

The Effects of Expectations on Perception: Experimental Design Issues and Further Evidence

Tyler Williams

Abstract:

Numerous studies have found that top-down processes can affect perceptions. This study examines some of the issues involved in designing field experiments aimed at discovering whether top-down mental processes affect perceptions, and, if so, *how* the influence takes place. Lee, Frederick, and Ariely (2006) (LFA) attempt to go further by testing whether expectations affect perceptions directly, by altering how sensory receptors and/or the brain's processing centers interpret an outside stimulus—or indirectly, for example, by changing the amount of attention paid to the outside stimulus. In order to test the robustness of the findings in LFA, this paper reports the results of a field experiment similar to the one analyzed in LFA. The field experiment, designed to address some potential confounding factors in this type of research, confirms that expectations can alter perceptions. However, it also shows that heterogeneity across individuals can play a role in determining the nature of this effect, a finding that complicates the interpretation of results such as those in LFA. To frame the analysis, this paper discusses the difficulties in designing this type of experiment, makes some improvements to existing designs, and suggests some ways of eliminating the confounding influences that remain.

JEL Classifications: D43, L13, L93

Tyler Williams is a senior research assistant at the Federal Reserve Bank of Boston's Research Center for Behavioral Economics and Decision-Making. His email address is tyler.williams@bos.frb.org.

This paper, which may be revised, is available on the web site of the Federal Reserve Bank of Boston at <http://www.bos.frb.org/economic/wp/index.htm>.

The views expressed in this paper are those of the author and not necessarily those of the Federal Reserve Bank of Boston or the Federal Reserve System.

I am grateful to Florian Ederer, Drazen Prelec, Lorenz Goette, and Suzanne Lorant for helpful comments throughout the course of this study.

This version: September 13, 2007

1 Introduction

A substantial number of studies on a variety of sensory stimuli have shown that people’s sensory experiences are determined not only by bottom-up processes (that is, through the impact of external stimuli on individuals’ sensory organs), but also by top-down processes, such as expectations and prior desires. The evidence for such joint effects is strikingly broad. Studies have shown that top-down processes affect responses to visual stimuli (Biederman, 1972; Palmer, 1975), assessments of a person’s ability (Darley and Gross, 1983; Jones et al., 1968), judgments of extended events (Brief et al., 1993; David et al., 1997), the enjoyment of a film (Klaaren, Hodges, and Wilson, 1994; Geers and Lassiter, 2005), and cigarette preferences (Friedman and Dipple, 1978). Friedman and Dipple (1978), for example, manipulate the brand name of identical cigarettes across subjects, and find that males (females) enjoy smoking the cigarettes more when the cigarettes have a masculine (feminine) name. Since changing the name of the cigarettes does not change the bottom-up inputs from smoking the cigarette, this result implies that top-down processes (such as expectations or concerns for one’s image) must affect perception in some way.

In the realm of food and drink, expectations-triggering information, generally in the form of brand names or labeling, affects preferences for, among other items, cola (McClure et al., 2004), turkey (Makens, 1965), seltzer (Nevid, 1981), beer (Allison and Uhl, 1964), nutrition bars (Wansink et al., 2000), coffee (Olson and Dover, 1978), yogurt and cheese spreads (Wardle and Solomons, 1994), and vanilla ice cream (Bowen et al., 1992).¹ Most of these studies measure the importance of top-down processes by comparing a control treatment, in which individuals are given no information about the stimulus prior to experiencing it, with an experimental treatment, in which individuals receive information or some other kind of cue for expectations prior to sampling the stimulus. In the classic study by Allison and Uhl (1964), for example, subjects in the control treatment sample various beers without brand labels, while subjects in the experimental treatment drink the beer from labeled bottles. Allison and Uhl find significant differences in preferences for the same beer between the two treatments. Again, the stimulus itself, and hence its bottom-up influence on perception, is held constant between the treatments; thus, the difference in preferences between the two treatments implies that top-down processes must

¹For even more studies, including many that focus on the efficacy of various advertising techniques, see the review by Deliza and MacFie (1996).

somehow change evaluation of the stimulus.

However, despite the seemingly ubiquitous importance of top-down processes to perception, the usual design in these experiments does not identify exactly *how* top-down processes change individuals' observations and/or judgments. There are two general ways that information or expectations could play a role. First, it may be that top-down processes act indirectly on perception and decision-making by clouding the memories of an experience, creating an experimental demand effect, or creating a need for image consistency. The evidence from Allison and Uhl (1964), for example, could easily be explained by one of these mechanisms. That is, it could be that subjects tried to choose the "right" beer when they knew the brands, in an attempt to please the experimenter; or, subjects may have chosen the brand that they generally buy, so as to maintain consistency with their self-image. Numerous studies have shown, by giving individuals information that triggers top-down processes after the subjects are confronted with a stimulus, that top-down processes do affect sensory perception indirectly through one of these channels (Snyder and Uranowitz, 1978; Cohen, 1981; Stangor and McMillan, 1992; Cowley and Janus, 2004).² However, these designs do not prove that top-down processes have only an indirect impact, since they do not include a treatment in which subjects receive information before they experience the stimulus.

Top-down processes may also influence perception indirectly by changing the level of attention or intensity of search during an experience. One example of how this might occur is the affective expectation model (AEM), developed by Wilson and Klaaren (1992), which states that affective expectations induce a readiness to perceive an experience as being positive or negative; the theory then posits that, when actually faced with the experience, people may substitute expectations for paying close attention to the attributes of the stimulus (so as to conserve observation and mental-processing resources). Thus, if the true valence of the experience differs from the expected valence, then the AEM predicts that people will sometimes fail to exert the necessary effort to detect the discrepancy. For example, if subjects are told that a movie is awful, when watching it they may not evaluate its true merit as carefully as they would if they were given no information about the movie prior to viewing it. In fact, if subjects have negative expectations for an experience, it seems plausible that they might actively exert effort to avoid perceiving the actual attributes of the experience. In the Allison

²Although these studies focus on the memory explanation, their designs generally cannot rule out as explanations demand effects or the need for image consistency.

and Uhl (1964) beer study, for example, subjects with negative expectations for a beer based on its brand may have swallowed it very quickly or tried to ignore their perceptions of its taste in order to avoid the anticipated negative flavor.

The AEM therefore suggests that experiencing relatively unknown or ambiguous stimuli will result in incorrect assimilation of the experience with expectations more often than experiencing known stimuli will, since more effort is required to perceive the true nature of a stimulus when it is unfamiliar. A few empirical studies have found support for this claim by showing that proxies for ease of detection are correlated with sensitivity to expectations.³ However, the evidence for this specific prediction of the AEM is not yet definitive, since these studies cannot rule out the possibility that ease of detection may be endogenously tied to expectations. That is, it could be the case that individuals who can more easily observe the details of some stimulus also have preset expectations about the stimulus.

Alternatively, expectations may directly influence the perceptions that people have when presented with a stimulus (Lee, Frederick, and Ariely, 2006). For example, in the Allison and Uhl (1964) study, the brand information may have actually set subjects' taste buds and/or taste processing centers to respond differently to the beer in the experimental treatment than in the control treatment. This hypothesis has strong implications for decision-making theory. It states that the interpretation of an experience—as determined by individuals' sensory organs and the part of the brain that interprets signals from these organs—may change without any change in attention to, or uptake of, the stimulus, when information is provided before observation. Under this hypothesis, individuals have not overlooked anything when their perceptions are drawn toward their expectations (as is assumed in the AEM); their brains are simply primed to experience the stimulus differently. In the context of decision-making theory, this direct-effect hypothesis suggests that an individual's *correct* interpretation of the stimuli around him depends, in part, on his state of mind.

Lee, Frederick, and Ariely (2006), Hoch and Ha (1986), and Braun-LaTour and LaTour (2005) provide some evidence in favor of the direct-effect hypothe-

³Geers and Lassiter (2005) find that experience with a stimulus (in their case, a movie clip), which should allow for easier detection of discrepancies between expectations and experience, is associated with lower sensitivity to expectations. Deliza, MacFie, and Hedderley (1996) find that pre-measured ability to rate sweet and bitter solutions is associated with lower sensitivity to expectations when subjects taste similar solutions, while Deliza, MacFie, and Hedderley (2003) find that higher Need For Cognition scores predict lower sensitivity to superficial packaging features of passion-fruit juice.

sis,⁴ although only Lee, Frederick, and Ariely (LFA) choose the third hypothesis as an explanation of their results. Partly because Hoch and Ha (HH) and Braun-LaTour and LaTour (BL) were not focused on testing this hypothesis, all three of these papers have issues that potentially confound their ability to separate the two hypotheses. This paper critiques the methodology and evidence in those studies, focusing on LFA, and conducts a new field experiment that attempts to address some of the simpler issues that arise in their designs. The results of this experiment very weakly support the conclusion drawn in LFA that top-down processes can affect perception directly. Specifically, I find very preliminary evidence supporting LFA’s findings for men, but not for women, implying that any effect of top-down processes depends on individuals’ characteristics and/or context. In other words, the results in this paper suggest that the type of situation in which top-down processes might have an effect is somewhat restricted. However, even this conclusion stands on weak ground, because of the particular difficulty in distinguishing direct effects from indirect effects through attention and search channels. A number of other design challenges and confounding explanations, some of which are addressed in this paper, are outlined below. In short, the difficult question of whether perceptions are directly altered by context-generated top-down processes is likely far from resolved.

My discussion of these issues is organized into five additional sections. The next section outlines the general design framework used to test the indirect- and direct-effect hypotheses and discusses the main issues complicating the separation of these hypotheses. Section 3 identifies the problems in LFA, BL, and HH’s designs; section 4 describes this paper’s experimental methodology; section 5 presents the results; and section 6 concludes with an in-depth discussion of how future experiments might solve the issues raised throughout this paper.

2 Design Issues

LFA, HH, and BL use three between-subjects treatments that arguably test whether expectations⁵ change perceptions directly. In the first treatment (referred to as the “before” treatment henceforth), subjects receive expectations-

⁴See also Levin and Gaeth (1988) and Marks and Kamins (1988), although these papers have a somewhat different focus and thus lack good control conditions.

⁵The rest of the paper will focus on how expectations influence perceptions, since this is the top-down channel addressed by LFA, HH, BL, and the field study presented here. While the analysis may apply to other top-down processes as well, such extension is beyond the scope of this paper.

generating information before they sample some stimulus and then are asked to provide feedback on the stimulus. In the second treatment, subjects receive the same information after they sample the stimulus, but before they are asked to provide feedback (the “after” treatment). In the third treatment (the control treatment), subjects receive no additional information about the stimulus that they are asked to sample.

Based on the previously cited studies, which employ either a before treatment or an after treatment, both treatments may alter reported experiences. In particular, any difference between the after treatment and the control treatment may be due only to indirect effects of expectations on perception that occur after the experience is over. That is, in the after treatment, there can be no effect of expectations on perceptions *during* the uptake of the stimulus. The before treatment includes these post-experience indirect effects, as well as the effect of expectations on the experience itself, including changes in attention or search effort. Thus, if the before treatment changes behavior significantly more or differently than the after treatment (compared with the control), then expectations must somehow affect perception during the experience, either by changing the degree of attention paid to the experience or by actually changing perceptions directly.

This design seems effective at separating post-experience indirect effects from indirect effects through attention and effort channels, and direct effects. However, distinguishing between these latter two mechanisms is much more difficult. To eliminate changes in attention as an explanation, one must be able to control how much attention is paid to the sensory experience, which seems difficult to do. This study does not address this problem head-on, although the conclusions offer some potential solutions.

Apart from the confounding influence of attention, there are a number of underlying design details that have potentially important effects on the results of this type of experiment. First, it is important to ensure that subjects do not glean unintended information from the descriptions of the stimuli in the after and control treatments. If stimuli are given suggestive names or placed in informative context *and* these characteristics interact with the information given to subjects in the before treatment, then the information that subjects effectively obtain from the before treatment will be different from the information in the after treatment. Such differences could lead to variation in responses between the before and after treatments. Ideally, in order to narrow in on the indirect-effect and direct-effect hypotheses, the names and context of the stimuli should

give subjects as little information as possible, to minimize the possibility of introducing this confounding influence.

Second, in conducting the experiment, it became clear that the after treatment was difficult to administer. In a pilot study, subjects were told before experiencing the stimulus (in all treatments) that the experimenter wanted to know which of two samples they preferred. However, this made it difficult to ensure that the release of information in the after treatment occurred after the subjects had finished experiencing the samples but before they had expressed their preference, since many blurted out their preferences before they were prompted (even when they were specifically told to wait until asked before stating their preference). Thus, in order to ensure that individuals would not reveal their preference prematurely in the main study, the experimenter informed subjects in all treatments that they would be asked a question after they had sampled the stimuli, but they were not told ahead of time what the question would be.⁶

Third, heterogeneity across individuals may generate misleading results with this type of design. The basic intuition behind this argument is that using differences in aggregate percentages to compare choices in the three between-subjects treatments can overlook differences in how the treatments affect different individuals. These treatments involve at least four different factors that may vary across subjects. First, individuals may differ in the valence of their expectations resulting from the same information. In this experiment, for example, which asks subjects to sample full-strength lemonade and diluted lemonade, many subjects probably formed different expectations for one sample versus the other when they learned the contents of the samples. Second, individuals' responses to expectations may depend heterogeneously on their valence. For example, some subjects may respond more strongly to negative stimuli than others. Third, some subjects may respond more strongly to information in the before treatment relative to the after treatment, while others may respond more strongly to information in the after treatment. Finally, subjects may simply differ in their blind response to the stimulus. Correlations among these four sources of heterogeneity, as well as their potential interactions with the timing of the provision of information in the experiment, could generate almost any pattern of

⁶Another way to fix this problem would be to tell participants what they would be asked, but to leave them alone in a room to do the sampling and have them receive information and record their answers on paper. Without anyone to talk to, subjects would be unable to verbally commit to a preference before receiving information in the after treatment. Unfortunately, my recruiting method, as outlined in section 4, did not allow for this procedure. I discuss other possible solutions to this problem in the conclusions.

results. Section 3 offers an example of how LFA’s results could be generated by individual heterogeneity.

The best way to control for individual heterogeneity across these characteristics would be to incorporate a within-subjects design. However, since the different information timing in the three treatments essentially requires a between-subjects design, a second-best solution is to try to measure as many individual characteristics as possible (such as Need For Cognition scores, as in McClure et al. (2004)) and to control for them in the data analysis as proxies for individual heterogeneity in the experiment. In this study, in order to obtain a sufficient sample size, each subject engaged in the experiment for only a few minutes. However, subjects did complete a short survey asking for their gender, age, and lemonade-preference/experience data in order to allow for some testing of whether individual characteristics matter for the sensitivity to information in the before and after treatments. As noted below in section 5, the results confirm the earlier evidence that information sensitivity is heterogeneous across individuals. More specifically, males and females exhibit a very different pattern of preferences across the three treatments in the data. Thus, individual characteristics do seem to predict different information sensitivities in the before and after treatments, suggesting that (unless a within-subjects design can be devised) future research should attempt to control for as many individual characteristics as possible when studying this issue.

These issues, along with some possible solutions, are described in more detail in the conclusion.

3 Literature Critique

In their study, LFA ask participants to sample two beers (one normal, one with balsamic vinegar added) and choose which they prefer. In addressing whether changes in attention can explain their results, LFA claim that it is unlikely that participants spent more or less time or attention on tasting the two beers when they knew their contents, since this is a very simple activity. BL generally take a more agnostic view on this psychological issue and focus instead on the simple question of whether the timing of learning information changes the importance accorded to it relative to the importance of the stimulus; however, the sensory experience in BL’s study—tasting a sample of orange juice tainted with water, vinegar, and salt—is also quite simple. HH use a more

complex stimulus: comparing many different brands of shirts. In fact, they note that participants in the before treatment spend more time examining the shirts than participants in the other two treatments do.⁷ Thus, it seems unlikely that HH can determine whether expectations change attention or instead have a direct effect on perception. Whether either LFA or BL distinguish between these two possibilities is also debatable. Particularly in LFA’s study, it may be that subjects actually exerted effort in the before treatment to avoid tasting the vinegar-tainted beer, since most had a negative expectation for this beer’s flavor.⁸ The concluding section addresses this issue in more detail, since the field experiment employed here relies on the same argument that LFA make against attention as a confounding explanation.

BL do not aim to answer the questions raised here and, perhaps for this reason, they also have a seriously complicating design problem not discussed in section 2. At the beginning of the reporting stage in BL’s experiment, subjects must try to identify the orange juice that they originally drank from among five similar samples of orange juice. Although the ability to identify the treatment juice may be a useful measure, this test unfortunately changes the important sequencing of experience and information in the before and after treatments. By adding an extra period of experience (albeit a complicated one with five similar samples) following the before, after, and control treatments, BL give another opportunity, in all treatments, for information to have an impact during the experience. Ignoring this issue, BL do find results consistent with the conclusion that information directly changes perception. However, given the importance of this design issue, their evidence remains only suggestive.

Thus, I will focus on the study by LFA, which was designed to determine whether expectations change perceptions and, therefore, suffers from the fewest limitations to addressing this question. However, LFA’s experiment still includes some of the confounding design features described in section 2.

First, in LFA’s experiment, before subjects sample the two beers, they are told that the unmodified beer is “regular Budweiser” (or Sam Adams) and that

⁷This result differs from that posited by the AEM. However, the information given to participants in HH was an exaggeration, stating that J. C. Penney shirts were of very high quality. If participants were somewhat familiar with the quality of different brands of shirts, such information may have caused confusion and led them to examine shirts more carefully. Alternatively, HH argue that individuals engage in “hypothesis testing” when they have expectations about an experience.

⁸LFA ask a separate sample of individuals whether adding vinegar to beer would make it taste better or worse. Within this sample, 80 percent thought that plain beer would taste better.

the beer with vinegar added is “MIT brew.” In the before and after treatments, they are also told (before and after sampling, respectively) that MIT brew contains a few drops of balsamic vinegar. However, as discussed in section 2, the first step—telling the subjects the names of the two beers—may activate additional top-down processes that differ between treatments and thus muddle the management of information timing between treatments. For example, it may be that subjects in the after and control treatments have positive expectations for a beer (with unknown contents) named MIT brew, particularly in comparison with Budweiser or Sam Adams, while subjects in the before treatment, who know that MIT brew is simply Budweiser or Sam Adams with vinegar added, have negative expectations for the MIT brew. The relative nature of these expectations is important, since subjects must choose between the two beers. In fact, LFA note that many of their subjects complained that Budweiser doesn’t deserve to be called beer, suggesting that its name actually could create negative expectations for individuals in the after and control treatments compared with their expectations for the mystery MIT brew. Under this logic, LFA’s finding that a much smaller percentage of subjects prefer MIT brew in the before treatment than in the after or control treatments is not surprising (see Figure 1 for a comparison of preferences across LFA’s three treatments). As described in section 4, this issue of extra information is one that I address in my experiment, by using the same generic name for both lemonade samples in all treatments.

Second, regarding the difficulty in administering the after treatment, it is unclear exactly what procedure LFA used for their treatments. However, given the subtleties involved in this aspect of the design, differences in when subjects knew what they would be asked may explain some of the discrepancies between our results.

Third, although LFA do not measure any individual characteristics, individual heterogeneity in expectation valence, sensitivity to expectation valence and the timing of receipt of information, and blind preferences may all impact their results. For example, it seems plausible that some subjects would expect vinegar-tainted beer to taste better (especially if they don’t really like the regular beer) and that some would expect it to taste worse (which is what LFA intended their subjects to think).⁹ That is, the information provided may not have the same valence across individuals. Additionally, individuals probably

⁹As mentioned in an earlier footnote, LFA actually find that 80 percent of a separate subject pool felt that adding vinegar to beer would make it taste worse. This result leaves a non-trivial 20 percent who felt that vinegar would enhance the beer’s flavor.

have different sensitivities to information and the timing of receipt of information. Finally, while the majority of individuals in LFA's sample preferred vinegar-tainted beer when tasting the beers blind, 40 percent still preferred regular beer. By themselves, these heterogeneities do not weaken LFA's conclusions. However, interactions among the three may create problems. That is, subjects' sensitivity to information may depend not only on the timing of information, as LFA assume, but also on the valence of the information and on the subjects' blind preference between the two beers (as well as on interactions among the three factors). It is not difficult to construct interacting relationships among the effects on sensitivity of these three factors (valence, timing, and blind preference) that could generate results similar to LFA's without supporting their conclusions.¹⁰

Such relationships may not be the most likely explanation behind LFA's results, but they are not impossible. In fact, Deliza and MacFie (1996) review literature showing that individual variation in pre-information confidence levels (with respect to the judgment task), need for closure, investment in the task, levels of self monitoring, and scores on the Private Body Consciousness test all correlate with the effect of information on preferences. The evidence cited above concerning the AEM model also shows that the sensitivity of perceptions to information generally depends on the specific characteristics of individuals. Thus, if these heterogeneities also correlate with the expected valence of the information, all that is necessary to generate the confounding relationships described above is for these individual differences to correlate with the information effects differently in the before and after treatments. Looking at the multitude of relevant personal characteristics, this combination of correlations does not seem to be very far-fetched.

The next section discusses the details of my methodology, which is designed

¹⁰Consider the following hypothetical: Assume that individuals who expect vinegar-tainted beer to taste worse assimilate their experience to this expectation in the before and after treatments, while those who expect the vinegar-tainted beer to taste better are affected by information only in the after treatment. Then, moving from the control to the after treatment, there should be some preference shift from vinegar-tainted to regular beer, due to the effect of information acting on the group for whom vinegar has negative valence, and some preference shift in the opposite direction, due to the effect of information acting on the group for whom vinegar has positive valence. These effects may balance each other somewhat. However, if the positive-valence group is not affected by information in the before treatment, then the preferences may shift in favor of regular beer in the before treatment. This pattern could generate results very similar to LFA's, and it does not assume that anyone is more strongly affected by information in the before treatment than in the after treatment. Such an example is not entirely implausible, as asymmetries between responses to negative and positive information are common (for an oft-cited example, see the evidence on loss aversion concerning hypothetical risk questions in Kahneman and Tversky (1979)).

to address, at least in part, the issues described in this and the previous section. Section 5, focuses specifically on how the differences between this paper’s methodology and LFA’s may have led to the differences in our results.

4 Methodology

The subjects in this study are volunteers who passed by a free lemonade stand on a bike/walking path near Davis Square, in Somerville, MA. Eighty individuals were recruited during a few August and July afternoons using a sign reading, “Free Lemonade! Participate in a brief research study.”¹¹ Only 75 subjects are included in the summary statistics presented in section 5, since five could not choose between the two samples of lemonade that they were offered.¹² Two more individuals are also excluded from the regressions that control for individual characteristics after failing to answer all of the survey questions.

The experimental procedure is based on that used by LFA, with some minor modifications designed to address the issues raised in section 3. When individuals expressed interest in participating in the study, they were told that they would taste two small samples of lemonade (about two ounces each). Both lemonade samples were made from frozen concentrate, one according to the recipe on the can, and the other with the same recipe, but with four ounces of extra water per 28 ounces of regular (that is, recipe-based) lemonade. To ensure that the subjects did not receive extraneous information through the names given to the two lemonades, both were simply called “samples of lemonade” in all three treatments.¹³

In the before treatment, subjects were told, before tasting, that one sample

¹¹In order to encourage participation and to control loosely for temperature effects, the experiment was conducted on days with a forecasted high temperature of over 80 degrees Fahrenheit. Although the design suggests a bias toward subjects who really like lemonade, many subjects did not realize that the study involved tasting lemonade (that is, they thought the lemonade was a reward given for participation). Some of these subjects actually declined the offer of lemonade before they were informed that sampling the lemonade was required.

¹²Including these individuals in the summary statistics and the regressions does not change the general conclusions drawn from the analysis.

¹³If subjects suspected that the two samples were different in the after and control treatments, they may have felt that the experimenter was trying to deceive them, perhaps leading to increased attention during the sampling process in order to validate this hypothesis (see the earlier footnote concerning the study by Hoch and Ha (1986)). However, this outcome seemed worth risking in order to eliminate any confounding information effects, since “lemonade” is also arguably a general term that describes varying types of lemonade. Anecdotal evidence confirms this claim: no subjects explicitly suspected this type of deception, although many seemed unsure whether the lemonade samples were the same or different, suggesting that they received no information from the names of the two samples.

was regular lemonade and that the other was slightly diluted; in the after treatment, this information was given immediately after tasting. Subjects in all three treatments were told, before tasting, to sample as much of the two lemonades as they wanted and that they would be asked a question when they were finished. As described in section 3, in order to control the timing of subjects' responses in the after treatment, it was important that subjects not know exactly what they would be asked. Otherwise, subjects tended to give their responses before they were told the difference between the two samples in the after treatment.¹⁴ Finally, once subjects stated that they were finished sampling, they were told the difference between the two samples in the after treatment and, in all treatments, were asked which sample they preferred.

After completing the experiment, subjects filled out a short questionnaire that asked for their age, their gender, how many times they drink lemonade per week (on average), and how much they like lemonade (on a one to seven scale). Women are oversampled somewhat (47 women to 33 men in the full sample), and most of the individuals in the study are under 30 years old, drink lemonade less than twice a week, and like the taste of lemonade. Although these questions do not completely address the heterogeneity concerns raised above, they do allow for tests of individual heterogeneity in the information effects among the three treatments by looking at whether differences among treatments in decision-making differ according to the measured individual characteristics. Additionally, the chosen pair of stimuli (lemonade and diluted lemonade) has valences that are probably quite variable across individuals, creating good potential for such heterogeneity.

5 Results

In order to assess the treatment effects in this design, this section follows LFA and compares the proportions who chose each of the two samples of lemonade between treatments. Perhaps because of the differences between the design used here and LFA's, this experiment did not generate the same general distribution of choices among the three treatments (see Figures 1 and 2). LFA's results show that information in the after treatment moves preferences slightly toward regular

¹⁴Some subjects asked directly what they should be looking for when sampling. The same request was repeated to these individuals, asking them to try as much as they wanted of each one and stating that they would be asked a question afterwards. This procedure may have confused some subjects, but it was necessary to ensure consistency in the after treatment.

beer, while information in the before treatment does so significantly. In contrast, information in the after treatment in this experiment increases choices of diluted lemonade, while information in the before treatment has essentially no effect compared with the control treatment. Specifically, 50 percent of individuals in the control treatment and 53.8 percent of individuals in the before treatment preferred the diluted lemonade, compared with 60 percent of individuals in the after treatment. Although the proportion choosing the diluted sample is a bit higher in the after treatment, none of the pairwise differences among the three treatments are statistically significant. Thus, information seems to have no effect on preferences in the before treatment and very little (if any) effect on preferences in the after treatment. The regressions in Table 1—both with and without demographic and lemonade-preference controls—confirm the lack of an information effect in either treatment.¹⁵

From a simple perspective, these results may be due to the basic differences between the lemonade experiment and LFA's. By giving subjects no extraneous information about the lemonade samples in the control treatment, one potentially spurious generator of differences among the treatments was eliminated. Therefore, one possible conclusion to draw from these results and LFA's is that the differences between the before and control treatments observed by LFA are mostly due to this extra information, rather than to the presence and timing of information. The procedure here may also differ somewhat from LFA's in that subjects were not told the question that they would be asked before tasting the lemonade. It is unclear whether this could have generated substantially different results between our studies, although this design was necessary to ensure consistent application of the after treatment.

These two procedural differences may explain a lot of the differences between our results, but it is worth noting that the results in this paper for the before and control treatments are somewhat unexpected, given the large body of literature showing an effect of information on preferences when the information is provided before the sensory stimulus (see the discussion in section 1). Although surprising, this difference actually provides some indirect support for the hypothesis that information effects are individually heterogeneous. The information in this experiment did not have a clear valence,¹⁶ while the information given in most

¹⁵All regressions in Tables 1 and 2 are linear probability models. Using a probit specification instead does not change the results substantially.

¹⁶Although diluted lemonade may seem unappetizing to many, someone taking a break from a jog to participate in this study might very well have had a positive expectation for the diluted sample.

of the experiments referenced earlier had an intended valence. If individuals with different valences for the information had different (blind) preferences for the two lemonades or were affected differently by the information in this study, then these varying preferences or information effects could easily have led to the observed non-result. In LFA’s study and other studies, the information was designed to be strongly negative, meaning that the effect of negative information should dominate and create observable differences between the information treatments and the control treatment.¹⁷

Since the data include a few personal characteristics, it is possible to run a simple test of whether heterogeneity explains the non-result. The logic in this test is that, if individual heterogeneity in valences or information effects correlates with any of the individual variables that were measured, then these variables can proxy for the underlying heterogeneity in the experiment. The results of the regressions suggest that male and female subjects seem to behave differently in the experiment. As suggested by the significant coefficient on the male dummy in the second regression model in Table 1, a higher proportion of men (67.7 percent) than women (45.5 percent) prefer diluted lemonade ($p = 0.029$, unpaired t -test¹⁸). Interestingly, this difference does not hold across treatment groups. As shown in Figure 3, a much larger proportion of men than women prefer diluted lemonade in the control and before treatments (for the control treatment difference, $p = 0.045$; for the before treatment, $p = 0.018$), while an equal percentage (60 percent) of both men and women prefer diluted lemonade in the after treatment.¹⁹ Although gender is only a proxy for individual heterogeneity in the information effects across treatments, these results suggest that heterogeneity is in fact present. Probably because of the small size of the sample, regression analysis does not yield definitive evidence in support of this conclusion (see the tests for different treatment effects by gender in Table 2). Still, the summary-statistic results, coupled with the previous evidence cited above, suggest that individual heterogeneity can explain the lack of an overall information effect in this experiment and confound the results of this study.

¹⁷It is important to note, however, that individual heterogeneity can still play a confounding role in LFA’s results, as described in section 3. Also, as mentioned above, the strong negative valence in LFA’s design may have caused subjects to exert effort—for example, by drinking the vinegar-tainted beer quickly—to avoid experiencing the beer’s actual flavor.

¹⁸All p -values given in this section are generated using unpaired t -tests.

¹⁹It is important to note that bad luck through random assignment into the three treatments resulted in only five male subjects participating the after treatment (against 20 women) and only eight women participating in the control treatment (against 16 men). A more even sampling may have generated a clearer comparison, particularly in the regression analysis.

Finally, it is worth noting that the pattern of male responses actually matches LFA’s results quite closely (see Figures 1 and 3). Although in this study the across-treatment differences for males are not significant, and, based on the issues raised to this point (including the sample size), it seems overly ambitious to claim that these results are conclusive, still, confirmation of the finding that information does affect perceptions directly, even if only in a greatly restricted context (for example, in the context “males sampling diluted lemonade” or “people tasting beer with vinegar added”), would be an important result. That is, it is possible that LFA simply found a special example where information has strong direct effects. This study’s results for males suggest that other such examples may exist.

6 Conclusions

This paper calls attention to a number of issues working against the current design of tests of whether information truly changes perceptions. Such a result could have major implications for decision-making theory, since it implies (under standard utility maximization, at least) that an individual’s state of mind determines the nature and strength of the perceived stimuli to which the individual responds in maximizing his utility. With an individual’s potentially fluctuating state of mind already leading to variable preferences over time, such a theory may be very hard to model and understand fully. It appears that it is quite difficult to show concretely that top-down processes directly change sensory uptake. This study, like those of LFA, HH, and BL, offers some evidence that information can have such an effect, but none of these studies can rule out all of the confounding explanations discussed here. These explanations include extraneous information, difficulties in designing the after treatment, individual heterogeneities, and attention effects, and there are probably others that have not been considered.

Although the design of the experiment reported in this paper mostly addresses the problems of extra information and after-treatment design, solving these problems completely may not be very difficult. Instead of trying to name stimuli generically as done here, a technique that may lead to confusion or suspicion about whether the stimuli are truly different, one could have subjects experience the stimuli in a manner that eliminates the need to explain to them beforehand what to expect in the after and control treatments. For example, the

stimulus could be a movie preview at a campus screening that would require no explanation; subjects would simply observe the preview passively without any name or context being assigned to the preview. A similar design could eliminate issues with the after treatment. If subjects in the movie-design study were given paper (or computers or cell phones for text messaging) and directed by instructions on the movie screen to record their responses without speaking to anyone, they would have little opportunity to reveal their preferences before the information could be given.

However, the issue of individual heterogeneity is probably more difficult to solve. The results of this study show very different choice patterns for men versus women across the three treatments, suggesting that heterogeneity really is a problem for this design. As discussed earlier, a within-subjects design would overcome this issue easily, but does not seem feasible, since giving subjects information more than once would inherently ruin the timing of information. Perhaps the best solution might be to measure as many characteristics as possible so as to control for them when measuring the information effects. A more creative solution might be to use a within-subjects design, but to use multiple stimuli (for example, beer, lemonade, and wine) and to ensure that each subject receives a different stimulus in each treatment. If the stimuli were randomly assigned across the three treatments, such a design could potentially control for heterogeneous interactions between information valence and sensitivity to information and information timing. One final possibility would be to control for heterogeneity in information valences by asking individuals, after the experiment, the valence that they associated with the information. In order to get an accurate response, one could use, for example, political propaganda as the stimulus, and political leanings (which would be unlikely to change during the experiment) as a proxy for the information valence.

Beyond all of these difficult issues, controlling for differences in attention and search effort may be the hardest. Using all of the design improvements employed so far, subjects might still choose to pay more or less attention to a stimulus in the before treatment than in the after or control treatments, as the affective expectation model predicts. HH actually observe such differences in the before treatment (although in a manner opposite to the AEM predictions); their subjects spend much more time examining the stimulus when they have already received the information than in the after and control treatments. Similarly, Petty, Cacioppo, and Schumann (1983) show that information effects on attitudes depend greatly on the subjects' level of involvement in the experience.

These attention effects are difficult to control for, but a first attempt might be to raise the subjects' level of involvement as high as possible or simply to encourage subjects to pay very close attention to the experience. Attempting to maximize attention in one of these two ways seems to be the simplest approach to getting similar attention levels from all subjects across treatments.

This question of subjects' attention levels seems to be the most difficult to deal with, although the other issues raised should not be neglected either. As Deliza and MacFie (1996) and LFA have noted, more research on information effects is needed before one can draw definitive conclusions. In particular, further study may alert us to better methods of controlling for attention levels and for the other issues discussed here. One possible avenue for new research would be to use brain-imaging tools, such as functional magnetic resonance imaging (fMRI), in order to determine which parts of the brain are most active during each stage of the three treatments (for a discussion of this type of research, see Camerer, Loewenstein, and Prelec (2005)). It is difficult to say whether such imaging could address all of the issues raised in this paper, but it could certainly suggest which of the potentially confounding factors are most relevant.

At the very least, a larger body of results using different stimuli and measurement techniques should offer some preliminary bounds on the possible range of situations in which top-down processes can change perception directly. As evidenced by the many confounding issues discussed in this paper, this question is not simple, and its solution may require a variety of approaches. However, its implications are important.

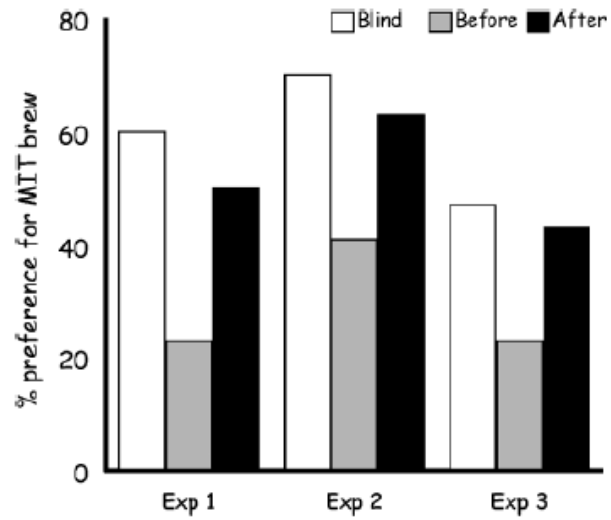
References

- Allison, Ralph I., and Kenneth P. Uhl. 1964. "Influence of Beer Brand Identification on Taste Perception." *Journal of Marketing Research* 1:36–39.
- Biederman, Irving. 1972. "Perceiving Real-World Scenes." *Science* 177:77–80.
- Bowen, Deborah J., Naomi Tomoyasu, Marin Anderson, Maureen Carney, and Alan Kristal. 1992. "Effects of Expectancies and Personalized Feedback on Fat Consumption, Taste, and Preference." *Journal of Applied Social Psychology* 22:1061–1079.
- Braun-LaTour, Kathryn A., and Michael S. LaTour. 2005. "Transforming Consumer Experience." *Journal of Advertising* 34:19–30.
- Brief, Arthur P., Ann Houston Butcher, Jennifer M. George, and Karen E. Link. 1993. "Integrating Bottom-Up and Top-Down Theories of Subjective Well-Being: The Case of Health." *Journal of Personality and Social Psychology* 64:646–653.
- Camerer, Colin, George Loewenstein, and Drazen Prelec. 2005. "Neuroeconomics: How Neuroscience Can Inform Economics." *Journal of Economic Literature* 43:9–64.
- Cohen, Claudia E. 1981. "Person Categories and Social Perception: Testing Some Boundaries of The Processing Effects of Prior Knowledge." *Journal of Personality and Social Psychology* 40:441–452.
- Cowley, Elizabeth, and Eunika Janus. 2004. "Not Necessarily Better, but Certainly Different: A Limit to the Advertising Misinformation Effect on Memory." *Journal of Consumer Research* 31:229–235.
- Darley, John M., and Paget H. Gross. 1983. "A Hypothesis-Confirming Bias in Labeling Effects." *Journal of Personality and Social Psychology* 44:20–33.
- David, James P., Peter J. Green, Rene Martin, and Jerry Suls. 1997. "Differential Roles of Neuroticism, Extraversion, and Event Desirability for Mood in Daily Life: An Integrative Model of Top-Down and Bottom-Up Influences." *Journal of Personality and Social Psychology* 73:149–159.
- Deliza, Rosires, and Halliday J. H. MacFie. 1996. "The Generation of Sensory Expectation by External Cues and its Effect on Sensory Perception and Hedonic Ratings: A Review." *Journal of Sensory Studies* 11:103–128.

- Deliza, Rosires, Halliday J. H. MacFie, and Duncan Hedderley. 1996. "Information Affects Consumer Assessment of Sweet and Bitter Solutions." *Journal of Food Science* 61:1080–1084.
- . 2003. "Use of Computer-Generated Images and Conjoint Analysis to Investigate Sensory Expectations." *Journal of Sensory Studies* 18:465–486.
- Friedman, Hershey H., and William S. Dipple, Jr. 1978. "The Effect of Masculine and Feminine Brand Names on the Perceived Taste of a Cigarette." *Decision Sciences* 9:467–472.
- Geers, Andrew L., and G. Daniel Lassiter. 2005. "Affective Assimilation and Contrast: Effects of Expectations and Prior Stimulus Exposure." *Basic and Applied Social Psychology* 27:143–154.
- Hoch, Stephen J., and Young-Won Ha. 1986. "Consumer Learning: Advertising and the Ambiguity of Product Experience." *Journal of Consumer Research* 13:221–233.
- Jones, Edward E., Leslie Rock, Kelly G. Shaver, George R. Goethals, and Lawrence M. Ward. 1968. "Pattern of Performance and Ability Attribution: An Unexpected Primacy Effect." *Journal of Personality and Social Psychology* 10:317–340.
- Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 47:263–292.
- Klaaren, Kristin J., Sara D. Hodges, and Timothy D. Wilson. 1994. "The Role of Affective Expectations in Subjective Experience and Decision-Making." *Social Cognition* 12:77–101.
- Lee, Leonard, Shane Frederick, and Dan Ariely. 2006. "Try It, You'll Like It: The Influence of Expectation, Consumption, and Revelation on Preferences for Beer." *Psychological Science* 17:1054–1058.
- Levin, Irwin P., and Gary J. Gaeth. 1988. "How Consumers Are Affected by the Framing of Attribute Information Before and After Consuming the Product." *Journal of Consumer Research* 15:374–378.
- Makens, James C. 1965. "Effect of Brand Preference Upon Consumers' Perceived Taste of Turkey Meat." *Journal of Applied Psychology* 49:261–263.

- Marks, Lawrence J., and Michael A. Kamins. 1988. "The Use of Product Sampling and Advertising: Effects of Sequence of Exposure and Degree of Advertising Claim Exaggeration on Consumers' Belief Strength, Belief Confidence, and Attitudes." *Journal of Marketing Research* 25:266–281.
- McClure, Samuel M., Jian Li, Damon Tomlin, Kim S. Cypert, Latané M. Montague, and P. Read Montague. 2004. "Neural Correlates of Behavioral Preference for Culturally Familiar Drinks." *Neuron* 44:379–387.
- Nevid, Jeffrey S. 1981. "Effects of Brand Labeling on Ratings of Product Quality." *Perceptual and Motor Skills* 53:407–410.
- Olson, Jerry C., and Philip A. Dover. 1978. "Cognitive Effects of Deceptive Advertising." *Journal of Marketing Research* 15:20–38.
- Palmer, Stephen E. 1975. "The Effects of Contextual Scenes on the Identification of Objects." *Memory and Cognition* 3:519–526.
- Petty, Richard E., John T. Cacioppo, and David Schumann. 1983. "Central and Peripheral Routes to Advertising Effectiveness: The Moderating Role of Involvement." *Journal of Consumer Research* 10:135–146.
- Snyder, Mark, and Seymour W. Uranowitz. 1978. "Reconstructing the Past: Some Cognitive Consequences of Person Perception." *Journal of Personality and Social Psychology* 36:941–950.
- Stangor, Charles, and David McMillan. 1992. "Memory for Expectancy-Congruent and Expectancy-Incongruent Information: A Review of the Social and Social Developmental Literatures." *Psychological Bulletin* 111:42–61.
- Wansink, Brian, Sea Bum Park, Steven Sonka, and Michelle Morganosky. 2000. "How Soy Labeling Influences Preference and Taste." *International Food and Agribusiness Management Review* 3:85–94.
- Wardle, Jane, and Wendy Solomons. 1994. "Naughty but Nice: A Laboratory Study of Health Information and Food Preferences in a Community Sample." *Health Psychology* 13:180–183.
- Wilson, Timothy D., and Kristin J. Klaaren. 1992. "Expectation Whirls Me Round": The Role of Affective Expectations on Affective Experiences." In Margaret S. Clark, ed. *Review of Personality and Social Psychology: Emotion and Social Behavior*. Newbury Park, CA: Sage, vol. 14, pp. 1–31.

Figure 1: Results from LFA



Notes: Figure taken from LFA. The three sets of bars represent three variations in LFA's study used to elicit preferences between their two beer samples.

Figure 2: The Effect of Information on Preferences For Diluted Lemonade

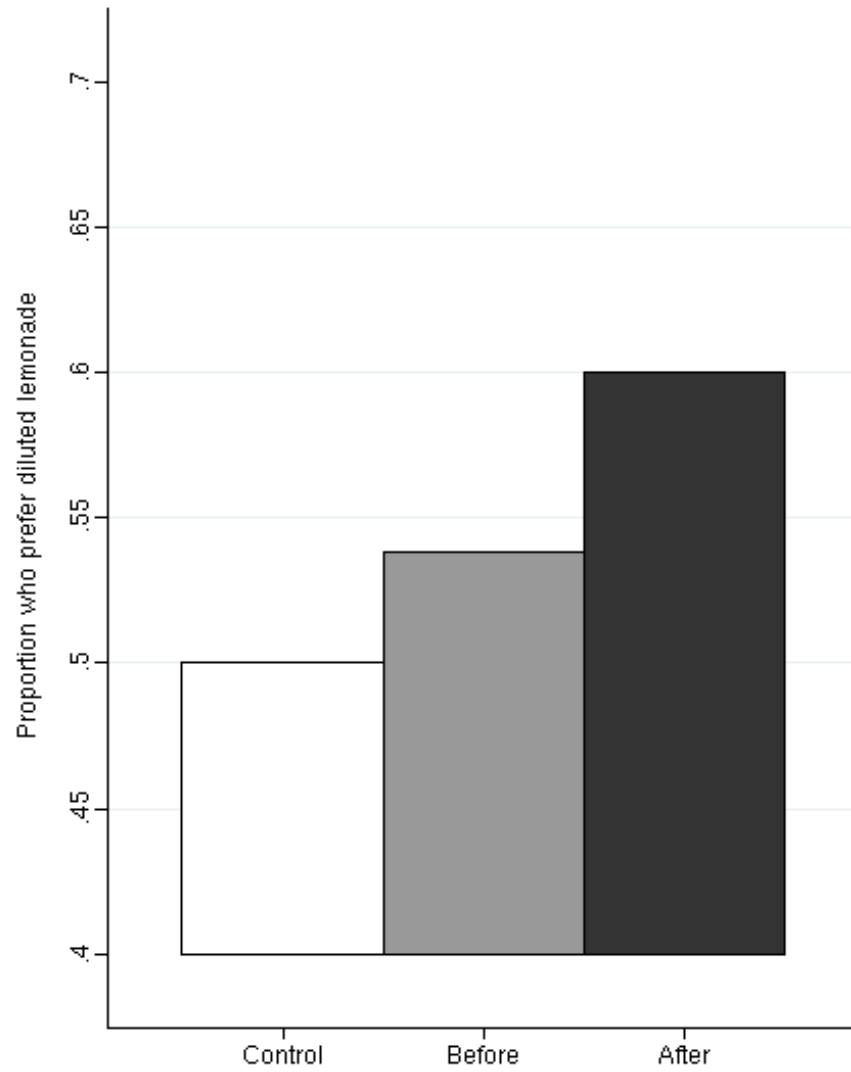


Figure 3: The Effect of Information on Preferences by Gender

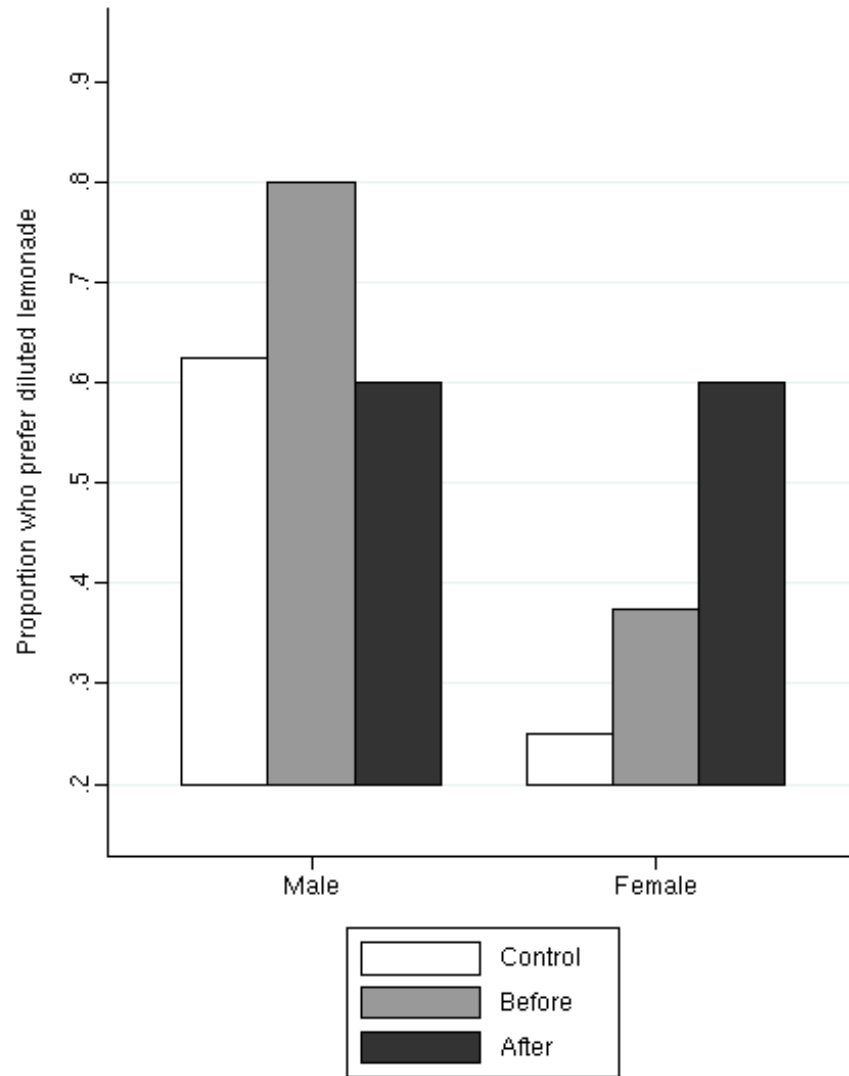


Table 1: Main Regressions

	(1)	(2)
Before treatment	0.038 (0.144)	0.104 (0.139)
After treatment	0.100 (0.144)	0.229 (0.158)
Male		0.287** (0.130)
Age		-0.002 (0.006)
Times per week		0.023 (0.040)
Like lemonade		-0.077 (0.053)
r ²	0.007	0.105
N	75	73

Notes: ** indicates significance at the 5 percent level. The dependent variable in both regressions is equal to one if the subject preferred the diluted lemonade and equal to zero if they preferred the full-strength lemonade. The specifications shown are linear probability models; probits yield similar results. The variable “Times per week” is based on categorical responses. Subjects could answer that they drink lemonade less than once per week, one to three times per week, four to six times, or more than seven times. The ranges are coded as 0.5, 2, 5, and 7 in the regression. “Like lemonade” is a variable ranging from one to seven, where one means that the subject does not like lemonade at all, and seven means that the subject likes it a lot.

Table 2: Interaction Regressions

	(3)	(4)
Male*Before treatment	0.175 (0.198)	0.145 (0.191)
Male*After treatment	-0.025 (0.251)	-0.040 (0.274)
Female*Before treatment	0.125 (0.212)	0.129 (0.203)
Female*After treatment	0.350* (0.205)	0.359* (0.197)
Male	0.375* (0.212)	0.387* (0.206)
Age		-0.002 (0.007)
Times per week		0.021 (0.039)
Like lemonade		-0.083 (0.051)
Are male/female profiles significantly different?	$p = 0.369$	$p = 0.400$
r ²	0.107	0.132
N	75	73

Notes: * indicates significance at the 10 percent level. The dependent variable in both regressions is equal to one if the subject preferred the diluted lemonade and equal to zero if they preferred the full-strength lemonade. The specifications shown are linear probability models; probits yield similar results. For independent variable definitions, see Table 1. The test for whether the male and female treatment effects are different is a joint comparison of the pairs of before- and after-treatment coefficients by gender.