Testing for Cognitive Dissonance
Evidence from Children

By
Steinar Holden
Department of Economics
University of Oslo
Steinar.holden@econ.uio.no
July 2008

Abstract
In a recent paper, Egan, Santos and Bloom (2007) report from an experiment which they claim provides evidence for cognitive dissonance among preschoolers. The subjects are given a choice between two alternatives (two different stickers) that the subjects previously have given the same rating of liking. The subjects are then given the choice between the unchosen sticker and a third sticker that was originally also given the same rating. The children preferred the third sticker over the unchosen one, which Egan et al interpreted as a reduction in the liking of the unchosen sticker, i.e. as evidence for cognitive dissonance. In this article I argue that the two alternatives should not be viewed as equally attractive, in spite of the previous equal rating. This implies that the preference for the third sticker is not caused by cognitive dissonance. Furthermore, I report the results from a variation to their experiment that supports my argument.

Introduction

The theory of cognitive dissonance has been subject to a lot of research since it was proposed by Leon Festinger in 1957 (Festinger, 1957). Broadly, the theory says that people feel uncomfortable or subject to stress if their cognitions, like knowledge, attitudes, emotions, or behavior, are in conflict with each other. One strand of this research is the postdecisional dissonance studies pioneered by Jack Brehm (1956). Brehm showed that female students, after having made a difficult decision, displayed decision rationalization by favoring the alternative they had chosen more strongly.

In a recent paper, Egan, Santos and Bloom (2007) test for decision rationalization or cognitive dissonance among preschoolers, by giving them a choice between two alternatives (two different stickers) that the subjects previously have given the same rating of liking. The subjects are then given the choice between the unchosen sticker and a third sticker that was originally also given the same rating of liking. The children preferred the third sticker over the unchosen one, which Egan et al interpreted as a change in the attitude toward the unchosen sticker, deeming it less valuable, i.e. as evidence for cognitive dissonance.

However, the interpretation of their experiment is more complicated than they acknowledge. Even if the children have given the two stickers the same rating of liking, it may still be the case that the children prefer one over the other. Even within the same level of rating, stickers are not necessarily equally attractive. This implies that the preference for the third sticker need not be caused by cognitive dissonance. In this article I discuss how their evidence should be interpreted. Furthermore, I report the results from a variation to their experiment that sheds additional light on the interpretation. In independent work, Chen (2008) has put forward essentially the same criticism of common tests of cognitive dissonance. (Egan et al also did experiments with monkeys, but the same argument applies to this experiment.)

The experiment of Egan et al

In the experiment of Egan et al, 30 4-year-old children were shown a number of commercially available foam stickers of various shapes (e.g. dolphin or dragonfly). The children were asked to rate the stickers using a smiley-face rating scale with six faces, i.e. six rating levels. The aim of this part of the experiment was to identify a series of triads, defined as three stickers (A, B and C) that a child had given the same rating of liking. In the next phase, each child was given the choice between two stickers A and B in a triad, i.e. which the child had given the same rating. Then the child was given a second choice between the sticker not selected (either A or B) and the third sticker in the triad, C.

As the three stickers in a triad were given the same rating, one would expect that in a choice between two of the stickers, there would be a 50% probability for choosing each of them. The hypothesis of interest was whether the 50% probability would be changed by the prior decision. Would the children experience dissonance in choosing one equally preferred sticker over the other, implying that they afterwards would like the unchosen sticker less because of the prior decision? In other words, if a child chose A over B in the first choice, would this decision induce the child to prefer C over B in the next choice? It turned out that in 63% of the cases the subject chose the novel alternative, C, rather than the one not selected in first choice (A or B). Egan et al interpreted this as indicating that the children demonstrated a decrease in their preference for the sticker not chosen in the first choice, i.e. as evidence of decision rationalization and cognitive dissonance.

However, while the children had given the three stickers in a triad the same rating level, this does not necessarily imply that the children viewed the stickers as equally attractive. First, children may have a finer preference scale for stickers than six levels, implying that stickers given the same rating are not equally attractive. Second, rating a number of stickers according to a rating scale is not the same as choosing between two stickers. Possibly, children put more effort and concentration into a real choice of which sticker to take home, rather than the more "academic" process of rating the stickers.

If we assume that the stickers are not equally attractive, but rather that the children, when given the choice, are able to rank the three stickers, there are six possible rankings. These are ABC, ACB, BAC, BCA, CAB, and CBA. Ex ante, all these rankings are equally likely. However, with the additional knowledge of the first choice, in which we for concreteness assume that the child chose sticker A, some of these rankings are no longer possible. Specifically, we can delete the rankings where B is better than A, as these rankings are inconsistent with the first choice. One is then left with three possible rankings. These are 1) ABC 2) ACB 3) CAB. We observe that B is ranked above C in only one of them, while C is ranked above B in the latter two. With constant preferences, i.e. without any form of decision rationalization, these three possibilities are still equally likely. Thus, it follows that without any decision rationalization, one would expect the subjects to prefer the third sticker C in 66.7% of the cases. In fact, the experimental outcome from Egan et al's study of 63.0% is lower than the expected outcome of 66.7% without decision rationalization. Thus, if anything it would suggest a change in preferences going in the "wrong direction", but presumably not significant.

The intuition here is that while A, B and C are equally attractive in expected terms ex ante, the fact that a subject prefers A to B provides new information that A is likely to be somewhat

more attractive, and B likely to be somewhat less attractive. Thus, subjects are likely to prefer A over C (with probability 2/3), and C over B (with probability 2/3), cf. the possible rankings above.

Egan et al also undertook a no-choice condition. Here, the subjects were presented with two stickers A and B, and then received one of them randomly, i.e. without a choice. Then they were allowed to choose between the unreceived alternative and a third sticker C. This time the children chose C in 47.2% of the cases. In this condition, the children were not expected to experience dissonance, as they themselves never made a choice between the two stickers. Thus 50% was the expected outcome in Egan et al's study. However, also according to the "ranking argument" made above, 50% would be the expected outcome: As the subjects have not ranked A and B, all the six rankings listed above are possible. Then C is better than the unreceived alternative, A or B, in 50% of them.

To explore the validity of this critique, I reran Egan et al's experiment with one variation, which consisted of adding a new stage. After rating the stickers, when being presented with two stickers A and B from a triad, the children were asked whether they liked the stickers equally much, or whether they liked one sticker better than the other. Then the children were asked which sticker they would like to take home, etc. The idea here was that if children responded that they liked one sticker better than the other, the ranking argument above would clearly apply, and 66.7% would be the expected outcome without any cognitive dissonance.

The novel experiment

Thirty 4- and 5-year-olds participated in the study, 13 girls and 17 boys. Children were recruited from two preschools in Eiksmarka, close to Oslo, Norway. They were tested in their preschool, while sitting at a desk across from the experimenter. The experiments were undertaken by two research assistants, in addition to the author of this article.

Following Egan et al, children's preferences for different stickers were assessed using a smiley-face rating scale that included six faces, from sad to very happy, corresponding to six levels of liking. While many of the children already were familiar with smiley-faces as a measure of liking, the experimenter nevertheless ensured that all understood the scale. This was confirmed by appropriate responses to three queries by the experimenter: "Let's say I like a sticker a whole lot/not at all/somewhere in the middle. Which face should I put it with?".

When the children had shown that they understood the rating scale, they were presented with stickers one by one and asked to match to the faces. We used commercially available adhesive foam stickers with various pictures, like faces, animals, cars, stars, etc. Most children continued to rate stickers until they become fatigued, while some of the eager ones were stopped after rating almost 40 stickers. All children rated at least three triads, i.e. at least three times three stickers with the same rating level.

When the children had rated the stickers, they participated in one of two conditions, either the *choice* condition or the *no-choice* condition. In the choice condition, the child was given the choice between two stickers, A and B, randomly chosen from a triad. The stickers were put on a plate in front of the child, and the experimenter asked "Do you like these stickers equally much, or do you like one sticker better than the other? And if so, which one do you like better?" When the child had responded to this/these questions, he or she was asked which of the stickers he or she would like to take home. (Chosen stickers were put in an envelope

bearing the child's name, to be taken home at the end of the day.) Next, the child was given the choice between the unchosen alternative (either A or B) and the third sticker in the triad, C. This process continued until the child had chosen between all the triads.

In the no-choice condition, the experimenter displayed two stickers A and B from a triad to the child, and said, "Look, here are two stickers. You'll get this sticker; we'll put it in your envelope." (One sticker was randomly chosen and put in the envelope.) Then the child was presented with the choice between the remaining sticker, A or B, and the third sticker in the triad, C.

For all children in both conditions, after the rated triads were exhausted, we also undertook an extra experiment. The extra experiment was identical to the choice condition explained above, except that the stickers were not previously rated, and thus did not have the same rating of liking. Thus, each child was given the choice between two stickers, A and B. When the child had chosen one of the stickers, he or she was given a new choice between the unchosen sticker (A or B) and a third unrated sticker C. This procedure was repeated 5-6 times. The motivation for this experiment was to explore children's choice between stickers that were not previously given the same rating, and thus would be much more different in regard to attractiveness. In this setting one would expect that most choices were fairly easy, implying that no decision rationalization would take place. If children had stable and consistent preferences over the stickers, one would then expect the "ranking" argument presented above to apply. Thus, children were expected to choose the novel alternative, C, in two thirds of the cases.

Results

In the choice-condition, the 16 children on average rated 8.8 triads, while in the no-choice condition, the average over 14 children was 5.4 triads. The somewhat lower number in the no-choice reflected the available time before the children were leaving for home. Following Egan et al, for each child we computed a percentage preference for the novel option, C, over the unchosen (choice-condition) or unreceived (no-choice condition) option, A or B. Thus, a child choosing C for three out of five triads would have a percentage preference of 60%. Like Egan et al, we then computed the mean percentage preference for C for each of the conditions.

For the 16 children in the choice condition, the mean percentage preference for the novel option C was 51.9%. This was considerably below the expected value of two-thirds. In fact, this is significantly lower than 66.7% according to a one-sample t-test, $t(15) = 2.53$, p-value = 0.02, two-tailed. In the no-choice condition, the mean percentage preference for C was 57.4%. This is not significantly different from the hypothesized mean of 50%, $t(13) = 1.09$. While the difference in the percentage preference for C between the two conditions goes in the opposite direction of what was expected, the difference was not significant in an unpaired two-sample t-test, $t(28) = 0.62$, p-value = 0.54.

We then consider the children's response in the choice condition on whether they liked the two stickers in a triad equally much. In only 20.6% of the cases, 29 of 141, did the children respond that they liked the two stickers equally much. Nine children did this only once, four children did this 2-3 times, and two children did it five times. For these 15 children, we calculated the mean percentage preference for the novel option. Thus, for each of these children we calculated the percentage preference for the novel option, C, and we calculated the mean over all the 15 children, which was 62.0%. On the assumption that the stickers A and B were equally attractive, in spite of the children subsequently choosing one over the

other, the ranking argument above would not apply. Thus, the expected preference for C would be 50%. However, the mean of 62.0% is not significantly above 50%; a t-test gives $t(14) = 1.06$, p-value = 0.31. Given that nine of the children responded only once that the two stickers were equally attractive, implying that their percentage preference would be zero or unity, the statistical uncertainty of the mean percentage preference is very large.

In the large majority of cases (79.4%) where the children responded that they liked one of the stickers A and B better than the other, we also calculated the percentage preference of the novel option C for each child. Furthermore, we calculated the mean preference for all the 15 children who responded this. This was 53.9%, considerably lower than the hypothesized value of 66.7%. It is also significantly lower in a t-test; $t(14) = 2.57$, p-value = 0.02. However, an unpaired two-sample t-test revealed no significant difference between the preference for C depending on whether or not the child had responded that s/he liked one of the stickers better than the other: $t(28) = 0.66$, p-value = 0.52.

Finally, we report the results from that extra experiment, where all the 30 children chose between stickers that they had not previously rated. Again, the children first chose between two stickers A and B. Then they chose between the sticker not chosen in the first choice, and a third sticker C. On average, the children undertook 5.7 "rounds". For each child we calculated the preference for C. We also calculated the mean preference over all the 30 children, which was 68.2%, i.e. very close to the hypothesized value of 66.7%. This is significantly above 50% in a t-test; $t(29) = 5.15$, p-value = 0.00.

Discussion

In almost four out of five cases, the children responded that the two stickers that they had previously given the same rating, were nevertheless not equally attractive. This result undermines the interpretation given in Egan et al, that the preference of the novel option reflects decision rationalization. When the two stickers in the first choice are not equally attractive, the expected preference for the novel option is no longer 50% even with constant preferences. As argued above, the expected preference for the novel option, without any decision rationalization, is 66.7%. This is consistent with the results of the extra experiment reported above, where the children chose between unrated stickers, and where the mean preference for the novel option was 68.2%. The upshot is that the preference for the novel option would have to be above 66.7%, and not only above 50%, to indicate decision rationalization.

In the choice condition, and also in the sub-sample where the children answered that they liked one of the stickers A and B better than the other, the preference for the novel option of 51.9% and 53.9% was significantly lower than the hypothesized value of 66.7%. This difference goes in the opposite direction of what would have been implied by decision rationalization. One speculative hypothesis is that the children after the first choice, "wanted the other one too". However, one should recall that the children had already given the stickers the same ranking, implying that the difference in liking between the stickers would in any case be small. This gives more room for arbitrariness in the decision of the children, which would bias the outcome towards 50%. This interpretation is consistent with the results in the extra experiment. In this experiment the stickers were more different, implying that there would be less scope for arbitrariness, and hence less or no bias towards 50%. Indeed, here the preference for the novel option was 68.2%.

In the few cases (20.6%) where the children responded that they liked the stickers equally much, the mean preference for the novel option was 62.0%. On the assumption that the stickers were indeed equally attractive, the hypothesized value would be 50%. Thus, the results go in the direction of decision rationalization. However, one should recall that the number of observations is very low, implying that the statistical uncertainty is vast (as also indicated by the p-value of 0.31). Note also that even if the children responded that the stickers were equally attractive, they may still have a slight preference for one over the other. Then the "ranking" argument above applies, and the expected preference of the novel option would be 66.7%. Under this interpretation the evidence goes in the "wrong direction".

Our experiment does not answer the question why the children in most cases displayed a strict preference ranking over stickers that they had previously given the same rating of liking. From the experiments, it was not our impression that there were too few levels in the rating. Thus, a finer rating scale would not necessarily have helped. Our conjecture is rather that the children find it easier to choose between two stickers than to rate stickers according to a rating scale. Furthermore, we would suggest that the choices that the children make between the stickers give a better indication of their preferences than does the ranking.

Overall, this article makes three points: First, subjects in most cases prefer one item over another, even if they have previously given them the same rating of liking. Second, the expected preference for the novel option in the experimental design of Egan et al, with no decision rationalization, is 66.7%, and not 50%. Third, the experiment shows no evidence of decision rationalization. This is consistent with the results reported by Egan et al, even if my interpretation is different.

Acknowledgment

The author would thank Åshild Auglænd Johnsen and Marius Østli for their help running the experiment, Åsa Rosèn and Kalle Moene for helpful discussions on the experimental design, and the children and personnel at the preschools at Fossum and Eiksmarka for their kind participation.

References

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

Chen, K. (2008). Rationalization and Cognitive Dissonance: Do Choices Affect or Reflect Prefences? Mime, Yale University.

Egan, L.C, L.R. Santos and P. Bloom (2007). The Origins of Cognitive Dissonance. Evidence From Children and Monkeys. *Psychological Science,* 18, 978-983.