we could easily expect an increase in the spread of preferences of at least one ranking point, completely absent CD. This difference is very close (if not larger) than what we see in most studies. Further complicating the interpretation of FCP studies, in many instances subjects who rank A over B but then choose B over A are thrown out. These studies are problematic for an even more serious and obvious reason.

Free-Choice Paradigms that Compare Cultures, Groups and Treatments:

Several papers which use the FCP procedure I describe above also find different size effects across cultures or additional treatments. For example, a study by Kitayama, Snibbe, Markus & Suzuki (2004) found Japanese subjects display less spreading than Canadians, and several studies have found that subjects display less spreading if they are first asked to self-affirm (Steele et al. 1983, 1988, 1993, Stone 1999, Hoshino-Browne et al. 2005). These comparisons are extremely suggestive, but may still be contaminated by the selection problems I describe above if the strength of selection varies systematically across groups.

For example, consider the claim that Japanese subjects show less cognitive dissonance than Canadians. One possible confound is that even if never asked to make choices, Japanese subjects may rate and re-rate goods more consistently that do Canadians. For instance, for cultural reasons Japanese subjects may try harder to correctly represent their preferences during the rating and re-rating portions of a FCP experiment. If this is true, then their subsequent choices contain less information and induce less selection; the FCP would mistakenly conclude they display less dissonance. This possible problem suggests a testable prediction; a Japanese control group (rank and re-rank with no intervening choice so no CD) should demonstrate more inter-rank correlation that a comparable Canadian group.

More broadly, several studies look at how either exogenous demographics (culture, status, gender), or randomized treatments (self-esteem boosts, choosing for yourself or a friend) interact with the measured spread of preferences coming out of a FCP. Given the problem I identify, there is no way while using a FCP measure, to separately measure how a treatment affects CD and how affects the accuracy of reported preferences, even with randomization of treatments and comparisons across different timings of the treatment. For example, consider the measured effect of self-affirmation on the spreading of preferences (studies by Steele and co-authors have found that self-affirmed individuals display less spreading).

These results are extremely suggestive, however self-affirmation has been shown to alter performance on a number of different tasks, complicating the issue of interpretation. If self-affirmed subjects either: more accurately report their preferences when ranking objects, or display more stable preferences over the span of a FCP experiment, this would generate a false finding of CD. Furthermore, if the effects of a self-affirmation treatment are short lived, this would also explain the observed timing of these effects, such as those observed in Steele (1993). Essentially, any comparison of FCP measured dissonance across groups or treatments is compromised by the fact that the FCP may not be an unbiased measure of dissonance in either group, and that these groups or treatments may directly interact with the selection bias that the FCP conflates with dissonance.

Discussion and Conclusion:

The bias I identify suggest that a reexamination of the literature on choice-induced cognitive dissonance may be in order. FCP papers of the first form I discuss, I believe, are highly problematic, as they magnify any imprecision in the measurement of preferences into a strong finding of cognitive dissonance. All other FCP papers are of the second form and are less obviously problematic, but still suffer from the same fundamental problem. Every FCP paper tests whether chosen goods improve in measured value, and vice-versa for unchosen goods. I show this is in fact to be expected even absent cognitive dissonance. Therefore, absent strong, unstated assumptions about why and how subjects' ratings change, these papers cannot be taken to accurately measure choice-induced cognitive dissonance. This is not to say that cognitive dissonance does not exist, nor are the problems I identify insurmountable.

For example, if a subject is allowed to choose between their 7th and 9th ranked goods, a method which avoids the problem I identify would compare the movements of 7th ranked goods to in experimental subjects to the movement of 7th ranked goods in control subjects, regardless of which good the subject chose. This "intent to treat" approach would rely on the fact that experimental subjects are likely to choose the 7th ranked good, without contaminating the comparison of experimental and control subjects by using information from their actual choices.

Another procedure that avoids the problem I discuss modifies the classic FCP by adding an additional step to the typical control treatment, as illustrated in Table Three.

Table Three: A Modified Free-Choice Method

	Stage 1:	Stage 2:	Stage 3:	Stage 4:
Experimental Group:	Rank	Choose	Rank	
Classic Control Group:	Rank		Rank	
Modified Control Group:	Rank		Rank	Choose

The classic FCP compares changes in rankings of goods chosen by experimental subjects to the movement of similarly ranked goods for subjects in the classic control group. The problem I identify, of course, is that the good a subject chooses is not random, and should be expected to predict changes from stage 1 to 3, even if it does not affect that ranking. What this suggests, is that even if a subject were to make their choice in a fourth stage (after both rankings), their chosen and unchosen items would still spread between stage 1 and 3, even though their choices could not have caused such a change. In other words, if stage 4 choices appear to "cause" the spreading of preferences between stage 1 and 3 in the modified control group, we know that the FCP has produced a spurious result. An additional benefit of this modified control group is that it allows for a basic decomposition exercise. If the spreading found in the experimental group is similar to that found in the modified control group, then cognitive dissonance is not at work. If the spreading in the experimental group is larger that that in the modified control group, we then have an idea as to how much of that spreading is caused by selection, and how much can be attributed to cognitive dissonance.

These and several other simple procedures would eliminate the problem I identify; however no paper in this literature has employed such a procedure. Given this, the methodological problems I identify may help explain why recent studies have found cognitive dissonance in unlikely subjects, and suggest that mere choice may fail to induce cognitive dissonance despite a fifty-year-old literature claiming otherwise. This may change how we think about the causes and origins of cognitive dissonance, as studies that were taken to discredit ego-defensive and self-signaling accounts of the phenomena may not actually contradict these accounts. Given these methodological concerns, a reevaluation of the literature on choice-induced cognitive dissonance may be called for.

References:

1. Aronson, J., G.L. Cohen & P.R. Nail, Self-Affirmation Theory: An Update and Appraisal, in *Cognitive dissonance: Progress on a pivotal theory in social psychology*, E. Harmon-Jones, J. Mills, Eds. (APA Books, Washington, D.C., 127-147. 1999)
2. Bem, D.J., *Psychological Review*. **74**, 183-200, (1967)
3. Brehm, J.W. Postdecision changes in the desirability of alternatives. Journal of *Abnormal Psychology*, **52** (3), 384 -389, (1956)
4. Egan, L.C., L.R. Santos & P. Bloom, The Origins of Cognitive Dissonance: Evidence From Children and Monkeys. *Psychological Science*. **18** (11), 978-983, (2007)
5. Festinger, L.A., *A Theory of Cognitive Dissonance* (Stanford Univ. Press, Stanford, CA, 1957).
6. Gerard, H.B. & G.L. White, Post-decisional reevaluation of choice alternatives. Pe*rsonality and Social Psychology Bulletin*. **9** (3), 365-369, (1983)
7. Harmon-Jones, E. & J. Mills, Eds., *Cognitive Dissonance: Progress on a Pivotal Theory in Social Psychology* (APA Books, Washington, D.C., 1999)
8. Hoshino-Browne, E., A.S. Zanna, S.J. Spencer, M.P. Zanna, S. Kitayama & S. Lackenbauer, On the cultural guises of cognitive dissonance: The case of Easterners and Westerners. *Journal of Personality and Social Psychology*, **89** (3), 294-310, (2005)
9. Jones, E. E., Major developments in social psychology during the past five decades. *The Handbook of Social Psychology*, ed. by Lindsay, G., Aronson, E. New York: Random House, 47-107, (1985)
10. Kitayama, S., A. Snibbe, H.R. Markus & T. Suzuki, Is There Any ''Free'' Choice? Self and Dissonance in Two Cultures, *Psychological Science.* **15** (8), 527-533, (2004)
11. Lieberman, M.D., K.N. Ochsner, D.T. Gilbert & D.L. Schacter, Do amnesiacs exhibit cognitive dissonance reduction? The role of explicit memory and attention in attitude change. *Psychological Science*. **12**, 135-140, (2001)
12. Lyubomirsky, S. & L. Ross, Changes in attractiveness of elected, rejected, and precluded alternatives: A comparison of happy and unhappy individuals. *Journal of Personality and Social Psychology*, **76**, 988-1007, (1999)
13. Nalebuff, B., Puzzles: Choose a curtain, duelity, two point conversions, and more. *Journal of Economic Perspectives*, **1** (2), 157-163, (1987)
14. Ross, L. & R.E. Nisbett, *The Person and the Situation: Perspectives of Social Psychology* (Temple Univ. Press, Philadelphia, 1991)
15. Shultz, T.R., E. Léveillé & M.R. Lepper, Free Choice and Cognitive Dissonance Revisited: Choosing Lesser Evils Versus Greater Goods. *Personality and Social Psychology Bulletin*. **25** (1), 40-48. (1999)
16. Steele, C. M. The psychology of self-affirmation: Sustaining the integrity of the self. In L. Berkowitz (Ed.), *Advances in experimental social psychology* **21**. 261-302. New York: Academic Press, (1988)
17. Steele, C.M. & T.J. Liu, Dissonance processes as self-affirmation. *Journal of Personality and Social Psychology*. **45**, 5-19, (1983)
18. Steele, C.M., S.J. Spencer & M. Lynch, Self-Image Resilience and Dissonance: The Role of Affirmational Resources, *Journal of Personality and Social Psychology,* **64**, 885-896, (1993)
19. Stone, J., What Exactly Have I Done? The Role of Self-Attribute Accessibility in Dissonance, in *Cognitive dissonance: Progress on a pivotal theory in social psychology*, E. Harmon-Jones, J. Mills, Eds. (APA Books, Washington, D.C., 127-147, 1999)

Appendix One: A Simple Numeric Example

I will illustrate how the free choice paradigm will detect a spreading of preferences between chosen and unchosen goods, even when subjects do not suffer from cognitive dissonance. First, I set up a basic model of a FCP experiment and describe the simplified behavior of subjects who do not suffer from cognitive dissonance. I then show using Bayes' Rule that what the FCP tests for is to be expected in this example. I conclude by discussing the generality of this example.

A FCP Experimental Design:

Assume that the experiment first asks subjects to rate 15 goods from best to worst, with 1 being best and 15 being worst. The experiment then asks subjects to choose between the goods they initially ranked 7 and 9. Finally, subjects are asked to re-rank all fifteen goods, and the spread between their chosen and unchosen goods is measured.

Experimental Hypothesis:

The FCP claims that subjects show Choice-Induced Cognitive Dissonance if the spread between their chosen and unchosen goods increases from the first stage to the third stage. We will show how this will happen even if subjects display no cognitive dissonance.

A Simple Model of Subject's Rankings:

Assume that subjects ratings are unbiased (do not regress to the mean) and are not affected by cognitive dissonance. Never the less, subject's rankings shift across time; in every FCP experiment that has been conducted this has been the case, even among "control" subjects who are not asked to make an intervening choice. To fix ideas, let us imagine a simple way in which subject's rankings could shift. Call the good a subject initial ranks 7 good A, and the good they ranked 9 good B. For simplicity, assume that a subject's third stage ranking of either A or B is just their initial ranking plus or minus 2, with equal probability. Since the shift up or down in rankings are of equal probability, subjects do not display an bias in their rankings; on average their rankings do not move. In the third stage then, we have four equally-likely ways that a subject can rank A and B, displayed in Appendix Table One.

Appendix Table One: Simplified Ranking Behavior of Unbiased Subjects

Goods:	First Stage rank:	Second Stage choose:	Third Stage re-rank (assume 1st rank ± 2):			
			Case 1:	Case 2:	Case 3:	Case 4:
A	7	choose	9	9	5	5
B	9	A or B	7	11	7	11

Note that the spread between A and B is 2 in cases two and three, is 6 in case four, and is -2 in case one. Since each case is equally likely, on average the spread between A and B stays constant at 2.

Now, as an intermediate step towards computing the expected spread of chosen alternatives, let us ask what the probability is (in each of our four cases) that the subject chose A or B in the second stage. That is, if the experimenter only observed the first and third stage behavior of unbiased subjects, what would the experimenter beliefs be about what a subject chose in stage two?

A Subject's Choices Conditional on their Rankings

Notation:

Let ChA and ChB denote choosing A and B in the second stage, respectively. This means that what we are solving for is the probability a subject chooses A (or B) in the second stage conditional on how the re-rank the good. So for example, P(ChA | Case1) asks: of the subjects who switch the ranking of A and B in stage 3, what fraction chose A in stage two? I will now discuss these conditional probability across our four different cases, and assume simple numbers for each value to aid in exposition.

Case One:

Let us first look at case one. In case one a subject initially gave goods A and B rankings of 7 and 9, then re-ranked them 9 and 7, switching their ranking.

First Stage	Second Stage	Third Stage
A: 7	choose	A: 9
B: 9	A or B	B: 7

Given this, what is the probability the subject chose A in the second stage? It seems natural to assume that since we have two rankings from the subject which are exactly the opposite of each other, that:

$$P(ChA \mid Case1) = 1/2.$$

We assume this for simplicity, though the exact number is not critical for anything but simplicity.

Case Two:

In case two, a subject initially gave goods A and B rankings of 7 and 9, then re-ranked them 9 and 11, worsening both item's ranks by 2.

First Stage	Second Stage	Third Stage
A: 7	choose	A: 9
B: 9	A or B	B: 11

Given this, what is the probability the chose A in the second stage? It seems natural to assume that since the subject has twice ranked good A over B by 2 places, that he should be more likely to choose A than subjects in case one, where subjects switched their rankings. Therefore if subject's rankings tell us something about how they value these goods, it seems natural to assume that:

$$P(ChA \mid Case2) > P(ChA \mid Case1)$$

For simplicity let us assume $P(ChA \mid Case2) = 3/4$, though all that is important is that it be strictly greater than $P(ChA \mid Case1)$.

Case Three:

In case three, a subject initially gave goods A and B rankings of 7 and 9, then re-ranked A and B 5 and 7, improving the rank of both items by 2.

First Stage	Second Stage	Third Stage
A: 7	choose	A: 5
B: 9	A or B	B: 7

Given this, what is the probability a subject chose A in the second stage? Just like in case two, the subject has twice ranked good A over B by 2 places. It seems natural to assume that:

$$P(ChA \mid Case3) = P(ChA \mid Case2)$$

This would imply that P(ChA | Case3) is also 3/4, though all that is important is that it be strictly greater than P(ChA | Case1).

Case Four:

Finally, in case four a subject initially gave goods A and B rankings of 7 and 9, then re-ranked A and B 9 and 11, widening the difference in rank between the two items by 6.

First Stage	Second Stage	Third Stage
A: 7	choose	A: 5
B: 9	A or B	B: 11

Given this, what is the probability a subject chose A in the second stage? It seems natural to assume that since the subject first ranked good A over B by 2 places then increased this spread considerably (to 6), that he should be even more likely to choose A than those who spread of rankings stayed the same. Therefore it seems natural to assume:

$$P(ChA \mid Case4) > P(ChA \mid Case3) = P(ChA \mid Case4) > P(ChA \mid Case2)$$

For simplicity let us assume P(ChA | Case4) = 1, though all that is important is that it be greater than in any of the first three cases.

Now that we have these 4 numbers, we can solve completely for the behavior of the subjects in our toy model. Note that subjects do not display cognitive dissonance, all we have assumed is that their rankings tell us something about how they are likely to act when asked to choose between goods. One number in specific we can now solve for, is the expected number of subjects who should choose A or B after initially ranking them 7 and 9.

Prior Probability of Choosing A

The percent of subjects who will choose A is just the fraction of subjects who will choose A in each of our four cases, weighted by one-fourth then summed. Using the simple number we have assumed for each of our four cases, this says that:

$$P(ChA) = (½ * ¼) + (¾ * ¼) + (¾ * ¼) + (1 * ¼) = 3/4.$$

That is, given the simple numbers we have assumed three quarters of subjects will choose A in the second round, and one quarter will choose B.

Now let us directly solve for what the free choice paradigm tests for, the expected spread between chosen and unchosen goods for our subjects.

The Expected Spread in Rankings between Chosen and Unchosen Goods

Subjects who Choose Good A

First, assume the subject chooses good A in round two. What then do we expect the spread of A and B to be in round three? This is just the spread between A and B in each of the four possible cases, multiplied by the probability of those cases after seeing the subject choose A in round two. This is:

- 2 * P(Case 1 | ChA) + 2 * P(Case 2 | ChA) + 2 * P(Case 3 | ChA) + 6 * P(Case 4 | ChA)

Now note that Bayes' Rule gives us a way of computing those conditional probabilities. Specifically, for any case X,

$$P(\text{Case X} \mid \text{ChA}) = P(\text{ChA} \mid \text{Case X}) * P(\text{Case X}) / P(\text{ChA})$$

Plugging in the simple numbers we assumed, this gives us that:

$$
\begin{aligned}
P(\text{Case 1} \mid \text{ChA}) \;&=\; (\tfrac{1}{2} * \tfrac{1}{4}) / \tfrac{3}{4} \;=\; 1/6, \\
P(\text{Case 2} \mid \text{ChA}) \;&=\; (\tfrac{3}{4} * \tfrac{1}{4}) / \tfrac{3}{4} \;=\; 1/4, \\
P(\text{Case 3} \mid \text{ChA}) \;&=\; (\tfrac{3}{4} * \tfrac{1}{4}) / \tfrac{3}{4} \;=\; 1/4, \\
P(\text{Case 4} \mid \text{ChA}) \;&=\; (1 * \tfrac{1}{4}) / \tfrac{3}{4} \;=\; 1/3.
\end{aligned}
$$

Therefore the expected spread of rank between A and B if the subject chooses A in round two is equal to:

$$-2 * 1/6 + 2 * 1/4 + 2 * 1/4 + 6 * 1/3 = 8/3,$$

or two and two-thirds. Recall that overall, the spread between goods A and B among subjects is only 2. Subjects who also choose A in the second stage, however, will on average increase the spread between A and B, even if they do not suffer from cognitive dissonance.

Subjects who Choose Good B

Now, assume the subject chooses good B in round two. What then, do we expect the spread of B and A to be in round three? This is just the spread between B and A to be in each of the four possible cases, multiplied by the probability of those cases after seeing the subject choose A in round two. This is:

  2 * P(Case 1 | ChB) - 2 * P(Case 2 | ChB) - 2 * P(Case 3 | ChB) - 6 * P(Case 4 | ChB)

Just like before, Bayes' Rule and the simple number we assumed lets us compute the relevant probabilities, giving us that:

$$P(\text{Case X} \mid \text{ChB}) = P(\text{ChB} \mid \text{Case X}) * P(\text{Case X}) / P(\text{ChB})$$

and hence:

$$
\begin{aligned}
P(\text{Case 1} \mid \text{ChB}) &= (\tfrac{1}{2} * \tfrac{1}{4}) / \tfrac{1}{4} = 1/2, \\
P(\text{Case 2} \mid \text{ChB}) &= (\tfrac{1}{4} * \tfrac{1}{4}) / \tfrac{1}{4} = 1/4, \\
P(\text{Case 3} \mid \text{ChB}) &= (\tfrac{1}{4} * \tfrac{1}{4}) / \tfrac{1}{4} = 1/4, \\
P(\text{Case 4} \mid \text{ChB}) &= (0 * \tfrac{1}{4}) / \tfrac{1}{4} = 0.
\end{aligned}
$$

Therefore the expected spread of rank between B and A if the subject chooses B in round two is equal to:

$$2 * 1/2 - 2 * 1/4 - 2 * 1/4 - 6 * 0 = 0.$$

Recall that overall, the spread between goods B and A among subjects is -2; however, those subjects who choose contrary to their initial ranking and choose B in the second stage will increase the spread between B and A by two ranking points to 0, even if they do not suffer from cognitive dissonance.

Averaging both of these effects together, subjects are expected to increase the spread of the chosen good to their unchosen good by

$$3/4 * 2/3 + 1/4 * 2 = 1.$$

That is, overall, subjects will "spread" the ranking between their chosen and unchosen goods by an average of 1 ranking point (or 50% of the initial spread), even though they do not suffer from cognitive dissonance, and even though their ratings were unbiased. That is, what the FCP tests for is to be expected even in the absence of cognitive dissonance.

Discussion:

Note that the particular numbers we have assumed are not important, they simply make the computations easy to follow. All we rely on here is Bayes' Rule, and a basic intuition that when subjects rank items, that may tell you something about how they feel about those items, and hence tell you how they might choose between these items when asked. The first of our assumptions is simply a mathematical property of probabilities, the second is a necessary condition for the rankings provided by subjects to be meaningful.

Put another way, what this example illustrates is another way of conceptualizing what the FCP tests when it claims to test for cognitive dissonance. Imagine that we saw how two subjects ranked goods in the first and third stage, but were not told how they had chosen

in the second stage. Subject one switched their rankings of goods A and B, while subject two significantly widened them to 5 and 11.

First Stage	Second Stage	Third Stage	
		Subject 1	**Subject 2**
A: 7	subjects chose	A: 9	A: 5
B: 9	A or B	B: 7	B: 11

Key Conclusion:

Bayes' Rules implies that for the FCP to be a valid test of cognitive dissonance, it must be the case that if subjects one and two do not suffer from cognitive dissonance, then they would be equally likely to have chosen A in the second stage. That is, unless we are willing to rule out that an unbiased subject two may be more likely that an unbiased subject one to have chosen A, the FCP will conflate rational behavior with cognitive dissonance.