Supplementary Material for “The neural basis of rationalization: Cognitive dissonance reduction during decision-making”

Johanna M. Jarcho\textsuperscript{1,2}

Elliot T. Berkman\textsuperscript{3}

Matthew D. Lieberman\textsuperscript{3}

\textsuperscript{1}Department of Psychiatry & Biobehavioral Science, University of California, Los Angeles, CA 90025, USA.

\textsuperscript{2}Center for Neurobiology of Stress, University of California, Los Angeles, CA 90025, USA

\textsuperscript{3}Department of Psychology, University of California, Los Angeles, CA 90025, USA
Pilot Study

Because decision-related attitude change had not been examined with fMRI, we conducted a pilot study to determine the feasibility of such a study. Constraints of event-related fMRI studies differ sufficiently from those involved in classic behavioral studies of decision-related attitude change (Brehm, 1956) that the standard paradigm needed to be modified. In behavioral studies, subjects have traditionally made a small number of decisions between similarly rated items, with statistical power generated by including dozens of subjects. In contrast, fMRI samples are typically small by comparison, with each subject being exposed multiple times to each type of trial. Thus, we designed a decision-related attitude change paradigm in which subjects made many more decisions than in a typical behavioral study.

Our goals in this pilot study were twofold. First, we wanted to determine whether a modified paradigm with multiple decision-making trials would produce group-level effects similar to studies using a classic paradigm. It is possible that repeated decisions by individual subjects might contaminate or interfere with the typically observed effect of decision-related attitude change. Second, we wanted to determine the proportion of subjects that demonstrated significant within-subject effects, and thus exhibited reliable attitude change across many trials. This is critical because only those subjects who produce reliable effects will be used in the event-related fMRI analyses conducted in the primary study. Just as neuroimaging studies of the placebo response typically exclude non-responders from analyses (Mayberg et al., 2002; Sarinopoulos et al., 2006), we also planned to exclude those who did not demonstrate significant levels of post-decision attitude change (‘non-responders’) from our neuroimaging analyses. Determining the
METHODS

Subjects

Thirteen right-handed, English-speaking subjects (4 male; mean age 20 ± 1.58 years) were recruited from the UCLA campus. All subjects gave informed consent to participate according to a protocol approved by the University of California, Los Angeles Institutional Review Board, and received 1 hour of class credit for their time, plus an additional $10 of compensation. Three additional subjects were eliminated from analyses because of technical difficulties.

Procedures

Except when noted, behavioral methods are the same as those used in the study reported in the main text (depicted in Supplementary Fig. S1). All portions of the study were completed in a behavioral testing room. Once subjects provided their initial ratings of how much they liked 140 names and 140 paintings, they exited the testing room and completed filler questionnaires while the experimenter prepared stimuli for the decision-making phase of the study. This delay between initial ratings and decision-making was designed to be similar in duration to the delay subjects would experience in an fMRI study (i.e., the positioning of the subject and structural scanning that precede functional scans). Subjects were then invited back into the testing room, where they were told they would be presented with pairs of names and pairs of paintings, and they were to choose
the item in each pair they preferred. Just as in the primary study, when subjects were presented with pairs of names, they were instructed to imagine they were choosing a name for their own future child, and to select which name in each pair they would rather name that child. When choosing between paintings, subjects were led to believe that they would receive posters of two paintings they indicated they preferred.

The pairs of similarly rated items were presented over two counterbalanced, 5-minute runs, with one run of names and the other of paintings. Each run consisted of 40 5-second trials during which subjects indicated by button press which item in each pair they preferred. Once subjects made this choice a blank screen replaced the stimuli for the duration of the 5-second trial.

After decision-making, subjects once again rated how much they liked all 280 items on a 1-100 scale. At the end of the study, subjects learned they would not receive posters and were instead compensated an additional $10.

**Quantification of attitude change scores**

An ‘attitude change score’ was calculated for each item by subtracting the initial rating from the final rating of an item, such that a positive score indicated attitudes toward an item had become more positive during the course of the study. Stimuli that comprised each of the 80 pairs presented during decision-making were classified, post hoc, as a selected or rejected item. The 120 items that were rated twice, but not included in the decision-making phase were classified as ‘no choice’ items. Results are depicted in Supplementary Fig. S2.
Quantification of post-decision attitude change scores at the subject level

Previous behavioral studies find that subjects demonstrate decision-related attitude change effect on average. It remains unknown what proportion of subjects consistently show the effect from trial to trial because there have not been enough trials in past studies to measure this within-subjects effect. To assess this, we computed a paired samples t-test for each subject to determine if attitude changes scores were greater for selected compared with rejected items across all trials. Significant differences indicate the subject demonstrated a general tendency to produce decision-related attitude change. To control for the possibility that attitude change was simply due to repeated ratings, attitude change scores were also computed for 120 items that were rated twice, without an intervening choice.

RESULTS

Magnitude of attitude change did not differ by the order in which decision-making runs occurred, nor was it affected by whether stimuli were presented during the first or second half of initial and final rating phases. Initial ratings of names and paintings were not significantly different, and attitude change scores did not vary based on whether stimulus pairs were comprised of paintings or names. Given this, data were pooled together across runs and across paintings and names for the remainder of analyses.

The first goal of the study was to determine whether the adapted paradigm produced decision-related attitude change at the group level. Attitude change scores were significantly more positive for selected items (M=2.27, SD=5.02) than rejected items (M=-3.61, SD=6.71; t(12)=3.08, p<.01). Since items were paired based on the similarity
of their initial ratings, this difference indicates that following decision-making, selected items came to be viewed more positively, relative to rejected items, (Supplementary Fig. S2). Attitudes about no choice items did not change (M=-1.39, SD=3.97; t(12)=1.26, p=ns).

The second goal of the pilot study was to determine what percent of subjects produced reliable decision-related attitude change across trials. Paired samples t-tests compared the attitude change score for selected and rejected items on a subject-by-subject basis, and revealed that the majority of subjects (69.3%) demonstrated a significant level of attitude change across trials. The mean difference in attitude change scores between selected and rejected items for subjects with significant attitude change ("responders") was 7.90 (SD=7.40), and the mean for other subjects ("non-responders") was 1.34 (SD=1.90). Based on these data, we expected 60%-70% of subjects sampled from a similar undergraduate population to be responders in the fMRI study.

DISCUSSION

We sought to develop an experimental paradigm that was able to produce decision-related attitude change while conforming to methodological constraints posed by fMRI. A classic experimental paradigm known to produce decision-supporting attitude change (Brehm, 1956) was modified such that the number of items subjects rated, as well as the number of subsequent decisions made about those items, was increased dramatically, while the amount of time provided for decision-making was limited to 5 seconds. Additionally, the stimuli that subjects rated and made choices about were drawn from multiple domains. Despite these modifications, significant levels of attitude change
occurred, regardless of the order in which choices and ratings were made, and the type of stimuli that were presented. The ability to replicate previous behavioral results, even with a relatively small sample size, suggests the modified paradigm utilized in the pilot study is a valid paradigm that can be implemented in the context of fMRI studies decision-related attitude change. Additionally, we observed that 69% of subjects produced significant within subject attitude changes effects. This suggests that in our main fMRI study, 60%-70% of subjects should have enough trials with attitude change to conduct within subject event-related analyses.

**Supplementary data from fMRI study**

Both classic and current models of cognitive dissonance suggest conflict sets cognitive dissonance reduction, and thus attitude change, in motion (Festinger, 1957). As such, the absence of a relationship between activity in dorsal anterior cingulate, a brain region most commonly associated with conflict (Botvinick et al., 2004), is somewhat puzzling. Also somewhat puzzling is the negative relationship between attitude change and activity in the anterior insula, a brain region associated with negative affect and distress (Sanfey et al., 2003; Tabibnia et al., 2008), which is often considered the other factor necessary to drives cognitive dissonance reduction (Festinger, 1957). Exploratory analyses altering the time window investigated for each trial may help account for these apparent discrepancies.
Exploratory analysis of the relationship between attitude change and functional brain activity preceding decision-making

Exploratory event-related regression analyses focused on brain activity that occurred during the initial period of each trial. Specifically during the time that elapsed between initial presentation of each pair of stimuli, and the subject’s button press indicating which item in the pair they preferred. Duration of this initial period was automatically recorded via computer, and was used to define the time frame of analysis for each trial.

Each trial was modeled as an event with the attitude change score entered as a parametric modulator. Each trial, which varied in duration based on response time, was convolved with the hemodynamic response function; the attitude change score for that trial was then used as a scaling factor for the hemodynamic response function. This analysis produced an activation map of regions that were significantly associated with attitude change for each subject during the time between initial presentation of stimuli, and the button press that signified a decision had been made. These maps were then entered into a random-effects analysis at the group level.

RESULTS AND DISCUSSION

Although absent in analyses that included the entire 5-second trial, when analyses were limited to the time prior to subjects’ decision on each trial (M=2.19 sec, SD=.38), left dorsal anterior cingulate was positively associated with attitude change (t(11)=4.35). Whereas activity in anterior insula was negatively correlated with attitude change across the entire 5-second trial, when analyses were limited to the time
that elapsed prior to subjects’ decision on each trial, left anterior insula activity was also positively correlated with attitude change (-30, 32, -2; k=20; t(11)=3.52, p<.005).

Thus, activity in dorsal anterior cingulate, which is often associated with conflict, is positively associated with attitude change, but only early during the decision-making process. Activity in the anterior insula, which occurred very early in the decision-making process, was also positively associated with attitude change, but the extent to which it was ultimately dampened down over the entire trial, was negatively associated with attitude change. Although exploratory in nature, these data provide support for the idea that greater conflict or discomfort may be associated with greater attitude change. Indeed, in a recent study, a similar relationship between activity in anterior insula, dorsal anterior cingulate, and attitude change was demonstrated in the context of performing counter-attitudinal behavior, a situation also thought to elicit cognitive dissonance (van Veen et al., 2009), further supporting the theory that conflict or distress set attitude change in motion. However, unlike the current study, van Veen and colleagues did not identify potential neural mechanisms activated during counter-attitudinal behavior associated with the resolution of distress that presumably facilitate attitude change. This suggests that the conflict aroused during counter-attitudinal behavior may ultimately be resolved at a later point in time, perhaps during re-evaluation of attitudes. In contrast, decision-making may involve a more rapid resolution of discomfort or distress generated early in the decision-making process, as demonstrated by the inhibitory relationship between right IFG and regions such as anterior insula, associated with attitude change.
SUPPLEMENTARY REFERENCES


**Supplementary Figure Captions**

Supplementary Figure S1. Diagram of experimental design. (a) Initial Rating: Outside the fMRI scanner, subjects rated how much they liked paintings and names on a 1-100 point scale. (b) Decision-Making: Once inside the scanner, subjects indicated by button press which one of two similarly rated items they would rather hang in their home (paintings) or name their child (names). (c) Final Rating: After exiting the scanner, subjects re-rated all items.

Supplementary Figure S2. Pilot Study: Average change in attitudes from pre-choice to post-choice ratings. Items were categorized post hoc based on whether the item was ‘selected’ or ‘rejected’ during decision-making, or was rated twice but excluded from the decision-making phase of the study (‘No Choice’).
Supplementary Figure S1

a

**Initial Rating**
How much do you like each item?

b

**Decision Making**

Which would you hang in your home?

5 sec

~2.5 sec

~2.5 sec

~2.5 sec

~2.5 sec

x 40

Final Rating

How much do you like each item?

x 140

x 140
Supplementary Figure S2

![Supplementary Figure S2](image-url)