Essays in Behavioral Decision Making

by

Kristina S. Shampanier

M.S. Mathematics, Moscow State University, Moscow, Russia, 2001
M.A. Economics, New Economic School, Moscow, Russia, 2002

Submitted to the Alfred P. Sloan School of Management
in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Management

at the

Massachusetts Institute of Technology

May 2007
[June 2007]

Signature of Author ....
Alfred P. Sloan School of Management
August 11, 2006

Certified by ....
Dan Ariely
Alfred P. Sloan Professor of Behavioral Economics
Thesis Supervisor

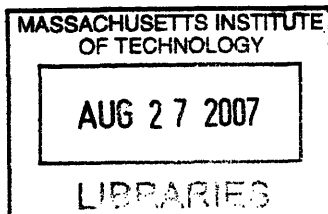Accepted by ....
Birger Wernerfelt
J. C. Penney Professor of Management Science
Chair, Doctoral Program

Essays in Behavioral Decision Making

by

Kristina S. Shampanier

Submitted to the Alfred P. Sloan School of Management
on May 2, 2007 in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy in
Management

ABSTRACT

Essay 1: Zero as a Special Price: the True Value of Free Products

When faced with a choice of selecting one of several available products (or possibly buying nothing), according to standard theoretical perspectives, people will choose the option with the highest cost–benefit difference. However, we propose that decisions about free (zero price) products differ, in that people do not simply subtract costs from benefits and perceive the benefits associated with free products as higher.
We test this proposal by contrasting demand for two products across conditions that maintain the price difference between the goods, but vary the prices such that the cheaper good in the set is priced at either a low positive or zero price. In contrast with a standard cost–benefit perspective, in the zero price condition, dramatically more participants choose the cheaper option, whereas dramatically fewer participants choose the more expensive option. Thus, people appear to act as if zero pricing of a good not only decreases its cost but also adds to its benefits. After documenting this basic effect, we propose and test several psychological antecedents of the effect, including social norms, mapping difficulty, and affect. Affect emerges as the most likely account for the effect.

Essay 2: Movies as a Mood Regulation Tool: Movie Watching Patterns Right After September 11.

Is a sad person more, less or equally likely than a happy person to pursue a "happy" activity rather than an "unhappy" one (e.g. prefer a comedy to a drama)? Surprisingly, the literature offers theories *and* laboratory evidence in favor of *all three* possibilities. In this paper I attempt to resolve the puzzle by moving out of the lab and analyzing the changes in movie watching patterns following the tragic events of Sep 11, 2001. Two data sets from the 7 weeks surrounding 9/11 are analyzed. One consists of US box office collections of top ten movies during the period. The other contains data on movie rentals in a rental store chain in Cambridge MA. The analysis suggests that the more private the

mood-regulating decision is (rental vs. movie going), the more likely is the person to use the movie as a mood repair tool. When the decision is more public (movie going), the appropriateness issues induce more mood congruent behavior.

Essay 3: Measuring Liking and Wanting

Recently neuroscientists have gathered a vast body of evidence that wanting (motivated preferences) and liking (non-motivated prferences) are not one and the same. We explore the possibility of measuring the two types of preferences uintrusiveley, in a behavioral lab. In particular we find that wanting and liking for viewing pictures of attractive people are not perfectly aligned and expessially for men.

Thesis Committee:

Dan Ariely (Chair)
Alfred P. Sloan Professor of Behavioral Economics
MIT Sloan School of Management

Leonard Lee
Assistant Professor of Marketing
Columbia Business School

Michael I. Norton
Assistant Professor of Marketing
Harvard Business School

Acknowledgements

The person who I would like to thank the most for helping me arrive at this point is my advisor at MIT, Dan Ariely for being a great advisor, friend, and a father figure. Dan, thank you for inspiration, encouragement, advice, mostly, thanks for believing in me when even I was not that sure.

Special thanks to Dan for contributing to all three essays of the dissertation, and to Leonard Lee and Jonathan Levav for contributing to the second essay.

I would also like to thank all the students in the program, the faculty and the friends, some in more than one category. To John Hauser, John Little, Drazen Prelec, Duncan Simester, Birger Wernerfelt, and Glen Urban: thanks for setting an exemplar of great academicians. To Shane: thanks for keeping an eye on the IQ level. To Allison, Barak, Catherine, Elan, Gui, Juanjuan, Mike, On, Rebecca, Robert, and Sharmila: thanks for being great colleagues. To Boris, Daniel, Han, Matt, Nathan, Ray, Rene, and Zoë: thanks for all the fun we shared. To Sharon, Hillary, Drew, Eric, Jay, Rosa, Sandra, and Taya: thanks for all the saved days. To, Leonard and Mike, thanks for being on my committee. To Anat and Nina: thanks for always being there. Thank you, Marketing Family, for all the lunches, discussions, for making me feel at home in E56 and E40, for being a family.

I would also like to thank my advisors at MSU, Alexey and Alexander Mikhalev, and my advisors at NES, Sergey Guriev, Konstantin Sonin, and Michael Schwarz for laying the foundation.

Special thanks to my parents – for laying the foundation for the foundation, – my grandmas, and my grandfather.

Thanks to my friends in Moscow, Jane, Masha, Vitalik and all the friends in Boston.

Thanks to my boyfriend Andrey, for putting up with me during my final year at MIT.

Essay 1
Zero as a special price: The true value of free products

Abstract

When faced with a choice of selecting one of several available products (or possibly buying nothing), according to standard theoretical perspectives, people will choose the option with the highest cost–benefit difference. However, we propose that decisions about free (zero price) products differ, in that people do not simply subtract costs from benefits and perceive the benefits associated with free products as higher.

We test this proposal by contrasting demand for two products across conditions that maintain the price difference between the goods, but vary the prices such that the cheaper good in the set is priced at either a low positive or zero price. In contrast with a standard cost–benefit perspective, in the zero price condition, dramatically more participants choose the cheaper option, whereas dramatically fewer participants choose the more expensive option. Thus, people appear to act as if zero pricing of a good not only decreases its cost but also adds to its benefits. After documenting this basic effect, we propose and test several psychological antecedents of the effect, including social norms, mapping difficulty, and affect. Affect emerges as the most likely account for the effect.

Zero as a special price: The true value of free products

1. Introduction

"The point about zero is that we do not need to use it in the operations of daily life. No one goes out to buy zero fish. It is in a way the most civilized of all the cardinals, and its use is only forced on us by the needs of cultivated modes of thought."

—Alfred North Whitehead

Initially invented by Babylonians not as a number but as a placeholder, the concept of zero and void was feared and denied by Pythagoras, Aristotle, and their followers for centuries. The most central objection of the early Greeks to zero was based on religious beliefs; they argued that god was infinite and therefore void (zero) was not possible. In addition to religious arguments, the early Greeks did not recognize their need for zero, because their mathematics were based on geometry, which made zero and negative numbers unnecessary. This failure to adopt the concept of zero likely impeded their discovery of calculus and slowed the development of mathematics for centuries.

The concept of zero as a number was brought to India by Alexander the Great, where it was first accepted. In India, unlike Greece, algebra was separate from geometry, infinity and void appeared within the same system of beliefs (i.e., destruction, purity, and new beginnings), and the concept of zero flourished. The notion of zero later found its way into Arabia and later immigrated to Europe. Because Aristotle had not accepted zero and because Christianity was partially based on Aristotelian philosophy and his "proof of God," zero was not widely embraced by the Christian world until the sixteenth century.[1]

---

[1] For a good source describing the history of zero, see Seife (2000).

8

In more recent history, the concept of zero enters into the understanding of multiple aspects of human psychology. In various domains, zero is used in a qualitatively different manner from other numbers; and the transition from small positive numbers to zero often is discontinuous.

Cognitive dissonance theory (Festinger and Carlsmith 1959) shows that getting a zero reward can increase liking for the task compared with receiving a small positive reward. Subsequent work reveals that changing a reward from something to nothing can influence motivation (Festinger and Carlsmith 1959) and switch it from intrinsic to extrinsic (Lepper, Greene, and Nisbett 1973), alter self-perception (Bem 1965), and affect feelings of competence and control (Deci and Ryan 1985). For example, Gneezy and Rustichini (2000a) demonstrate that introducing a penalty for parents who are late picking up their children from kindergarten can actually increase tardiness. Similarly, Gneezy and Rustichini (2000b) find that though performance in tasks such as IQ tests or collecting money for charity increases, as expected, with the size of a positive piece-wise reward, the zero reward represents an exception in which performance is greater when no reward is mentioned relative to when a small reward exists.

Related to these findings on motivation and incomplete contracts, it has also been shown that when prices are mentioned, people apply market norms, but when prices are not mentioned (i.e., the price effectively is zero), they apply social norms to determine their choices and effort (Heyman and Ariely 2004). As an illustration, Ariely, Gneezy, and Haruvy (2006) show that when offered a piece of Starburst candy at a cost of 1¢ per piece, students take approximately four pieces; when the price is zero, more students take

the candy, but almost no one takes more than one piece (i.e., decreased demand when prices are reduced).

Finally, in a different domain and in the most influential research on the psychology of zero, Kahneman and Tversky's (1979) work on probabilities indicates that when it comes to gambles, people perceive zero probability (and certainty) substantially differently than they do small positive probabilities. That is, whereas the values of the latter are perceived as higher than they actually are, perceptions of zero probability are accurate.

In this work, we extend research on the psychology of zero to pricing and examine the psychology of "free." Intuition and anecdotal evidence suggest that in some sense, people value free things too much. When Ben and Jerry's offer free ice-cream cones, or Starbucks offers free coffee, many people spend hours in line waiting to get the free item, which they could buy on a different day for two to three dollars. At first glance, it might not be surprising that the demand for a good is very high when the price is very low (zero), but the extent of the effect is intuitively too large to be explained by this simple economic argument. The goal of this paper is to examine the validity of this intuition, and to establish the causes of the phenomenon.

In a series of experiments, we demonstrate that when people are faced with a choice between two products, one of which is free, they overreact to the free product as if zero price meant not only a low cost of buying the product but also its increased valuation. In the next section, we describe a method to examine reaction and overreaction to free products. In Section 3, we detail two formal models: one that treats the price of zero as any other price, and one that includes a unique role for zero. The contrasts

between these two models provide some predictions for the effects of price reductions on demand. Then, in Section 4, we report experimental evidence in support of the zero-price model. We take a first step in finding the psychological causes that bring about the effect of zero price and test them in Section 5, then end with general conclusions and some questions for further research.

2. Measuring Reaction/Overreaction to Zero Price

To determine if people overreact to free products, we might simply test whether consumers take much more of a product when it is free than they buy of the product when it has a very low price (e.g., 1¢). However, though such behavior would be consistent with an overreaction to free, it also could simply reflect an increase in demand when price decreases. Similarly, it is not sufficient to show that the increase in demand when price falls from 1¢ to zero is greater than the increase in demand when the price drops from 2¢ to 1¢ because such a pattern of behavior could reflect a demand structure that is nonlinear in price (e.g., created by a valuation distribution in which more people value the product between 0¢ and 1¢ than between 1¢ and 2¢).

To measure reaction to zero and overcome these possible alternative interpretations, we examine whether people select a free product even when they must forgo an option that they "should" find preferable. We employ a method that contrasts two choice situations that involve a constant difference between two products' net benefits and use aggregate preference inconsistency as a measure of overreaction to the free product. The basic structure of this approach (and our experiments) is as follows: All subjects may choose among three options: buy a low-value product (e.g., one Hershey's

11

Kiss; hereafter "Hershey's"), buy a higher-value product (e.g., one Lindt truffle), or buy nothing. The variation across conditions that enables us to measure their reaction to the price of zero relies on two basic conditions: "cost" and "free." In the cost condition, the prices of both products are positive (e.g., Hershey's costs 1¢ and the Lindt truffle 14¢). In the free condition, both prices are reduced by the same amount, so that the cheaper good becomes free (e.g., Hershey's is free, and the Lindt truffle is 13¢).

We also consider how such constant price reductions might influence demand for these two products in a model in which zero is particularly attractive and a one in which zero is just another price so that we may better understand how this scenario might test whether the price of zero has some added attraction. According to a model in which the price of zero is particularly attractive, a price reduction from the cost condition to the free condition should create a boost in the attractiveness of the product that has become free and hence increase its relative demand. However, from the perspective of a model in which zero is just another price, because all changes in prices are the same, reducing one of the prices to zero should not create any unique advantage. In the next section, we examine these two models more formally and provide some testable predictions for distinguishing between them.

3. Formal Account of Standard Economic and Zero Price Models

We describe a "standard" model of how consumers behave in a situation in which they must choose between two products at certain prices (or buy nothing), as well as how their choices might change if both prices are reduced by the same amount. We then consider a special case of this situation in which the price decrease is equal to the original

smaller price; that is, the new smaller price is zero. Furthermore, we contrast this standard model with the *zero price model*, which is identical in all respects except that it assumes that when a product becomes free, its intrinsic value for consumers (or "benefit," in cost–benefit terminology) increases. After clarifying the different predictions of the two models regarding the observable behavior of consumers, we empirically test them in Section 4.

Consider a model with linear utilities in which a consumer must choose among three options X, Y, and N (we discuss the linearity assumption in detail subsequently). Option X refers to buying one unit of product X priced at P_X; option Y means buying one unit of product Y priced at P_Y; and option N means the consumer buys nothing. Suppose that the consumer values the first product at V_X and the second product at V_Y; he or she then will choose X if and only if

$$V_X > P_X \text{ and } V_X - P_X > V_Y - P_Y. \qquad (1)^2$$

The consumer will choose Y if and only if

$$V_Y > P_Y \text{ and } V_Y - P_Y > V_X - P_X. \qquad (2)$$

Finally the consumer will buy nothing (choose N) if and only if

$$V_X < P_X \text{ and } V_Y < P_Y. \qquad (3)$$

---

[2] Without loss of generality, we may assume that the probability that any of these or subsequent inequalities turns into an equality is zero.

Assume there are multiple consumers with $[V_X, V_Y]$ distributed over R^2; the three sets of inequalities determine three groups of consumers who choose each of the three options (see Figure 1a).

Now consider a situation in which both prices are reduced by the same amount ε. The new prices thus are equal to $[P_X - \varepsilon, P_Y - \varepsilon]$. How do the demand segments change? With the new prices, consumers who choose X are those with

$$V_X > P_X - \varepsilon \text{ and } V_X - P_X > V_Y - P_Y. \tag{1a}$$

Consumers choosing Y are those with

$$V_Y > P_Y - \varepsilon \text{ and } V_Y - P_Y > V_X - P_X. \tag{2a}$$

Finally, consumers choosing N are those with

$$V_X < P_X - \varepsilon \text{ and } V_Y < P_Y - \varepsilon. \tag{3a}$$

Comparing the two sets of formulas (or inspecting Figure 1b), we note that consumers who originally choose X keep choosing X, and consumers who originally choose Y keep choosing Y. Thus, according to this model, there should be no switching from one product to another. The only two possible changes in demand are that some consumers who originally buy nothing switch to either X (those with $V_X - P_X > V_Y - P_Y$ and $P_X - \varepsilon < V_X < P_X$) or Y (those with $V_Y - P_Y > V_X - P_X$ and $P_Y - \varepsilon < V_Y < P_Y$).

In short, according to this simple cost–benefit model, when prices decrease by the same amount, the costs decrease by the same magnitude for both products, whereas their benefits remain the same, and hence, the net benefits increase by the same amount. In turn, this model predicts that when the prices of both products drop by the same amount, both demands increase weakly (see Table 1).

Now consider a special case in which the price reduction, ε, equals the original smaller price, say P_X, so that the prices drop from $[P_X, P_Y]$ to $[0, P_Y - P_X]$. If zero is just another price, the preceding predictions remain valid. In our study setting, when prices decrease from the cost condition to the free condition, the proportion of consumers choosing each of the two products should increase weakly (see Figure 1c).

Next, consider the zero price model, which assumes that when a product becomes free, consumers attach a special value to it, that is, their intrinsic valuation of the good increases by, say, α. Note, the decision to add α to the benefit (intrinsic valuation) of the free good is rather arbitrary. All the predictions would go through just the same, if we assume that α is added directly to the net benefit of the free good or subtracted from its cost, or even added to the costs of all non-free goods (extra pain of paying). We will discuss he nature of α in more detail after the initial empirical findings are presented.

In this model, and in contrast with the standard model, some consumers switch from the more expensive good to the cheaper good if their valuations of the products satisfy the following set of inequalities. The first two inequalities imply the original choice of Y, and the second two inequalities lead to switching to X when its price is reduced to zero:

$$V_Y > P_Y,$$

$$V_Y - P_Y > V_X - P_X,$$

$$V_X + \alpha > 0, \text{ and}$$

$$V_X + \alpha - P_X > V_Y - P_Y. \qquad (4)$$

That is, as the prices fall from the cost condition to the free condition, the costs decrease by the same magnitude for both products, the benefit for the now free product increases more than that for the more expensive product, and the net benefit of the cheaper product becomes higher. In terms of demand, the zero price model predicts that as prices are reduced from the cost condition to the free condition, the demand for the cheaper good increases, and more importantly, the demand for the more expensive good may decrease as consumers switch from the more expensive product to the cheaper one (see Table 1, Figure 1d). We refer to the combination of the increase in the proportion of consumers choosing X and the decrease of those choosing Y when prices fall from $[P_X, P_Y]$ to $[0, P_Y - P_X]$ as *the zero price effect*. The prediction regarding the decrease in demand for the more expensive good represents the one observable difference between the two models, and thus, in our empirical section, we focus on it.

••• **Figure 1 & Table 1**•••

4. Testing the Phenomenon

In this section, we describe a series of experiments designed to test the validity of the zero price model and rule out some trivial economic explanations for the changes in demand that take place as the price of the cheaper good decreases to zero (i.e., from the cost condition to the free condition).

4.1. Experiment 1: Survey

Method. We asked 60 participants to make a hypothetical choice among a Hershey's, a Ferrero Rocher chocolate, and buying nothing (we provided pictures of both chocolates). Across the three conditions, the prices of the two chocolates decreased by a constant amount (for a description of all conditions across all the experiments, see the Appendix). In the cost condition, the prices of Hershey's and Ferrero were 1¢ and 26¢, respectively (1&26 condition). In the free condition, both prices were reduced by 1¢ and therefore were 0¢ and 25¢, respectively (0&25 condition). The third condition (2&27 condition) represents an additional cost condition in which the prices of goods increased by 1¢ above their prices in the first cost condition. The purpose of the 2&27 condition is to contrast the effect of a 1¢ price reduction that does not include a reduction to 0 (reduction from 2&27 to 1&26) with a 1¢ price reduction that does (reduction from 1&26 to 0&25).

Results and Discussion. We provide the results in Figure 2. As the prices decrease from the 1&26 condition to the 0&25 condition, the demand for Hershey's increases substantially ($t(31) = 3.8$, $p < 0.001$) while, more importantly, the demand for Ferrero decreases substantially ($t(31) = -2.3$, $p = 0.03$), in support of the zero price effect. The difference in demand between the 1&26 and 2&27 conditions is imperceptible

(Hershey's t(38) = -0.3, p = 0.76; Ferrero t(38) = 0, p = 1), which demonstrates that when all prices are positive, a 1¢ change in prices does not have a significant effect on demand. Only when one of the prices becomes zero does the observed perturbation take place. Thus, we observe (hypothetical) behavior consistent with the zero price model; participants reacted to the free Hershey's as if it had additional value.

••• **Figure 2** •••

4.2. Experiment 2: Real Purchases

Although the results of Experiment 1 suggest that consumers react to a price decease to zero differently than they do to other price reductions, their reaction pertains to a hypothetical situation, which means that it remains an open question whether consumers will behave in the same way when faced with real transactions. As a secondary goal, Experiment 2 includes another condition to test the robustness of the zero price effect. In this condition, the price reduction is much larger for the high-end candy, which gives participants a greater incentive to make choices opposite to the predictions of the zero price effect. Furthermore, this unequal price reduction provides a test of the notion that consumers divide, rather than subtract, costs and benefits (as we discuss subsequently).

Method. Three hundred ninety-eight subjects took part in the experiment. We use a Hershey's as the low-value product and a Lindt truffle (hereafter, "Lindt") as the high-value product. The experiment includes a free condition (0&14), a cost condition (1&15), and a second free condition (0&10). In the 0&14 and 0&10 conditions, the price of

Hershey's is 0¢, and the price of Lindt is 14¢ and 10¢, respectively. In the 1&15 condition, the price of Hershey's is 1¢, and the price of Lindt is 15¢.

A booth in MIT's student center contained two cardboard boxes full of chocolates and a large upright sign that read "one chocolate per person." Next to each box of chocolates was a sign lying flat on the table that indicated the price of the chocolate in that condition. The flat signs could not be read from a distance, and the prices were visible only to those standing close to the booth. We use the flat signs because we want to measure the demand distributions, including the number of people who considered the offer and decided not to partake. By placing the price signs flat next to the chocolates, we could code each person who looked at the prices but did not stop or purchase and classify them as "nothing."

Although field experiments have many advantages, this particular setup suffers a limitation in that the experimental conditions could not be randomized for each subject; instead we alternated the price signs (conditions) approximately every 45 minutes. When replacing the signs, we wanted to reduce the chance that students would notice the change (which would mix within- and between-subjects designs) and therefore instituted 15-minute breaks between each of the 30-minute experimental sessions.

Results and Discussion. As we show in Figure 3, the results are similar to the hypothetical choices in Experiment 1. As the prices decrease from the 1&15 condition to the 0&14 condition, demand for Hershey's increases substantially ($t(263) = 5.6$, $p < 0.001$), while demand for Lindt decreases substantially ($t(238) = -3.2$, $p < 0.01$). In addition, we find no significant difference between the demand for Hershey's between the 0&14 and 0&10 conditions ($t(263) = 0.5$, $p = 0.64$) and a marginally significant

difference in demand for the Lindt between the 0&14 and 0&10 conditions (t(271) = 1.5, p = 0.13). This marginal difference, however, is in the opposite direction of the expected effect of a price decrease on demand, which may be related to the higher number of participants who took nothing in the 0&10 condition. Together, these results show that the reduction of a price to zero is more powerful than a five-times larger price reduction that remains within the range of positive prices.

A somewhat surprisingly large proportion of people selected "nothing." This observed lack of interest could be due to the way we coded the choice of nothing; some people who might not even have noticed the offers (and thus effectively were not part of the experiment) could have been misclassified as buying nothing (instead of being considered nonparticipants). Another possible contributor to the choice of nothing could be transaction costs; buying a chocolate or even taking a free chocolate requires attention and time. Finally, in the experimental setting, the value of chocolate may have been either not positive or not sufficiently large for our participants.

If we take those whom we coded as nothing out of the analysis, the share of Hershey's increases from 27% in the 1&15 condition to 69% in the 0&14 condition and to 64% in the 0&10 condition. The demand for Lindt shows a complementary pattern: decreasing from 73% in the 1&15 condition to 31% in the 0&14 condition and 36% in the 0&10 condition. The difference between the cost and the free conditions is statistically significant (both ps < 0.001), but the difference between the two free conditions is insignificant (t(142) = -1.0, p = 0.31).

In summary, the results of Experiment 2 demonstrate that valuations of free goods increase beyond their cost–benefit differences, as we show with real transactions in a

field setting, and even when the price decrease for the high-value product is substantially larger than that of the low-value product. The observed drop in demand for the high-value good in such a case (from the 1&15 condition to the 0&10 condition) is theoretically even more impossible than in the case when prices decrease by the same amount.

Another advantage of the comparison of the 0&10, 0&14, and 1&15 conditions is that it sheds some light on the possibility that rather than evaluating options on the basis of their cost–benefit difference, consumers might consider goods on the basis of the ratio of benefits to costs (not a normative account). According to this interpretation, the net value of a free good is very high (strictly speaking, infinite) and therefore leads to the choice of the free good. However, the results of Experiment 2 weaken the possibility of this explanation in two ways. First, if our participants followed a strict ratio rule, and if we assume that everyone has at least an epsilon valuation for Hershey's, the choice share of the free chocolate should have been 100%, or at least 100% of those selecting any chocolate, which is not the case. Second, a less strict version of the ratio rule implies that the price reduction of the high-end chocolate from 15¢ to 10¢ (a 33% reduction) should have had a much larger effect on its share compared with the price reduction from 15¢ to 14¢ (a 7% reduction). This prediction does not bear out; there is no real difference in the changes in demand when the prices fall from 1&15 to 0&14 on the one hand and to 0&10 on the other hand.

••• **Figure 3** •••

4.3. Experiment 3: Cafeteria.

We acknowledge a possible shortcoming of Experiment 2; namely, the difference between conditions may not be confined to prices, such that the size of the transaction costs associated with the three options differs among conditions. Taking a free Hershey's or buying nothing means not only a zero monetary price but also no associate hassle of looking for change in a pocket or backpack. If transaction cost is a consideration in our setting, it could lead to a choice pattern that favors Hershey's when its cost is zero (in the 0&14 and 0&10 conditions), but not when both options involve a positive cost and hence a larger transaction cost (the 1&15 condition). We derive an initial indication that transaction cost is not the driver of the effect from the results pertaining to the hypothetical choices in Experiment 1. Because Experiment 1 does not involve real transactions, it does not involve any transaction costs, which implies the results will survive a situation without transaction costs. However, though these results are indicative, when respondents made their hypothetical choices, they might have considered transaction costs that would have been present if the choice they were facing had been real. Because the results of Experiment 1 cannot be interpreted conclusively and because transaction costs could be an important alternative explanation, we conduct Experiment 3, designed explicitly to control for possible differences in transaction costs. In this experiment, we hold the physical transaction costs constant for the three choices (high- and low-value chocolates and no purchase) and between the cost and free conditions.

Method. We carried out this experiment as part of a regular promotion at one of MIT's cafeterias, using customers who were already buying products at the cafeteria and

adding the cost of the chocolate to their bill as if it were any other purchase. By adding the cost to an existing purchase, we create a situation in which the chocolate purchase does not add anything to the transaction costs in terms of taking out one's wallet, looking for money, paying, and so forth.

The procedure of the experiment is generally similar to that used in Experiment 2: a box with two compartments, one containing Hershey's and the other containing Lindt, appeared next to the cashier. A large sign read "one chocolate per person," and we posted the price of each chocolate next to each compartment (varying across conditions). Customers who wanted one of the chocolates had its cost added to their bill. Thus, the transaction costs in terms of payment remained the same whether a customer purchased a chocolate, got a chocolate for free, or purchased nothing, because he or she still had to pay for the main purchase.

We manipulated the prices at two levels: 1¢ for Hershey's and 14¢ for Lindt in the cost condition, and 0 and 13¢, respectively, in the free condition. We switched the price signs (conditions) approximately every 40 minutes, with a 10-minute break between the experimental sessions. In this setting, it was difficult to separate customers who decided not to participate from those who did not notice the offer; therefore, all customers who passed by the cashier and did not select any of our chocolates were coded as "nothing." In total, 232 customers took part in this experiment.

Results and Discussion. As we show in Figure 4, in the condition in which Hershey's is free, the demand for Hershey's increases substantially ($t(189) = 4.7$, $p < 0.001$), while the demand for Lindt decreases substantially ($t(206) = -3.2$, $p = 0.001$). If we remove those whom we code as nothing from the analysis, the share of Hershey's

increases from 21% in the 1&14 condition to 71% in the 0&13 condition, whereas the share of Lindt decreases from 79% in the 1&14 condition to 29% in the 0&13 condition $(t(92) = 5.6, p < 0.0001)$.

Thus, the zero price effect is not eliminated when transaction costs are the same for all options and in both conditions, which provides strong evidence that the zero price effect is not produced solely by a difference in transaction costs.

••• **Figure 4** •••

4.4. Summary of the Initial Experiments

These initial experiments contrast the choices respondents make when the prices for both options are positive relative to a case in which both options are discounted by the same amount, such that the cheaper option becomes free. This methodology enables us to examine the reaction to free offers and indicates both an increase in demand for the cheaper product and a decrease in demand for the more expensive product, an effect we term the zero price effect.

Experiment 1 demonstrates that a 1¢ difference in price has an enormous influence on demands if it represents a difference between a positive and zero prices but not when it is a difference between two positive prices. Participants reacted as if a free Hershey's had more intrinsic value than a positively priced Hershey's. Experiment 2 validates this finding with real choices and argues against the ratio explanation. Finally, Experiment 3 demonstrates that the zero price effect is not driven by transaction costs.

Thus, we show that for prices, as for many other domains, zero is treated qualitatively differently from other numbers.

When we consider how zero might differ from other numbers, we posit two general answers: The first relies on the proposed model and assumes a unique benefit of the price of zero, which leads to a demand discontinuity at zero. A second approach is to model this process with a concave utility of money. In such a model, instead of evaluating options by V - P (i.e., value minus price), consumers evaluate them by V - v(P), where v is the prospect theory value function (Kahneman and Tversky 1979). To illustrate this point, consider the choices from Experiment 3: If the net benefit of a chocolate is defined by V - v(P), participants could switch from Lindt to Hershey's because $v(14¢) - v(13¢) < v(1¢)$. The utility of money is likely to be generally concave (Kahenman and Tversky 1979), so the question for our purposes is not whether it is concave but rather whether concavity may account for our findings. Moreover, the discontinuity in zero that we propose represents a special case of concavity; a function that is zero at zero and then "jumps" and is upward sloping and linear (or concave) is by definition concave. Our question therefore pertains to whether the effect of the price of zero is captured better by a continuous or discontinuous concave utility of money.

To examine the possibility that continuous concavity could be sufficient to account for the results, we consider the contrast between the two price reductions in Experiment 1: from 2&27 to 1&26 and from 1&26 to 0&25. A model claiming that a continuous concave utility function of money can account for the results would assume that consumers evaluate the options by V - v(P), that $v(26¢) - v(25¢) < v(1¢)$, and that this difference is sufficient to explain the large zero price effect documented in Experiment 1.

However, this model would have to assume also that $v(27\cent) - v(26\cent) < v(2\cent) - v(1\cent)$, and thus, we should expect an increase in demand for Hershey's and a decrease in demand for Ferrero in the 1&26 versus the 2&27 condition. Such demand changes should be smaller in magnitude than those between 1&26 and 0&25, but they would occur in the same direction. However, as we show in Figure 2, the results do not indicate anything of the kind. Although concavity is present in the utility of money, the type of concavity in our setting is more likely to exist because of a discontinuity at zero rather than continuous concavity alone (we provide further support for the discontinuous nature of the zero price in the Amazon gift certificates experiment and a flat screen televisions experiment described later).

5. Why Is Zero Price Special?

In the first part of this article, we demonstrate that zero price has a special role in consumers' cost–benefit analysis. In this section, we take another step toward exploring the psychology behind the zero price effect. In particular, we consider three possible explanations, which we label "social norms," "mapping difficulty," and "affect." On the basis of prior research and an additional study, we argue that the social norms explanation, though applicable in some cases, cannot account fully for the zero price effect, so we focus on distinguishing between the mapping difficulty and affect accounts. Overall, the results support the role of affect as a main cause for the effect of zero.

5.1. Social Norms

A possible psychological mechanism that could underlie the zero price effect deals with the norms that might accompany free products. Costly options invoke market exchange norms, whereas free products invoke norms of social exchange (Fiske 1992, McGraw, Tetlock, and Kristel 2003, McGraw and Tetlock 2005). Thus, evoked social norms may create higher value for the product in question. Heyman and Ariely (2004) offer one example in which they demonstrate that people are likely to exert higher effort under a social contract (no monetary amounts) than when small or medium monetary amounts are mentioned. Another example of the relationship between social and exchange norms appears in Ariely, Gneezy, and Haruvy's (2006) research, in which they examine the behavior of persons faced with a large box of candies and an offer to receive the candy either for free or for a nominal price (1¢ or 5¢). Not surprisingly, when the cost is zero, many more students take candy than when the price is positive. More interesting, when the price is zero, the majority of the students take one and only one candy, while those who pay to take candy take a much larger amount (effectively creating lower demand as prices decrease).

Together, these results suggest that social norms are more likely to emerge when price is not a part of the exchange, which could increase the valuation for a good and, in our experiments, increase the market share of the free chocolate. However, another condition in Heyman and Ariely's (2004) experiments suggests that the effect of social norms might not apply to our settings. When the elements of both social exchanges (e.g., a gift) and monetary exchanges occur (e.g., "Here is a 50¢ candy bar"), the results are very similar to those of a monetary exchange and different from those of a social exchange. Relating these findings to our setting suggests that it is highly unlikely

participants apply social exchange norms to one option in the choice set (free option) and monetary exchange norms to the other (cost option). Instead, participants probably apply the same set of norms to all choices in the set and thereby eliminate the effect of social exchange norms.

To test the ability of social exchange norms to account for the zero price effect further, we create an additional condition that enables us to disassociate the free cost from the social norms invoked by the lack of cost. That is, we offer the low-value chocolate for a small negative price (-1¢), which creates a transaction with no downside (no financial cost) but still mentions money and thus presumably does not invoke social exchange norms. To the extent that the zero price effect is due to the social nature of nonmonetary exchanges, a negative price, which has no social aspect, should not induce an increase in the intrinsic valuation of the products in the same way zero price does. However, if the zero price effect is not due to social exchange norms, demand in this condition should be very similar to that in the free condition.

Three hundred forty-two subjects took part in this experiment, which replicates the 1&14 and 0&13 conditions of Experiment 2 with the addition of a -1&12 condition, in which the price of Hershey's is -1¢ (participants received Hershey's plus a penny) and the price of Lindt is 12¢. The demands in the 1&14 and 0&13 conditions replicate our previous findings: Compared with the 1&14 condition, the demand for Hershey's in the 0&13 condition increases substantially from 15% to 34% ($t(193) = 3.4$, $p < 0.001$), and the demand for Lindt decreases substantially from 38% to 16% ($t(212) = -3.8$, $p < 0.001$). Of greater significance, we find that when prices drop from 0&13 to -1&12, the demand for Lindt remains 16% ($t(220) = 0.04$, $p = 0.97$), but the demand for Hershey's increases

from 34% to 50% (t(212) = -3.8, p < 0.001). Thus, in contrast with the social exchange norms explanation, the zero price effect remains even when we mention money for both options in the choice set. These results also suggest that a change in the cost–benefit analysis likely causes the shift in evaluations for the free (or small negative cost) product.

5.2. Mapping Difficulty

A second possible psychological mechanism that might explain the overemphasis on free options comes from the findings of Ariely, Loewenstein, and Prelec (2003, 2006), Hsee et al. (2003), and Nunes and Park (2003), which demonstrate that people have difficulty mapping the utility they expect to receive from hedonic consumption into monetary terms. In one set of studies that illustrates this mapping difficulty, Ariely, Lowenstein, and Prelec (2003) demonstrate that maximum willingness to pay (elicited by an incentive-compatible procedure) is susceptible to anchoring with an obviously irrelevant number—the last two digits of a social security number (Tversky and Kahneman 1974; Chapman and Johnson 1999). For example, students whose last two digits of their social security numbers were in the bottom 20% of a distribution priced a bottle of 1998 Cotes du Rhone wine at $8.64 on average, whereas those whose last two digits were in the top 20% priced the same bottle at $27.91 (see also Simonsohn and Lowenstein 2006). These results suggest it is difficult for decision makers to use their internal evaluations for products, so they resort to the use of external cues to come up with their valuations.

Mapping difficulty could play a role in our setting as well. To the extent that evaluating the utility of a piece of chocolate in monetary terms is difficult, consumers

might resort to a strategy that assures them of at least some positive surplus. Specifically, receiving a piece of the lower-value chocolate for free must involve positive net gain, but paying for a piece of the higher-value chocolate may or may not. To illustrate, imagine a situation in which a consumer's valuation for the lower-value chocolate is somewhere between 1¢ and 5¢ and his or her valuation for the higher-value chocolate is between 10¢ and 20¢. If this consumer were faced with the 1&14 condition, it would be unclear which of the options would give him or her a net benefit or the higher net benefit. However, the same consumer facing a 0&13 condition easily recognizes that the free option definitely provides a net benefit, so the consumer chooses that option. Thus, the zero price effect might be attributed, according to this perspective, to the uncertainty surrounding the overall benefit associated with costly options and the contrasting certainty about overall benefits associated with free options.

5.3. Affect

A third possible psychological mechanism that might account for the zero price effect pertains to affect, such that options with no downside (no cost) invoke a more positive affective response; to the extent that consumers use this affective reaction as a decision-making cue, they opt for the free option (Finucane et al. 2000, Slovic et al. 2002a, Gourville, and Soman 2005). We test this prediction directly with Experiment 5. The affective perspective also suggests the circumstances in which the zero price effect should be eliminated: If the cause of the zero price effect is a reliance on an initial (overly positive) affective evaluation, making a non-affective, more cognitive evaluation accessible might diminish the zero price effect.

To test which of these two psychological mechanisms (mapping difficulty, affect) is the more likely driver of the zero price effect, we conduct three more experiments. In Experiment 4, we attempt to reduce or eliminate the mapping difficulty to observe whether that diminishes or eliminates the zero price effect. In Experiment 5, we test the first proposition of the affective account, namely, that free offers elicit higher positive affect. In Experiment 6, we test whether forcing people to evaluate the options cognitively, and thereby making these evaluations available and accessible, eliminates the zero price effect.

5.4. Experiment 4: Halloween

Experiment 4 aims to test whether mapping difficulty could be driving the zero price effect. Therefore, we reduce mapping difficulty by making both sides of the transactions (i.e., that which participants stand to gain and that which they relinquish) commensurable. We predict that to the extent that mapping difficulty is the cause of the zero price effect, it will diminish when the two sides of the transaction match. We also predict that this type of manipulation will have no bearing if affect is the cause of the zero price effect.

Method. To reduce mapping difficulty, participants were able to exchange chocolate for chocolate rather than for money. Specifically, on Halloween, 34 trick-or-treaters at an authors' house were exposed to a new Halloween tradition. As soon as the children knocked on the door, they received three Hershey's (each weighing about 0.16 oz.) and were asked to hold the Hershey's they had just received in their open hand in front of them. Next, each child was offered a choice between a small (1 oz.) and a large

31

(2 oz.) Snickers bar. In the free (0&1) condition, they could simply get the small Snickers bar or exchange one of their Hershey's for the large Snickers bar. In the cost (1&2) condition, the children could exchange one of their Hershey's for the small Snickers bar or exchange two for the large Snickers bar. They also could choose not to make any exchanges.

Results and Discussion. As we show in Figure 5, the zero price effect remains strong even when the trade-offs involve commensurate products and exchange media ("money"). In the 0&1 condition, in which the small Snickers bar is free, demand for it increases substantially (relative to the cost condition), whereas demand for the large Snickers bar decreases substantially (t(31) = 4.9, p < 0.001). A follow-up experiment with adults, conducted at the MIT Student Center in a setting similar to Experiment 2, includes the 0&4 and 1&5 conditions for exchanges involving Hershey's for small and large Snickers, respectively. The results replicate the pattern of results of the Halloween experiment.

These results generalize our previous findings in five ways. First, they demonstrate that the attractiveness of zero cost is not limited to monetary transactions; there seems to be a general increase in attractiveness of those options that do not require giving up anything. Second, the results hold when the goods and exchange currency are commensurate—in this case, chocolate-based candy (for other results regarding commensurability, see Ariely, Loewenstein, and Prelec 2003; Hsee et al., 2003; Nunes and Park, 2003). Third, though a 1¢ price is not very common in the marketplace, the choice and trading of candy is more common (particularly in the context of Halloween), which adds ecological validity to our finding. Fourth, the results provide further support

that the physical hassle involved in transactions cannot account for the results. Fifth, this effect holds for adults as well as for children.

••• Figure 5 •••

As a further test of the mapping account for the effect of zero prices, we conduct another experiment in which both the products and the method of payment were money. The two products participants could choose from were $10 and $20 Amazon gift certificates (or "neither"). The prices for the gift certificates were varied at three levels: $5 and $12, $1 and $8, and $0 and $7, respectively, with the $20 certificate always costing $7 more than the $10 certificate. As the reader may guess, we find no differences in demand patterns between the 5&12 and the 1&8 conditions ($t(65) = 0.53$, $p = 0.6$), but demand for the $10 certificate rockets in the 0&7 condition ($t(65)= 6.9$, $p < 0.001$) while demand for the $20 certificate falls to zero (see Figure 6). Thus, the experiment further invalidated mapping difficulty as a source of the zero price effect; the effect survived a situation in which the product sold and the medium were both monetary.

This lack of difference in demand between the 5&12 and 1&8 conditions, together with the large shift in demand in the 0&7 condition, also argues against a ratio account. The ratios of the costs are much more favorable toward the $10 Amazon gift certificate in the 1&8 condition compared with the 5&12 condition (by approximately 3.3 times), so if participants actually used the ratio rule, we would have observed a large increase in demand for the $10 Amazon gift certificate in the 1&8 condition, which we did not.

The availability of multiple conditions with both positive prices in this experiment also helps us examine whether gradual price reduction to zero creates a continuous or discontinuous changes in demand and hence whether $v(P)$ is continuous at zero. Continuous change would most likely, result in at least a slight difference between the 5&12 and 1&8 conditions, and a (potentially larger) difference between the 1&8 and 0&7 conditions. The observed lack of the former difference adds suggests that discontinuity of $v(P)$ at zero might be a better account for our data.

••• **Figure 6** •••

In summary, the main reason for our Halloween and Amazon gift certificate experiments was to test whether the difficulty of mapping money onto experiences could be the cause of the zero price effect. We first replaced money as the exchange medium with chocolates, which presumably can be mapped more naturally onto other chocolates. We then replaced the product and the exchange medium with money. The results demonstrate that the zero price effect is not limited to goods-for-money exchanges and that it is unlikely to be explained fully by mapping difficulties.

5.6. Experiment 5: Smilies

The affect account has two basic components. The first is that free offers evoke higher positive affect, and the second is that people use this affect as an input for their decision-making process. In Experiment 5, we examine the first component: People experience more positive affect when facing a free offer compared with other offers.

Method. We asked 243 participants to evaluate how attractive they found an offer of a chocolate at a certain price. We manipulated the offer on four levels among participants: Hershey's for free (H0), Hershey's for 1¢ (H1), Lindt for 13¢ (L13), and Lindt for 14¢ (L14). Participants received a questionnaire with the details of the offer and a picture of the chocolate. At the bottom of the page, schematic pictures of five faces ("smilies") with different expressions appeared, varying from unhappy to very happy. Participants were asked to indicate their feelings toward the offer by circling one of the faces. If participants' attitude toward the offers reflected the offers' net benefits, the attitudes toward L14 and H1 should be slightly lower than those toward L13 and H0, respectively; and the difference between the attitudes toward L13 and L14 should be similar to the difference between H0 and H1. The affect argument, however, suggests that the attitude toward H0 should be much higher than that toward any other offer.

Results and Discussion. We depict the results in Figure 7. In line with the affect hypothesis, attitude toward the H0 offer is significantly higher than attitude toward any other offer (t(113) = 7.0, p < 0.001). Furthermore, we find no difference among the attitudes toward the other three offers (F(2, 178) = 0.35, p = 0.7). In support of the affect idea, the free good elicits more positive affect than standard cost–benefit analysis predicts.

Why does a free Hershey's elicit such higher positive affect relative to a 13¢ Lindt? Ex ante, it is possible that a Lindt at 13¢ provides a much better deal than a Hershey's at any price. In fact, when people carefully consider the pros and cons of these offers, they much more often come to conclusion that the value of 13¢ Lindt is higher than that of a free Hershey's (see Experiment 6). But, as the results of Experiment 5

demonstrate, it is also clear that free Hershey's creates much higher affective reaction. One reason for this could be that that the decision to take a chocolate for free is a much simpler decision, and that simplicity could be the driver of higher affect (Tversky and Shafir 1992, Luce 1998, Iyengar and Lepper 2000, Benartzi and Thaler 2002, Schwarz 2002, Diederich 2003, Gourville and Soman 2005). In particular, a free Hershey's involves benefits and no costs, while a Lindt for any positive price involves both benefits and costs – it is possible that options that have only benefits create more positive affect compared with options that involve both benefits and cots. Alternatively, much like the disutility of paying while consuming (paying for a vacation while experiencing it: Prelec and Loewenstein 1998), it is possible that options that involve both benefits and cots create a negative impact on affect due to the simultaneity of these two components, while options that have only benefits do not include this "penalty."

••• **Figure 7** •••

5.7. Experiment 6: Forced Analysis

In response to the high affective reaction to the free option in Experiment 5, we test whether consumers use this increased affect as a cue for their decisions, which in turn causes the zero price effect. In Experiment 6, we force participants to engage in a cognitive and deliberate evaluation of the alternatives before they choose and thereby make nonaffective, more cognitive evaluations available and accessible to participants. We assume that in these conditions, participants are more likely to base their evaluations on cognitively available inputs and therefore place a lower weight on the affective

36

evaluations. To the extent that the cause of the zero price effect is the affective component, such reliance on cognitive inputs should reduce the zero price effect.

Method. Two hundred students filled out a survey in which they made a hypothetical choice among three options. We also asked half the subjects to answer two questions before making the choice. The design was a 2 (chocolates' prices: 1&14 vs. 0&13) × 2 (survey type: neutral vs. forced analysis) between-subjects design.

The survey in the [1&14, neutral] condition asked participants to imagine that there is a chocolate promotion at the checkout counter of their supermarket and that they could either buy one Hershey's kiss for 1¢ or one Lindt truffle for 14¢. Participants indicated their preferred option (a Hershey's for 1¢, a Lindt for 14¢, or neither). The [0&13, neutral] condition mirrored the 1&14 condition, except that Hershey's and Lindt were offered for free and 13¢, respectively.

In the forced analysis conditions, after reading the introduction but before being asked for their hypothetical choice, participants were asked the following two questions: "On a scale from 1 (not at all) to 7 (much more) how much more do you like the Lindt truffles in comparison with Hershey's kisses?" and "On a scale from 1 (not at all) to 7 (much more) how much more would you hate paying 14¢ (13¢) in comparison with paying 1¢ (nothing)?" Participants circled a number from 1 to 7, anchored at 1 (not at all), 4 (about the same), and 7 (much more). After answering these questions, participants made their hypothetical choice among the three options.

Results and Discussion. We ran two logit regressions with the proportions of subjects buying Hershey's and Lindt as the dependent variables and the answers to the two questions as independent variables (forced analysis conditions only). Unsurprisingly,

preferring Lindt to Hershey's is related negatively to choosing Hershey's ($z = 3.1$, $p < 0.01$) and positively to choosing Lindt ($z = 3.0$, $p < 0.01$). Disliking paying more is related positively to choosing Hershey's ($z = 3.2$, $p = 0.001$) and negatively to choosing Lindt ($z = 3.1$, $p < 0.01$). Thus, participants' answers to the questions fall in line with their choices.

Next, we performed two ANOVAs with the proportions of subjects choosing Hershey's and Lindt as the dependent measures and the chocolates' prices, survey type, and the interaction term as independent variables. The ANOVAs reveal significant main effects of chocolates' prices (Hershey's $F(1, 196) = 9.7$, $p < 0.01$; Lindt $F(1, 196) = 8.7$, $p < 0.01$), no main effects of survey type (Hershey's $F(1, 196) = 2.0$, $p = 0.2$; Lindt $F(1, 196) = 1.6$, $p = 0.2$), and, most importantly, a significant interaction effect for the two factors (Hershey's $F(1, 196) = 4.5$, $p = 0.03$; Lindt $F(1, 196) = 5.1$, $p = 0.02$).

As we demonstrate in Figure 8, the zero price effect is replicated in the neutral conditions (Hershey's $t(97) = 3.7$, $p < 0.001$; Lindt $t(97) = -3.7$, $p < 0.001$) but not in the conditions in which subjects compare their quality and price options before choosing. In the forced analysis conditions, the direction of the effect remains the same, but the magnitude is much smaller and statistically insignificant (Hershey's $t(99) = 0.7$, $p = 0.5$; Lindt $t(99) = -0.6$, $p = 0.6$). These results support the basic affect mechanism we propose, according to which the affect invoked by the free option drives the zero price effect, but when people have access to available cognitive inputs, they base their decisions on those, and the benefit of zero largely dissipates.

Another potential interpretation of these results is that in three of our four conditions, subjects act "rationally"—the two forced analysis conditions and the [1&14,

neutral] condition. In the [0&13, neutral] condition, however, they act on the basis of the affect evoked by the zero price. In support of this idea, we find no significant difference among subjects' choices in the three rational conditions (Hershey's $F(2, 147) = 0.7$, $p = 0.5$; Lindt $F(2, 147) = 0.7$, $p = 0.5$), whereas the [0&13, neutral] condition differs significantly from them (Hershey's $t(83) = 3.8$, $p < 0.001$; Lindt $t(83) = 0.3.7$, $p < 0.001$).

••• **Figure 8** •••

6. General Discussion

We start with two models, one that treats zero as just another price and one that assumes free options are evaluated more positively. We propose a method to distinguish these two approaches and demonstrate in three experiments that the latter model is better able to account for our findings. Experiment 1 provides the initial evidence of the zero price model, and Experiment 2 supports the effect with a real buying scenario and clarifies that the effect could not be due to decision making based on cost–benefit ratios. Experiment 3 shows that the effect also could not be due to physical transaction costs.

After demonstrating the unique properties of zero price, we attempt to examine the psychological causes for this effect and propose three possible mechanisms: social norms, mapping difficulty, and affect. We discard the social norms explanation on the basis of findings (Heyman and Ariely 2004) that the mention of price invokes market-based transaction norms, which makes it unlikely that our scenario invokes social norms. We further discredit the ability of this account to explain our findings using negative prices that involve no cost but invoke prices. We then carry out three experiments to

explore which of the other two possible explanations is valid. Experiment 4 weights in against the difficulty of mapping explanation, and Experiments 5 and 6 provide support for the affective evaluation hypothesis.

In general, this research joins a larger collection of evidence that shows zero is a unique number, reward, price, and probability. Although our results suggest that the zero price effect might be accounted for better by affective evaluations than by social norms or mapping difficulty, zero and the price of zero remain a complex and rich domain, and all of these forces may come into play in different situations. In addition, other effects of zero might include inferences about quality, changes in signaling to the self and others, an effect on barriers for trial, and its ability to create habits. Therefore, much additional work is needed to understand the complexities of zero prices in the marketplace.

6.1. Alternative Explanations and Boundary Conditions

One of the limitations of our experimental conditions is that they are restricted to relatively cheap products and relatively unimportant decisions. Given this limitation, it remains an open question whether the zero price effect occurs when the decisions involve larger sums of money and more important decisions. To answer this question, at least partially, we distributed a survey in which participants responded to one of four hypothetical scenarios regarding purchasing an LCD flat-panel television. In these scenarios, participants were entitled to a large discount and had narrowed down their options to two: a cheaper 17" Philips and a more expensive 32" Sharp. The four conditions varied in terms of prices, such that the Sharp was always $599 more expensive than the Philips, and the prices of both sets decreased by approximately $100 across

conditions. From most expensive to least expensive, the conditions were 299&898, 199&798, 99&698, and 0&598. Comparing demand across these conditions, we find that the results (n = 120) generally resemble our previous findings. Demand for the smaller, cheaper television is 40% in the 299&898 condition, 40% in the 199&798 condition, 43% in the 99&698 condition, and 83% in the 0&698 condition. Concurrently, demand for the larger, more expensive television is 40% in the 299&898 condition, 33% in the 199&798 condition, 43% in the 99&698 condition, and 17% in the 0&698 condition. Overall, these results show that a shift in demand is apparent only when the price is reduced to zero ($F(3,98) = 3.24$, $p < 0.05$); otherwise, the effects of price reductions do not have a significant influence on the relative demand for the two televisions ($F(2,69) = 0.06$, $p = 0.94$), providing additional evidence against the continuous concavity argument.

Although these results suggest that the effect of the price of zero is not limited to small prices and meaningless decisions, some thought experiments also imply it might not be as simple with large, consequential decisions. For example, if we replace Hershey's and Lindt with Honda and Audi and change the prices from $28,000 and $20,000 to either $8,100 and $100 or $8,000 and $0, respectively, we suspect that relatively small prices such as $100 might be perceived within a just noticeable difference zone of zero, such that the effect of zero might be stretched to accommodate this price. Thus, the question of which prices people perceive as zero might not be simple, because it likely relates to the context of the decision and the original prices.

Another possible limitation of our setup is that our positive prices could seem suspicious. People in general are not accustomed to prices of 1¢, 13¢, or 14¢, whereas free samples often are a part of a promotion, which would make people more accustomed

to them. We selected such odd prices because we wanted to have a very small discount (1¢), while avoiding alternative accounts related to accumulation and disposal of small change across the different conditions (assuming that people are aversed to having many small coins fill their pockets). At the same time, these odd prices could have evoked suspicion, and our participants might have been making negative quality inferences about the cheap chocolates (the ones with odd prices) but not about the free chocolates. Three of the experiments cast doubt on this type of argument: In the Amazon gift certificates experiment the perceived quality of the gift certificates was unlikely to be influenced by price; in the Halloween experiment, all trade-offs were equally strange; and in the televisions experiment we gave an explicit explanation for the strange prices: "Luckily for you, you won a lottery that the store had conducted for its best customers. As a result, you are entitled to a huge discount on any product in the store."

To test this "negative inference from odd prices" alternative account more directly, we conducted two additional experiments. In one experiment we asked participants to make hypothetical choice among Hershey's, Lindt, and nothing but this time used prices that were less suspicious (0&15 and 10&25). The results replicate our previous findings, with demand for Hershey's increasing from 8% in the 10&25 condition to 65% in the 0&15 condition ($t(51) = 6.0$, $p < 0.0001$) and demand for Lindt decreasing from 45% in the 10&25 condition to 6% in the 0&15 condition ($t(54) = 3.8$, $p < 0.001$). In the second experiment we described in detail the setup of the Cafeteria Experiment (Experiment 3), and measured the inferences participants made about the products. Half of the participants read the description of the 0&13 condition, and the other half read the description of the 1&14 condition. After reading and viewing the

verbal and graphical descriptions, the participants are asked to describe their reaction to the promotion in an open-ended manner, followed by seven questions in which they are asked to rate the promotion on oddity and the chocolates on perceived quality, taste, and expiration date (relative to the same brand chocolates from a supermarket). The written protocols reveal that though participants mention that the promotion is odd (in particular, because of the "One chocolate per person" sign), or that the prices are odd; none of the participant spontaneously mentions the quality of the chocolates or makes any price-quality inferences. In addition, the rating in the seven questions reveal no differences in promotion oddity or inferences about chocolate quality (or taste, or expiration date) between the conditions. In general, even though the promotion is seen as somewhat odd by the participants, they do not make any differential inferences for the condition with low positive prices vs. the zero price condition.

Even though the zero price effect does not appear to be driven by the oddities of the prices we used, we do not assume that the price of zero effect will never interact with processes relating to consumers' inferences about quality. In many market situations, consumers might infer the expected quality of the product on the basis of such small prices, the price of zero itself, or the availability of free giveaway promotions (Simonson, Carmon, and O'Curry 1994).

Finally, the asymmetric dominance effect could offer another possible explanation for our findings (Huber, Payne, and Puto 1982). In our free conditions, the cheaper product always weakly dominates the buying nothing alternative, because they share the same cost (zero) and clearly differ in their benefits. In the cost conditions, no such asymmetric dominance relationship exists. If the zero price effect in our experiments is

driven by the asymmetric dominance effect, the relationship between the option to buy nothing and the cheaper chocolate (whether dominant or not) serves as the basic cause for the effect. Moreover, if we exclude the option not to buy anything, the asymmetric dominance relationship no longer exists, and any effect due to it should be eliminated. To test this asymmetric dominance explanation, we conducted a survey (n = 136) in which we excluded the buy-nothing option (which we could only do in a hypothetical choice study) and contrasted the zero price effect with the case in which participants had the buy-nothing option. The results replicate our standard findings: Free Hershey's experiences a demand boost (from 28% to 92%) while Lindt suffers a demand decrease (from 72% to 8%, t(50) = 6.8, p <0.0001), even in the absence of a dominated alternative. Moreover, these changes in demand are basically identical to the case in which the option to select nothing appears. Although the asymmetric dominance therefore is an unlikely explanation for our findings, there are other context effects ranging from product assortments to reference points in online auctions (e.g. Dholakia, and Simonson 2005, Leclerc, Hsee, and Nunes 2005) that could relate to these findings. Thus, we note that the more general questions of what context effects might be involved and influence prices of zero remain open and interesting.

6.2. Managerial Implications

The most straightforward managerial implication of our findings pertains to the increased valuations for options priced at zero. When considering promotions at a low price, companies should experiment with further discounts to zero, which likely will have a surprisingly larger effect on demand. At least one piece of anecdotal evidence supports

this claim. When Amazon introduced free shipping in some European countries, the price in France mistakenly was reduced not to zero but to one French franc, a negligible positive price (about 10¢). However, whereas the number of orders increased dramatically in the countries with free shipping, not much change occurred in France. This example also suggests that when trying to use bundling with a cheap good in order to bring up the sales of another good, it might be wise to go all the way down with the cheap good and offer it for free.

Another possible implication of the effect of zero might be in the domain of food intake. When designing food and drink products, companies can decide whether to create low caloric (or fat or carbohydrate) content or reduce these numbers further to zero. Assuming that the effect of zero generalizes to other domains, investing further effort to create a product with zero grams of fat might have a very positive influence on demand.

Decisions about zero might be more complex but also more relevant in domains in which multiple dimensions can occur separately but be consumed together. In the domain of prices, some examples might include cars or computers, for which price is composed of a sum of multiple components, some of which might be set at a standard price and some at zero. In the food domain, these components might be calories, grams of fat, carbohydrates, amount of lead, and so forth, such that some offer a standard amount and some are set to zero. To the extent that the effect of zero holds for individual dimensions that are a part of a complete product, it might be beneficial to consider it at such levels as well.

Appendix: The different types of goods, prices and dependent measures across experiments and conditions. **Go over this at the end**

Experiment	Dependent Variable	Condition	Low-Value Good	High-Value Good
Experiment 1	Hypothetical choice	0&25	Hershey's kiss for 0¢	Ferrero Rocher for 25¢
		1&26	Hershey's kiss for 1¢	Ferrero Rocher for 26¢
		2&27	Hershey's kiss for 2¢	Ferrero Rocher for 27¢
Experiment 2	Real choice	0&14	Hershey's kiss for 0¢	Lindt Truffle for 14¢
		0&10	Hershey's kiss for 0¢	Lindt Truffle for 10¢
		1&15	Hershey's kiss for 1¢	Lindt Truffle for 15¢
Experiment 3	Real choice	0&13	Hershey's kiss for 0¢	Lindt Truffle for 13¢
		1&14	Hershey's kiss for 1¢	Lindt Truffle for 14¢
Negative Price	Real choice	-1&12	Hershey's kiss plus 1¢	Lindt Truffle for 12¢
		0&13	Hershey's kiss for 0¢	Lindt Truffle for 13¢
		1&14	Hershey's kiss for 1¢	Lindt Truffle for 14¢
Experiment 4	Real choice	0&1	Small Snickers for 0 Hershey's	Large Snickers for 1 Hershey's
		1&2	Small Snickers for 1 Hershey's	Large Snickers for 2 Hershey's
Amazon gift	Real choice	0&7	$10 Amazon GC for $0	$20 Amazon GC for $7

46

certificates (GC)		1&8	$10 Amazon GC for $1	$20 Amazon GC for $8
		5&12	$10 Amazon GC for $5	$20 Amazon GC for $12
Experiment 5	Attitude	H0	Hershey's kiss for 0¢	-
		H1	Hershey's kiss for 1¢	-
		L13	-	Lindt Truffle for 13¢
		L14	-	Lindt Truffle for 14¢
	Hypothetical choice	0&13, neutral	Hershey's kiss for 0¢	Lindt Truffle for 13¢
		1&14, neutral	Hershey's kiss for 1¢	Lindt Truffle for 14¢
Experiment 6	Hypothetical choice and ratings	0&13, forced analysis	Hershey's kiss for 0¢	Lindt Truffle for 13¢
		1&14, forced analysis	Hershey's kiss for 1¢	Lindt Truffle for 14¢

References

Allenby, G. M., P.E. Rossi. 1991. Quality perceptions and asymmetric switching between brands. *Marketing Science* 10(3) 185-204.

Ariely, D., U. Gneezy, E. Haruvy. 2006. Social norms and the price of zero. Working paper, MIT.

Ariely, D., G. Loewenstein, D. Prelec. 2003. Coherent arbitrariness: Stable demand curves without stable preferences. *The Quarterly Journal of Economics* 118(1) 73-105.

Ariely, D., G. Loewenstein, D. Prelec. 2006. Tom Sawyer and the construction of value. *Journal of Economic Behavior and Organization* 60 1-10.

Ariely, D., D. Prelec.2006. Arbitrariness in the construction of value: Utility or mapping. Working paper, MIT.

Benartzi, S., R.H Thaler. 2002. *How Much Is Investor Autonomy Worth? Journal of Finance* 57(4) 1593 – 1616.

Bem, D. J. 1965. An experimental analysis of self-persuasion. *Journal of Experimental Social Psychology* 1(3) 199-218.

Chapman, Gretchen B., Eric J. Johnson. 1999. Anchoring, activation, and the construction of values. *Organizational Behavior and Human Decision Processes* 79 115-153.

Deci, E. L., R.M. Ryan. 1985. The general causality orientations scale: Self-determination in personality. *Journal of Research in Personality* 19(2) 109-134.

Diederich, A. 2003. Decision making under conflict: Decision time as a measure of conflict strength. *Psychonomic Bulletin and Review* 10(1) 167-175.

Dholakia, U.M., I. Simonson. 2005. The Effect of Explicit Reference Points on Consumer Choice and Online Bidding Behavior. *Marketing Science* 24 (2) 206–217.

Festinger, L., J.M. Carlsmith. 1959. Cognitive consequences of forced compliance. *Journal of Abnormal & Social Psychology* 58 203-210.

Finucane, M. L., A. Alhakami, P. Slovic, S.M. Johnson. 2000. The affect heuristic in judgments of risks and benefits. *Journal of Behavioral Decision Making* 13(1) 1-17.

Fiske, A.P. 1992. The four elementary forms of sociality: Framework for a unified theory of social relations. *Psychological Review*, 99, 689–723.

Gneezy, U., A. Rustichini. 2000a. A fine is a price. *Journal of Legal Studies* 29(1) 1-18.

Gneezy, U., A. Rustichini. 2000b Pay enough or don't pay at all. *Quarterly Journal of Economics.* 15(3), 791-810.

Gourville, J.T., D. Soman. 2005. Overchoice and Assortment Type: When and Why Variety Backfires. *Marketing Science.* 24 (3) 382–395.

Heyman, J., D. Ariely. 2004. Effort for payment: A tale of two markets. *Psychological Science.* 15(11), 787-793.

Houston, D. A., K. Doan. 1996. Comparison of paired choice alternatives and choice conflict. *Applied Cognitive Psychology*, 10(Spec Issue), S125-S135.

Hsee, C. K., F. Yu, J. Zhang, Y. Zhang. 2003. Medium maximization. *Journal of Consumer Research* 30(1) 1-14.

Huber, J., J.W. Payne, C. Puto. 1982 Adding asymmetrically dominated alternatives: violations of regularity and the similarity hypothesis, *Journal of Consumer Research* 9(1) 90-98.

Iyengar, S. S., M.R. Lepper. 2000. When choice is demotivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology* 79(6) 995-1006.

Kahneman, D., A. Tversky. 1979. Prospect theory: An analysis of decision under risk. *Econometrica*, XVLII, 263–291.

Leclerc, F., C.K. Hsee, J.C. Nunes. 2005. Narrow Focusing: Why the Relative Position of a Good in Its Category Matters More Than It Should. *Marketing Science* 24 (2) 194–205.

Lepper, M. R., D. Greene, R.E. Nisbett. 1973. Undermining children's intrinsic interest with extrinsic reward: A test of the "overjustification" hypothesis. *Journal of Personality & Social Psychology* 28(1) 129-137.

Luce, M. F. 1998. Choosing to avoid: Coping with negatively emotion-laden consumer decisions. *Journal of Consumer Research* 24(4) 409-433. Nunes, J. C., C.W. Park. 2003. Incommensurate resources: Not just more of the same. *Journal of Marketing Research* 40(4) 26-38.

McGraw, A.P., P.E. Tetlock. 2005. Taboo trade-offs, relational framing and the acceptability of exchanges. *Journal of Consumer Psychology* 15, 2-15.

McGraw, A.P., P.E. Tetlock, O.V. Kristel. 2003. The limits of fungibility: Relational schemata and the value of things. *Journal of Consumer Research* 30, 219-229.

Prelec, D., G. Loewenstein. 1998. The Red and the Black: Mental Accounting of Savings and Debt. *Marketing Science* 17(1), 4-28.

Schwarz, N. 2002. Feelings as information: Mood influences judgments and processing strategies. In T. Gilovich (Ed.), *Heuristics and Biases: The Psychology of Intuitive Judgment* (pp. 534-548). Cambridge University Press, New York.

Seife, C. *2000. Zero, the Biography of a Dangerous Idea.* Penguin Books, New York.

Simonsohn U., G. Lowenstein. 2006. Mistake #37: The effect of previously faced prices on current housing demand. *Economic Journal* 116(1) 175-199.

Simonson, I., Z. Carmon, S. O'Curry. 1994. Experimental evidence on the negative effect of product features & sales promotions on brand choice. *Marketing Science* 13 23-40.

Slovic, P., M. Finucane, E. Peters, D.G. MacGregor. 2002. The affect heuristic. In T. Gilovich (Ed.), *Heuristics and Biases: The Psychology of Intuitive Judgment* (pp. 397-420). Cambridge University Press, New York.

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

Tversky, A., E. Shafir. 1992. Choice under conflict: The dynamics of deferred decision. *Psychological Science* 3(6) 358-361.

Table 1: Predictions of the Standard Cost–Benefit Model and Zero Price Model.

		Standard Cost–benefit Model	Zero Price Model
Changes in valuations	Costs	Both costs decrease by the same amount	
	Benefits	Both benefits remain the same	Benefit of the low-value good increases
	Net benefits	Net benefits increase by the same amount	Net benefit of the low-value good increases more
Changes in demands		Some switching from nothing to something	
		No switching between goods	Some switching from high-value to low-value good
		Demand for the low-value good increases	Demand for the low-value good increases
		Demand for the high-value good increases	Demand for the high-value good decreases

Notes: The table illustrates predictions as the prices for two products move from $[P_X, P_Y]$ (where $P_X < P_Y$) to $[0, P_Y\text{-}P_X]$.

Figure titles

Figure 1: Segments of customers who choose options X, Y, and N as prices go down from [Px, Py] to [Px-ε , Py-ε], as predicted by the standard economic model with linear utilities and the zero price model.

Panel A presents the demand distribution when prices are [Px, Py].

Panel B presents the changes in segments of customers choosing options X, Y, and N when prices are reduced from [Px, Py] to [Px-ε , Py-ε].

Panel C presents the changes in segments of customers choosing options X, Y, and N when prices are reduced from [Px, Py] to [0, Py- Px] under the assumptions of the standard model.

Panel D presents the same changes under the assumptions of the zero price model.

Figure 2: Proportions of consumers choosing Hershey's and Ferrero Rocher chocolate across the three experimental conditions in Experiment 1.

Figure 3: Proportions of consumers choosing Hershey's and Lindt across the three experimental conditions in Experiment 2.

Figure 4: Proportions of consumers choosing Hershey's and Lindt across the two experimental conditions in Experiment 3.

Figure 5: Proportions of consumers choosing small and large Snickers Bars across the two experimental conditions in Experiment 4.

Figure 6: Proportions of consumers choosing the $10 and $20 Amazon gift certificates across the three experimental conditions in the follow-up to Experiment 4.

Figure 7: Affective ratings of the four offers in Experiment 5.

Figure 8: Proportions of consumers choosing Hershey's and Lindt across the experimental conditions in Experiment 6.
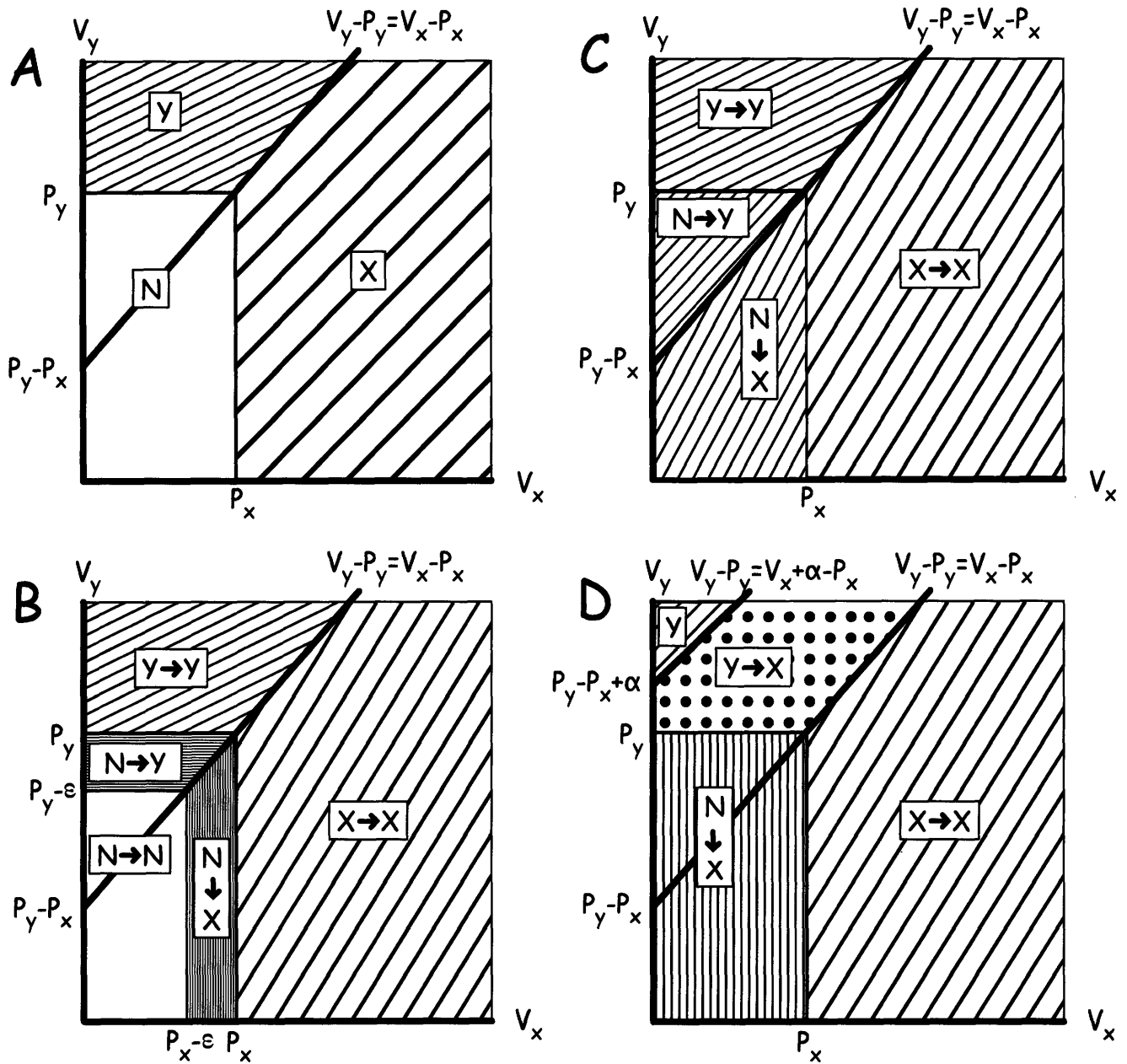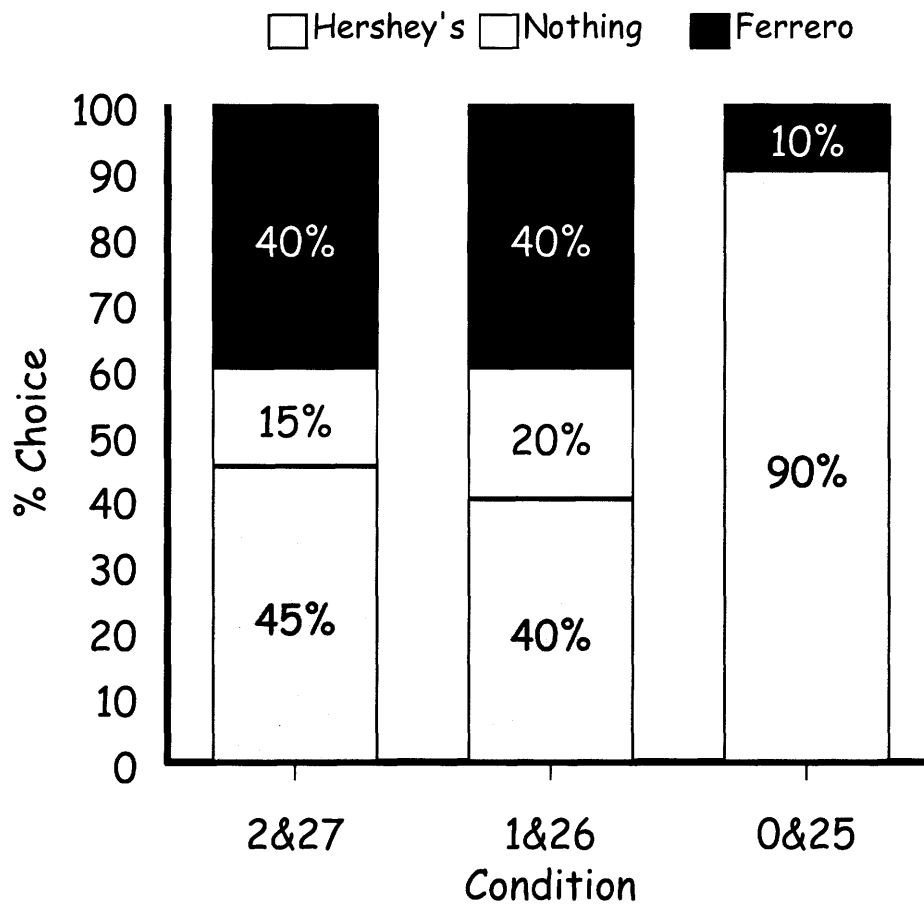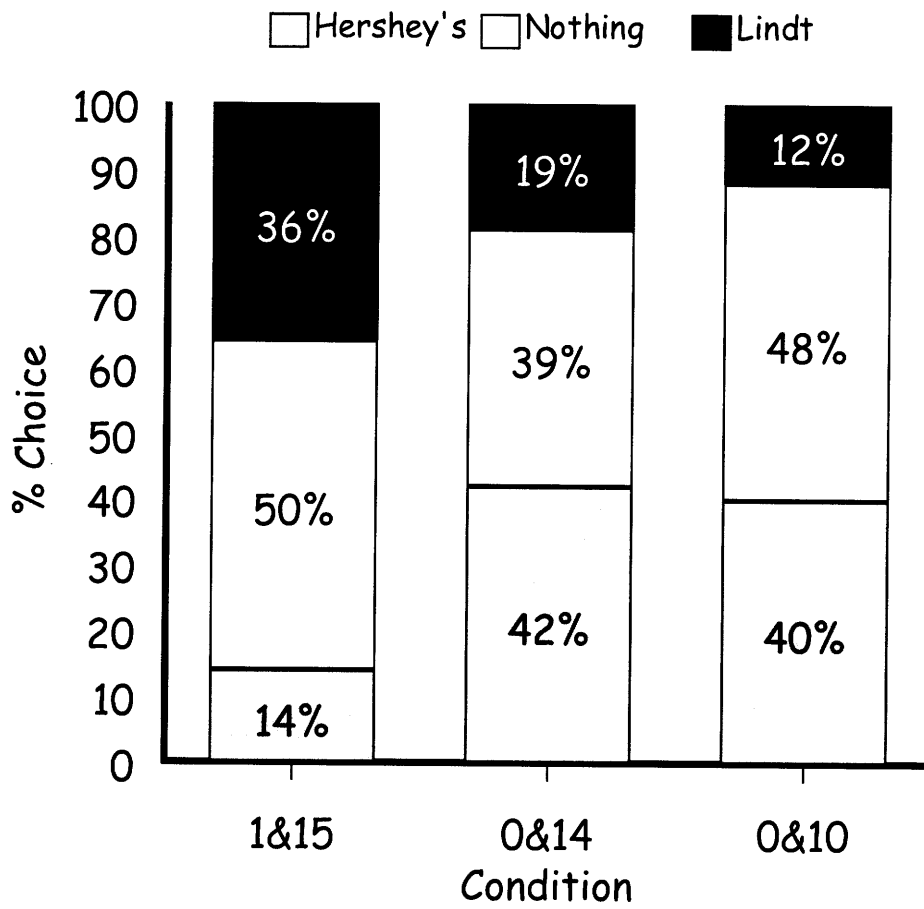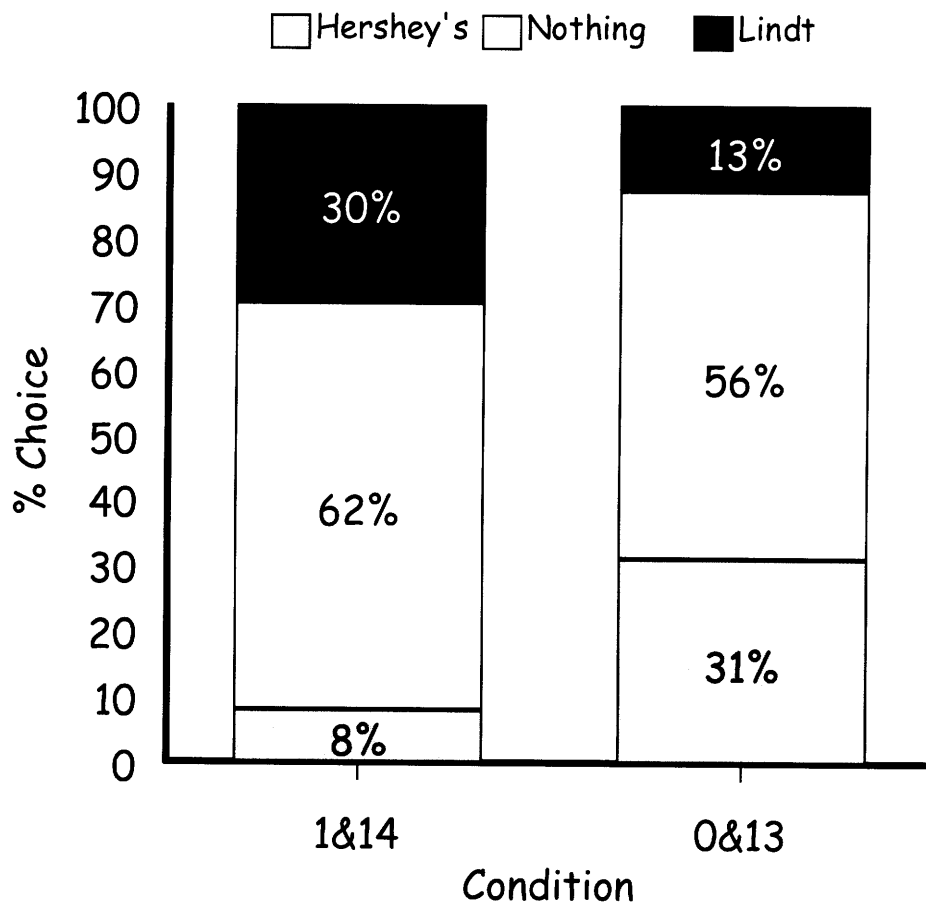
Figure 1

Figure 2

Figure 3

Figure 4

Figure 5

Figure 6

Figure 6 — Bar chart showing % Choice by Condition. Legend: $10 Gift cer (white), Nothing (white), $20 Gift cer (black).

Condition 5&12: 71% ($20 Gift cer), 29% (Nothing).
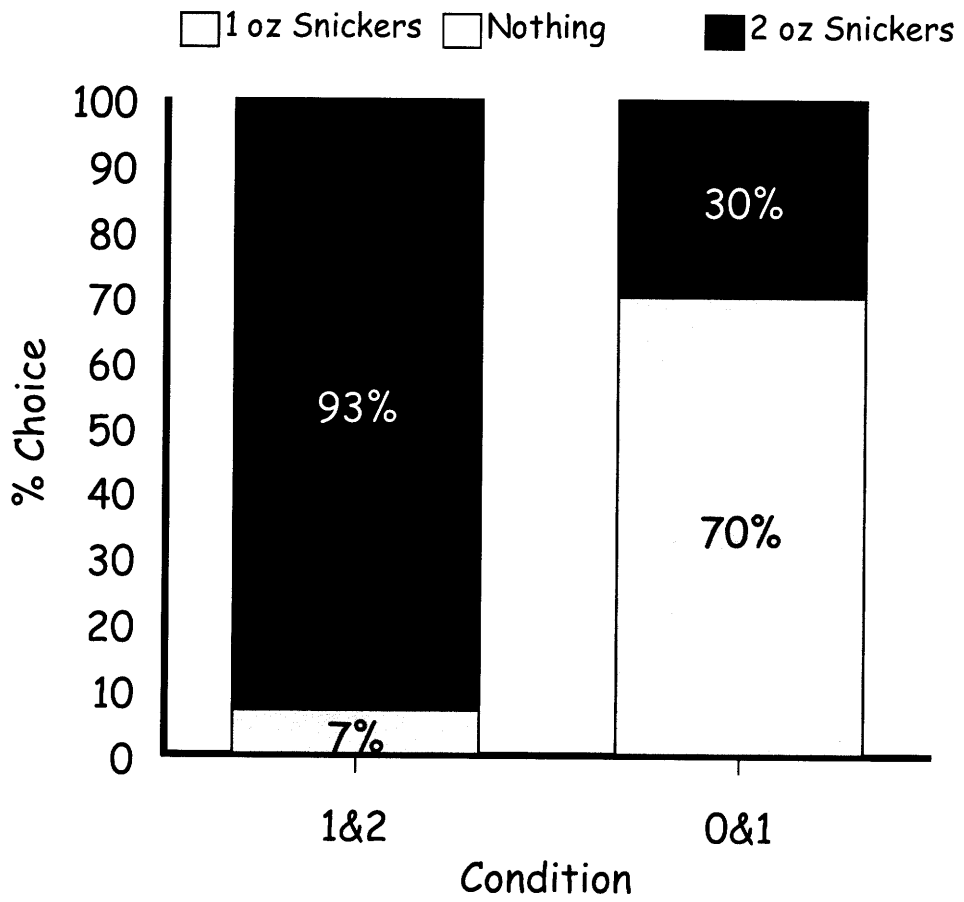Condition 1&8: 64% ($20 Gift cer), 36% (Nothing).
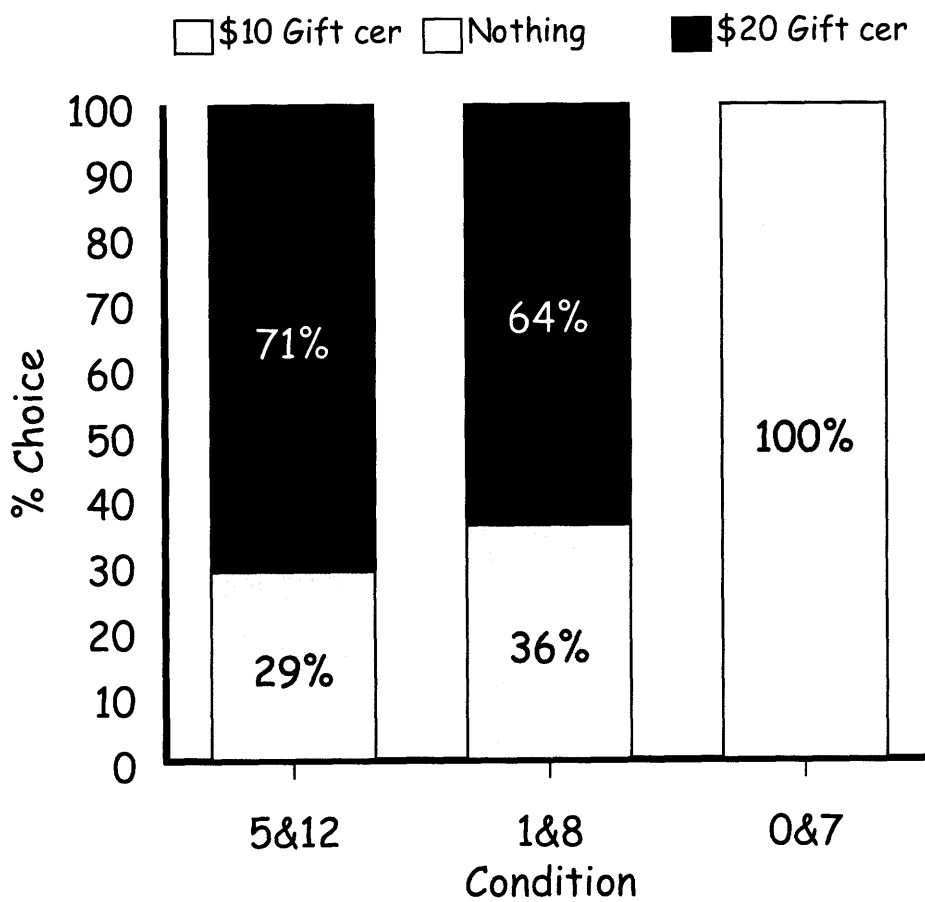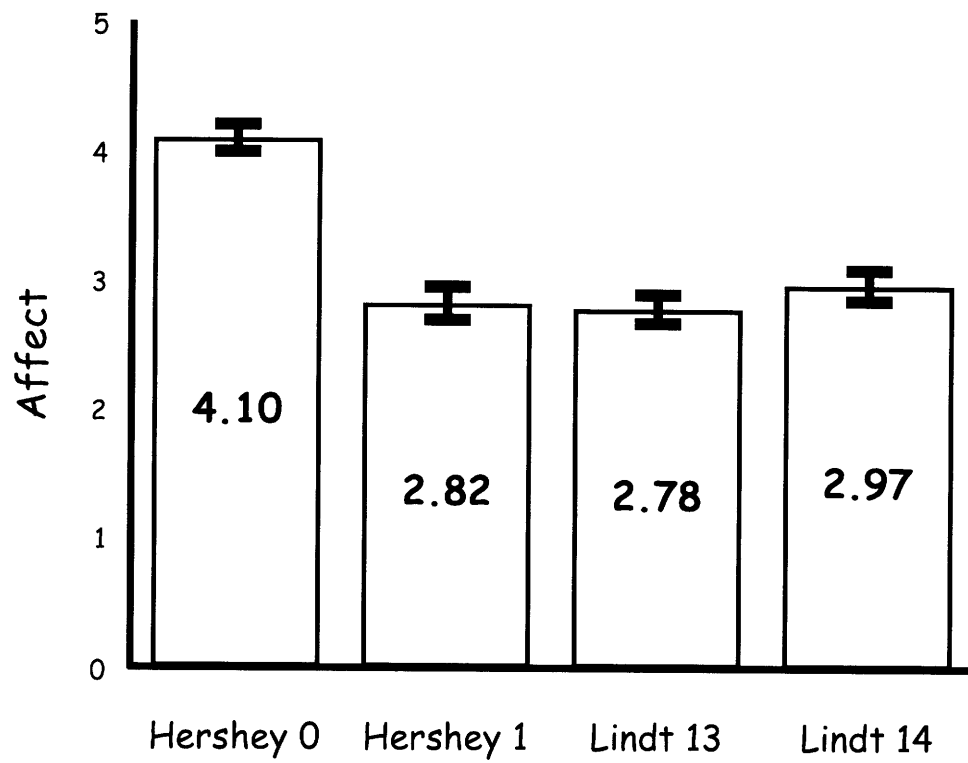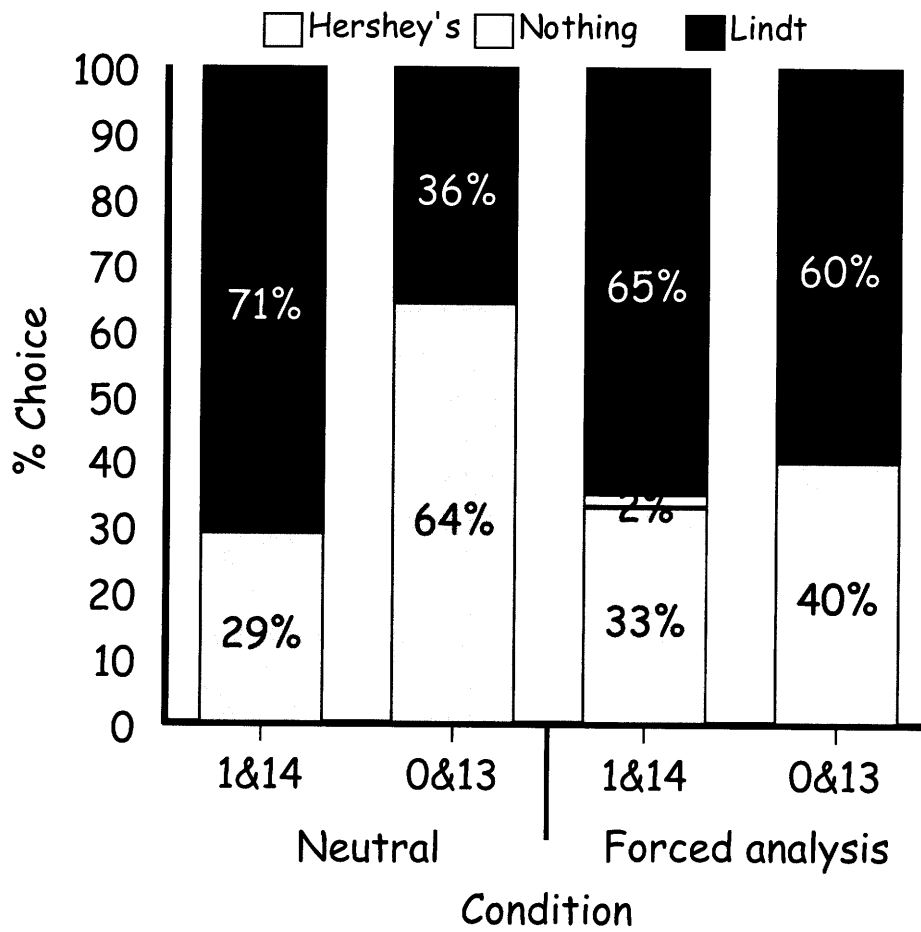Condition 0&7: 100%.

Figure 7

Figure 8

Essay 2
Movies as a mood regulation tool: movie watching patterns right after September 11.

Abstract

Is a sad person more, less or equally likely than a happy person to pursue a "happy" activity rather than an "unhappy" one (e.g. prefer a comedy to a drama)? Surprisingly, the literature offers theories *and* laboratory evidence in favor of *all three* possibilities. In this paper, we attempt to resolve the puzzle by moving out of the lab and analyzing the changes in movie watching patterns following the tragic events of Sep 11, 2001. Two data sets from the 7 weeks surrounding 9/11 are analyzed. One consists of US box office collections of top ten movies during the period. The other contains data on movie rentals in a rental store chain in Cambridge MA. The analysis suggests that the more private the mood-regulating decision is (rental vs. movie going), the more likely is the person to use the movie as a mood repair tool. When the decision is more public (movie going), the appropriateness issues induce to mood congruent behavior.

Movies as a mood regulation tool: movie watching patterns right after September 11.

Enormous body of literature on mood regulation and mood congruency reviewed in part below does not give an unequivocal answer to a simple question "what kind of movie will you want to see when you are sad?" More generally, will you always peruse positive emotions or, perhaps, you will be inclined to engage in activities in line with your current mood, including doing "unhappy" things when you feel unhappy. In this paper we will attempt to answer these questions by analyzing movie watching patterns on and right after Sep 11, 2001. More generally, we will attempt to answer the following questions:

Does mood regulation exist? When sad, will you watch a sad or a funny movie?

We will start with an overview of literature favoring various hypotheses with respect to the second question. It turns out that all three possible hypotheses could have merit; when sad, you could be more, less or equally likely as when you are happy to prefer a happy movie to a sad movie. We then present empirical evidence from the literature to demonstrate that experimentation does not resolve the problem, because all three hypotheses are supported by experimental evidence. Next, we move on to our own empirical evidence that includes two field studies related to movie watching patterns on and around 9/11 to see if real life data can bring resolution to the puzzle.

Predictions

At the fist glance, everybody should want to be happy at all times. As Andrade &
Cohen (2007) put it, "[I]ndeed, 'hedonism's' prime directive – i.e., people's tendency to
pursue pleasure and avoid pain – is one of the most well grounded assumptions in
psychology and consumer behavior." Isen (1985) provides an overview of literature
demonstrating that "positive affect...tends to facilitate the recall of positive material in
memory," whereas "sadness does not facilitate the recall of compatible material." Isen
(1985) proposes that people in positive mood are motivated to maintain their moods,
whereas those in negative mood are motivated to repair their mood. Thus, independent of
mood, a person should be more interested in activities that make her happier. In case of
the movies, other things being equal, a person should prefer to watch a comedy than a sad
movie.

One might wonder, if people should only seek happy material, why do dramas
exist at all and why do people watch them? A possible answer is that the level of movies'
"happiness" is not the only dimension that enters movie selection; there is also quality,
educational content, and probably others. In this paper, we do focus on the happiness
dimension, however; and the first hypothesis that we will consider is the following

H_1: *The preferences in choosing between a sad movie and a comedy will not*

depend on the chooser's mood. Other things being equal the happier movie will

be preferred.

However, the need to experience something positive might be stronger at certain situations than in the others. It is logical to assume that you need to be cheered up more when you are down than when all is well. At the same time, when all *is* well, one might be more tolerant to negative material, and it could be a good time to watch some high quality drama. For example, Zillmann (1988) suggests that a person in a bad mood "should be motivated to diminish the hedonic quality of this mood, to escape it altogether and if possible to enter into the hedonically opposite state...Individuals in bad moods should therefore be partial to consuming highly absorbing pleasant fare that features activities with little affinity to their experiential state." At the same time, "[p]ersons in a good mood are less in need of being cheered up. They can be expected to consume more of the alternatives to comedy." So the next, possible prediction is that people in bad mood are more likely to engage in mood-uplifting experiences than those in a good mood. In the case of the movies, then, other things being equal, a sad person should be more likely than a happy person to want to see a comedy rather than serious drama. Thus, the second hypothesis to consider is

H_2: A person in a sad mood is more likely to choose a comedy over a sad movie than a person who is in a good mood.

An alternative line of argument, suggests however, that when your are sad, almost anything can cheer you up, "the range of available activities would be almost entirely more positive (or less negative) than the person's present mood", whereas when in a good mood "most available activities would make the person feel worse" (Wegener & Petty,

72

1994). These arguments lead Wegener & Petty (1994) to put forward the *hedonic contingency hypothesis,* stating that people in a good mood should put more effort into mood maintenance than should people in a bad mood put effort into mood repair; "scrutiny of the hedonic consequences of potential future activities should become more usual, more practiced, and more likely in positive than in negative moods." This account results in a prediction that a happy person is more likely than a sad person to peruse a happy experience than a sad experience.

The same predictions can be derived from a different and unrelated account. According to prospect theory (Kahneman & Tversky, 1979), people experience losses more intensely than gains of the same amount. Thus, even a small decline of a good mood should feel more serious than a same size improvement in a bad mood; happy people should desire to remain happy much more than sad people should desire to improve their mood. Consequently, prospect theory also predicts that relative to a sad person, someone in a good mood should be more likely to prefer a comedy to a serious movie, because, other things being equal, the sad movie threatens the good mood more than the happy movie promises to improve the bad mood. Thus, both these accounts (hedonic contingency and prospect theory) predict somewhat counter intuitively that sad people will choose a sad movie more often than a happy person.

Note that both these accounts are based on asymmetries between sad and happy moods; in the first case, it is easier to loose a good mood than to repair a bad one, in the second case, it is more desirable not to loose a good mood than to repair a bad one. However, a third account with the same basic prediction brings symmetry into the picture. This account simply states that everybody is choosing movies and other

activities in line with their mood. This proposition can be supported by numerous consistency theories. For example, Caruso & Shafir (2006) write

"Affect can produce memories (Bower, 1981) and judgments (Isen *et al.*, 1978) that are congruent with one's current mood, and can ultimately predispose people to behave in affectively consistent ways (Leventhal, 1980). Thus, all else being equal, when choosing between movies, people in a sad mood might be expected to pick a heavy drama more often than people in a happy mood, who might prefer a comedy."

Thus, three separate accounts lead us to the third hypothesis

H₃: A person in a good mood is more likely to choose a comedy over a sad movie than a person who is in a sad mood.

Review of empirical evidence

As demonstrated above, various arguments and theories lead to three hypotheses that exhaust all possibilities available. Sad people in comparison to happy people could arguably be less, equally, or more likely to choose a sad activity over a happy one (e.g. a drama over a comedy). Not surprisingly then, the empirical studies on this topic are abundant. Problematically and surprisingly, they give support to *all three* hypotheses. In the following three sections we review empirical evidence on each of the three hypotheses. Given that the literature in general supports all three of them, it is no too surprising that some of the papers actually support more than one, thus the split into subsections is rather relaxed.

Evidence to support that mood valence has no effect of the choice of mood-relevant stimuli (H_1)

The majority of experiments that we will review consist of a mood manipulation followed by a choice of some mood-relevant activity (e.g. a movie) or recall of a memory. The researchers are usually interested in whether the choice (or recall) will be consistent or inconsistent with the mood. Caruso & Shafir (2006) use a similar experimental paradigm but the question they address is orthogonal to the usual one. Following Isen (1985), they suggest that because sad people "should" engage in mood improvement, and happy people "should" engage in mood maintenance, mood valence will have no effect on choices between sad and happy experiences, everyone should choose the happy ones. However, when people are not paying attention to their mood, they will not engage in any mood regulation activity, and thus they will be less likely to choose a happy activity. Note the choices will not be different depending on the mood valence, only mood salience is important, thus the predictions are in line with H_1.

In a series of experiments Caruso & Shafir (2006) offer their participants a choice between a high-quality serious dramatic movie (or audio) and a low-quality funny one. Consistent with the author's predictions, the participants' choices do not depend on their mood, but rather on whether their mood has been brought to their attention; those aware of their current mood are more likely to choose the comedy. For example in Study 3, participants are asked to choose between two audios; one is described as a low quality routine by an amateur comedian, the other as a recording of a Nobel Prize winning poet reading an excerpt from one of his most highly acclaimed poems. For half of the

participants, the choice is preceded by a rating of their mood in the end of an ostensibly unrelated study. In line with the authors' predictions, those who were asked about their mood before making their choice were more likely to choose the comedy, whereas the mood rating for these people was not related to the choice. The fact that only mood salience – but not mood valence – has an effect on the subsequent content choice in this and three other studies in the paper favors H_1.

There are several problems, however, for the finding to be conclusive. First, in none of the four experiments is mood actually induced; moods are simply measured in studies 2 and 3, and imagined in studies 1 and 4 (in addition, in studies 1, 2, and 4, the dependent measure is participants' intuition of how they would behave rather than actual behavior).

Obviously, people might act differently in an actual affective state compared with an imagined one. A bulk of research on predicting behavior in an affective state shows that these predictions are far from accurate. The findings on mispredictions of own and others' behavior in a "hot" state while being in the "cold" state (and vice versa) have been jointly dubbed "hot-cold empathy gap" (see Loewenstein, 2005 for a short summary or Loewenstein & Schkade, 1999; Van Boven & Loewenstein, 2003 for exhaustive reviews). One can point, however, that this literature discusses fleeting and very intense visceral states, such as hunger, thirst, sexual arousal, and drug craving, rather than less intense and more lasting moods. However, lack of intensity can be only narrowing the gap between intuition and behavior rather than closing it.

As for non-experimental mood measurement, it could be capturing even more stable states rather than more passing moods. For example, if depressed people are less

inclined to watch a comedy than people who are in a bad mood temporally, then the outcome favoring H_1, might be masking what in reality is H_2.

More support for H_1 is provided in Andrade (2005). With similar pretext and predictions as in Caruso & Shafir (2006), but with a somewhat better procedure the results are also similar. For example, in his Study 1, participants are first put in a sad, neutral or positive affective state (by watching a drama, or a documentary, or a comedy), and a manipulation check is conducted (thus, mood was salient for all participants). Then in an ostensibly unrelated study, participants are asked whether they would (hypothetically) sample chocolates in a promotion in exchange for filling out a survey. Participants are also asked whether they usually eat chocolate to feel better. The results show that those who do eat chocolate for mood repair purposes (mostly women) are equally likely to sample the chocolate whether they are in a positive or a negative mood (more likely than in neutral mood), thus supporting H_1. Similarly in Study 2 participants in bad and good mood are equally likely (more than neutral mood participants) to refuse to participate in a lengthy and thus mood-threatening survey. One problem that these studies have is that the dependent measure is a reflection of participant's intuition on how they would behave, rather than actual behavior. As with the independent factors, it could be that mood effects are very different in reality than when predicted or imagined. For example, when sad, you might think a chocolate would cheer you up, but given an opportunity to have the chocolate, you might feel too down to have it.

Evidence to support of mood repair (H_2)

Josephson *et al.* (1996) present support for both H_2 and H_3. Participants are induced with a sad or a neutral mood by watching a video and then asked to write down a memory, and when they are done, they are asked to write down another one. In line with H_3, sad participants are more likely, relative to participants in a neutral mood, to recall a sad memory. However, when asked to retrieve a second memory, the same participants with prior low depression scores are more likely to retrieve a happy memory and 68% of those who do so "mentioned mood repair as motivating the recruitment of the more positive second memory," thus supporting H_2.

Complicating matters even further, Parrott (1993) suggests and provides some empirical evidence that mood regulation can occur in any direction, including inhibiting positive moods and maintaining negative moods, for reasons such as motivating hard work. Parrott (1993) first provides a short overview of literature on mood congruence in memories, with ample evidence that people are likely to recall memories congruent with their current affective state. That is, a person who is happy (sad) is more likely to spontaneously recall happy (sad) memories (in line with H_3). According to the author, the literature "depicts mood congruence as an automatic result of associational links between the affective content of stored material and the present mood of the person." Next Parrott (1993) provides empirical evidence of mood *incongruent* recall and proposes a mood regulation explanation (in line with H_2). In one study, students who have just received their grades on an exam, are asked to recall three events from their high school years. The first of the recalled events is on average much happier for "unhappy" students (those whose grade was below expected) than for "happy" students (whose grade superseded expectation), with no difference between "conditions" in the two subsequent memories.

Three other studies have similar outcomes. In one, students are asked to recall three memories as they are entering the library. The "mood manipulation" is achieved by approaching some students on a sunny and some on cloudy days. In another study, mood was experimentally induced by music. In the third study with an additional neutral condition, neutral participants place in between sad and happy ones. Thus, at least four studies favor H_2.

The author suggests that in all these studies participants try to overcome their current mood by managing it with an incongruent memory. In addition to the natural hedonic explanation for subjects trying to overcome bad mood, Parrott (1993) suggests that subjects in a good mood try to overcome it in order to bring themselves into a more serious state required by the situation (school, library, music evaluation). He also suggests that previous results with congruent mood recall can be attributed in part to the fact that the participants in those experiments are often asked to bring themselves in a particular mood and probably keep working on that when producing memories. Thus, overall, Parrott (1993) overviews results that can be seen as evidence for H_2 and H_3 without a final resolution in favor of one of them.

Evidence to support of mood repair (H_3)

To start with, all the mood congruent recall literature (e.g reviewed in Parrott, 1993) can be seen as some evidence in favor of H_3. Indeed, if people have better recall of things congruent with their current mood, then their choice sets for behaviors will also be in line with their current mood and hence the action chosen is likely to be in line too. However, there is also more direct evidence in favor of H_3.

For example, testing their hedonic contingency hypothesis, Wegener & Petty (1994) give empirical support to H_3. In their findings, happy participants have a greater preference for happy movies than sad participants. In several experiments after various mood induction procedures (e.g. watching/reading a sad, happy or neutral video/article), student participants rank eight movies with respect to how much they want to see each of them. The major finding is that the four tapes that are ostensibly rated highly on the "happy" dimension by other students get higher rankings from happy than from sad participants.

For example, in Study 2 participants first read a sad, happy, or neutral article. Then they rank eight tapes according to how much they would like to watch each of them. The tapes have been ostensibly rated by fellow students on various dimensions, including "how happy the tape made students feel"; the average ratings are provided to the participants as they are ranking the tapes. In line with the predictions (and in line with H_3) happy participants rank happy movies as more desirable for viewing than sad and neutral participants (with no difference between sad and neutral).

In a similar study Handley *et al.* (2004) replicate this finding with the movies that are not claimed to be happy or sad, but simply use in the titles, (neutral) words previously paired with happy or sad words. That is, before the mood induction or the dependent measure, the participants watch word pairs on a screen one at a time, such that one word in the pair is always neutral, and the other is positive, negative, or neutral. This task is the same for all participants. Then the mood induction with a happy, sad or neutral movie and a manipulation check are performed. After the manipulation check each participant is asked to rank-order 8 fictitious movie titles on her desire to see the movies. The fictitious

titles are partially composed of the neutral words used in the first task. "Half of the titles contained two unique neutral words that were previously paired with two positive words and half of the titles contained two unique neutral words that were previously paired with two negative words." Results replicate those in Wegener & Petty (1994); happy participants have more desire than others to view the movies who's titles' words were previously associated with happy words.

In a somewhat similar experiment Zillmann *et al.* (1980) (described in Zillmann, 1988) also finds support for H_3, although they expected to find support for H_2. In the first part of the experiment, participants receive false – negative, positive or neutral –feedback in an emotion-recognition test. Participants are then either told that they are lacking an important social skill or are praised for very good performance, or are said to be at the level of most people (depending on the condition); thus bad, good or neutral mood is induced. After the mood induction, during an ostensible delay in the experiment participants have an opportunity to watch some TV; the available programming – situation comedies, action dramas, and game shows – is prerecorded and played back from an adjacent room.

In line with H_2, the authors predict that participants in a bad mood will have a strong preference for comedies, whereas those in a good mood will "consume more of the alternatives to comedy." In reality, whereas good-mood participants are less interested in comedies than neutral mood participants, contrary to predictions, sad-mood participants are not interested in comedies at all. "Not only did they not prefer comedy, they clearly shunned it." Zillmann (1988) explains the negative result by the nature of the humor in the stimuli. "The predominant form of humor in the prime-time comedy is hostile humor

that thrives on teasing and put-downs...These subjects had just suffered failure and belittlement themselves, and exposure to others' belittlement, although humorous, was likely to perpetuate their annoyance." Thus, although it could be the case that with a more generic mood induction procedure (i.e. recall of a memory) the results would be more consistent with H_2; as it is, we can say that the findings support H_3.

To sum up, the question of whether people will always pursue stimuli in line with their mood, or in the opposite direction, or whether mood valence does not matter at all, remains an open question. In this paper we move outside the lab, and look at the "real life" data. Two data sets are explored. Both data sets reflect possible temporary changes in movie preferences resulting from the 9/11 disaster (Galea *et al.*, 2002). The first data set (Study 2) reflects US box office collections by the top movies. The second set (Study 3) consists of data on movie borrowing from a local movie rental store chain in Cambridge, Massachusetts. However, we start our empirical enquiry by a simple experiment looking into intuitions rather than behaviors.

Study 1. Intuition.

Before starting the massive collection of the real life data, we ran a short survey to see what the intuition was with respect to the question "Depending on the mood, what kind of movies do people prefer to see?"

Method. Three hundred thirty seven UCLA students read the following

"We would like you to indicate which kind of movie or movies you think that each of the following people would go see given the kind of day that they're having.

Mr. A is having a very good day. Mr. A would likely go see a (circle all those that apply):

Sad movie

Happy movie

Comedy

Romantic comedy

Documentary

Other_____ "

The question was followed by two similar ones describing Mr.B having a neutral day, and Mr. C having a bad day.

Results. The proportions of participants (with standard errors) who circled "Sad movie," "Happy movie," "Comedy," and "Romantic Comedy," and "Documentary" are depicted in Figure 1. Among the movies that participants spontaneously suggested in the "Other" option were the following (numbers in parentheses are the number of observations for Mr.C/B/A):

1. "action", "action thriller", "action/suspense", "horror, action," "thriller"(10/10/17)

2. "adult", "porno" (3/0/0)

3. "anything"(0/0/1)

4. "drama" (2/2/1)

5. "horror", "horror, action," "scary" (5/2/8)

6. "indie flink" (0/1/0)

7. "none" (5/0/0)

8. "sci-fi," "sci fic" (0/1/2)

9. "Other" checked without any suggestion (3/5/1)

Given that the numbers were so small, no analysis was carried out with these data.

To compare predicted behaviors across moods, we run five logit regressions with the dummies corresponding to these five movie types as DVs and the dummies for "BadDay" and "GoodDay" as independent variables; we used robust standard errors and clustered observations for each participant to control for the within subject nature of the experiment.

Table 1 shows the coefficients, of these regressions with z-values in parentheses. Note that all the p-values that are below .05 are also below .01, and thus the significant coefficients will remain significant after the Bonferroni correction.

The regressions show that relative to people in a neutral mood, those having a bad day are predicted to be more willing to watch a sad movie, and less likely to watch a happy movie, a comedy or a romantic comedy. For someone having a good day, the predictions are almost the opposite, except that the happy movie is predicted to be wanted as much by a person having a neutral day. It is still true, however, that a person having a good day is predicted to be more likely to watch a happy movie than someone having a bad day ($\chi^2(1) = 31.45$, $p < 0.0001$). As far as the documentary is concerned, the person having a bad day or a good day are predicted to be equally (un)interested in it ($\chi^2(1) = 1.18$, $p = 0.28$), and less likely than a person having a neutral day.

Discussion. The results seem to be mostly in line with H3, people are predicted to prefer the movies in line with their mood.

Study 2. Box-office collections: Before and after 9/11.

The literature review, presented in the beginning of the paper offers contradictory empirical evidence on mood effects on choice of mood-relevant material. One possible

explanation for the contradictory empirical findings is that each experimental manipulation tackles only some aspects of mood at a time, whereas "in the wild," mood is very complex and all its factors interactively contribute to the mood-relevant behavior. Another aspect of lab results is that participants are allowed to choose among a limited number of activities none of which could be what the same participants would choose outside the lab for mood-regulation. Thus, both the manipulation and the DVs in those studies suffer from lack of ecological validity. A test that involves naturally occurring moods and naturally occurring choices of mood-relevant material might shed some light on what the "true" effect is. For this reason, we choose to examine how the natural decisions about natural mood regulation tools (selection of a movie in a theater or a rental store) were influenced by a naturally occurring mood shift resulting from the tragic events of September 11, 2001.

Data. The box office data were obtained from www.the-numbers.com. The data contain US box office collections of the top 9-20 movies on each day between August 20, 2001 and October 7, 2001. That is, the range encompasses the week of September 11, as well as three full weeks before and after. Apart from the gross box office collections, the data include the number of theaters where the movies were screened, and thus the per-theater collections as well. For most of the dates, the data were available for only the top 10 movies (only 9 top movies for September 9). To reduce selection effects, we excluded all data for movies ranking 11 or worse from the analysis, leaving us with a total of 489 observations; that is 29 movies with 1 to 49 observations per movie ($M = 16.9$). Our main dependent variable is the logged box office collection of a movie on a particular day.

In addition, we obtained "happiness ratings" of each of the movies in the data set. Five independent raters were asked to rate each of the 42 movies in the data set on a scale from 1(not at all happy) to 10 (very happy) with regards to how happy an average viewer would feel after seeing each movie. Four of the five raters were employees of a video rental store; the fifth was a self-proclaimed movie lover and expert. The raters were instructed to use their own viewing experiences for the ratings, or if they had not seen the movie, to use their friends' comments, trailers, etc. They could also leave the rating blank (only one rater used this option for only one movie). The Cronbach's reliability coefficient for the five ratings is 0.79; the average of the five ratings will be used for analysis (M = 3.96, SD = 1.62, min = 1.6, max = 8.2).

Analysis strategy. The major goal of the analysis is to determine how the dire events of 9/11 influenced the movie-going patterns of the public, and whether this influence differed for "happy" and "unhappy" movies. We explore two types of changes that could occur. First, we are interested how the overall collections are influenced by 9/11, that is, did fewer people go to the movie theaters on that and the following days? If going to a movie theater is considered a mood-lifting activity independent of the movie type, then this question is important. A drop in the box-office collections on and after 9/11 could be supportive of H_3; an increase could be supportive of H_2. Second, we are interested in how the movie going patterns changed differently for happier and unhappier movies. For example, an increase in happy movies' collections and a decrease in the unhappy movies' collections would be evidence in favor of H_2, whereas the opposite trend would be evidence in favor of H_3.

To achieve the first goal, we put the logged box office collections of the movies on the left hand side of a regression, and the dummies for 9/11 and the following days on the right hand side to see how (and if) the overall collections were influenced by 9/11. We also control (in this regression and in all later regressions) for the number of days that the movie has been playing in the theaters, the day of week, and include the dummies for September 2 and September 3, because the latter was Labor Day Monday (and thus exhibited patterns similar to Sunday, while Sep 2 was a Sunday acting as a Saturday). All the regressions also have fixed effects for movies, and robust standard errors are used. The behavior of the 9/11 dummy and the dummies for the following days, will allow us to see how the 9/11 influenced the box office collections in general.

To achieve our second goal, we add to the right hand side of the regression, the interaction effects between the happiness score and the dummies for 9/11 and the following days. These interaction effects will allow us whether 9/11 influenced differentially the collections of happier and less happy movies. If these interaction effects are insignificant, we will obtain evidence for H_1. If the interaction effects are significant, then we will obtain evidence for either H_2 or H_3, depending on the signs of the interactions and of the dummies for 9/11-9/16.[1]

Before the discussion of the results, it is important to note that the gross box-office data reflect both the demand of the public and the supply of the movie theaters. The latter reflects to some degree the intuition of the owners on the movie theaters for preferences of the public, and thus could also be of some interest in the discussion of how

---

[1] Note, the happiness score does not appear as one of the regressors (only in interactions with the dummies for 9/11-9/16), because it is already accounted for in the fixed effects.

the movie choices are influenced by mood. However, our main interest is still on the demand side, and that is probably better reflected in the per-theater data. Thus for every regression that we report, we run it with two possible DVs, the daily gross earnings and the daily gross earnings per theater (both logged). We also run regressions (we do not report them here) with the number of theaters where the movie was shown as the dependent variables and the same independent variables. These regressions show no effects of 9/11 and thus it is safe to assume that the effect we observe in the regressions with box-office (overall and per theater) collections as DVs reflect the demand side, and it explains why little difference will be observed between the behavior of these two DVs.

Results. The regression results are shown in Table 2; each column represents a set of coefficient estimates in one regression with t-values in parentheses and the DV shown in the top of the column. The first two regressions look on the impact of 9/11 on the overall box office collections. First of all, all the control variables behave as expected; the collections are larger on the weekends (e.g. Saturday brings approximately 168% more than Monday), the longer the movie is out, the less it collects (every additional day reduces the collections by approximately 5%), September 2 and September 3 did observe larger collections than expected, given the days of week. The controls behave similarly in all regressions, and we will not discuss them further.

The first two regressions also show that overall box office collections dropped on 9/11 by about 70%, 9/12 by about 16%, and by about 9/13 by about 20%, but not on the weekend[2]. In the next two regressions, the dummies for 9/11-9/16 are interacted with the happiness score. All the interactions are negative and all but one are significant; that is, the happier the movie is, the less it collected on these days in comparison with unhappy

---

[2] 9/11 was a Tuesday.

movies. For example, on 9/11 every additional point in the happiness score reduces the movie's collections by about 10% on average, which results in a difference of 60% between two movies with happiness scores of 2 and 8. The last two regressions basically show the same results but with a single dummy for all the days between 9/11 and 9/16.

Discussion. The results are in line with H_3 again. It seems that the events of 9/11 made people abstain from happy movies and pursue movies in line with their sad mood. Several technical issues with the data prevent us however from drawing the final conclusions. First, the data analyzed include only the top movies as opposed to the abundance of movies that were screened in that period. Although there is no reason to believe that the top movies are not representative of the population with respect to the patterns in question, it is still hard to make final conclusions based on a very small and biased (in terms of the DV) subsample. Second, as discussed above, the data reflect both demand for and supply of movies; although we are confident that the number of theaters showing each movie did not change significantly due to 9/11, we have no data on the number of shows per day – it is possible that the number of "happy" shows was reduced. Finally, going to a movie is a public experience. People often go with friends, and the choice of the movie is observed by them, other fellow movie goers, and the theater staff. It is quite reasonable to assume that the decision to go see a comedy or another happy movie on or right after a national tragedy might be considered as very inappropriate, even if it is a good mood-repair technique. Thus, it will be fruitful to supplement our analysis with analysis of a better data base. This is what we do in Study 3.

Study 3. Movie rentals: Before and after 9/11.

Given the identification problems in Study 2, the inferences we can make from its results are rather limited. The movie rental data set in Study 3, on the contrary, reflects the demand side only effects. In addition movie rental as opposed to movie going is a much more private choices, made on more regular basis, and from a much larger choice set. Thus, hopefully, Study 3 will provide us with more credible results and to some extent, will let us compare how more private movie choices versus more public ones were influenced by 9/11.

Data. The data were obtained from a Cambridge (MA) video rental store chain. In the original data set, the unit of observation is a transaction (rental of one VHS, DVD, video game, or piece of equipment). Each data point includes the date, title, category, and consumer ID. The data used in the analysis consists of all VHS and DVD rentals between August 29, 2001 and October 7, 2001. In the original data, the titles were categorized in very many categories (categories used by the store chain), some of which were very narrow (e.g. "Martin Scorsese"), so we use our own categories in the analysis. The two that we will focus on are "Comedy" and "Drama." The former consists of the original categories "Comedy," "British Comedy," "Classic Comedy," "Silent Comedy," "Stand-Up," "TV Comedies," and "Woody Allen;" the latter consists of the original categories "Drama," and "New Drama." When referring to categories "Comedy" and "Drama," we refer to our broader categories.

Preliminary Analysis. For the initial analysis we collapse all data so that we have one observation per day, containing the total number of movies rented that day and the number of movies rented in "Comedy" and "Drama" categories. During the 49 days in the data set, there were on average 731 movies rented per day ($SD = 315$), of which 108

(15%) are Comedies (*SD* = 54) and 100 (14%) are dramas (*SD* = 48). The analysis conducted with this short data set is represented in Table 3; each column represents a set of coefficient estimates in a regression with t- or z-values in parentheses. All the regressions have the same independent variables, that is, the controls (days of week and dummies for September 2 and September 3), and the variables of interest, the dummies for 9/11 – 9/16.

The first regression is a simple linear regression (with robust standard errors) with the total number of movies rented as the dependent variable. All the control variables behave as expected; there are more rentals on weekends than on Mondays (e.g. Saturdays observe 648 or 89% rentals more than other days, all else being equal), and the Labor Day weekend Sunday and Monday resulted in more rentals than normal Sundays or Mondays. Note, that there are more rentals on Tuesdays than on other weekdays, because all Tuesdays are promotional days ("rent two for the price of one").

Turning now to the variables of interest, we observe that the total number of movies rented on 9/11 fell by 150 (17% below a normal Tuesday) and was also below normal on 9/12, but rebounded on Friday, 9/14, and was 263 above normal (24% above a normal Friday) and remained abnormally high on the Saturday.

The next two columns of Table 3 represent two regressions (performed as seemingly unrelated regressions) with the number of Comedies and Dramas rented as the dependent variables. The pattern here is different from the general pattern. On 9/11 the number of these movies rented did not fall significantly whereas on the following Friday the number of Comedies rented increased by about 50% over a normal Friday and was also very high on Saturday. For Dramas, the rebound did not happen till Saturday and

was significantly lower than for Comedies (the test comparing the coefficients for 9/14 and 9/15 between the two equations results in χ^2 (2) = 24.79, $p < 0.0001$).

The last two columns of Table 3 represent regressions (performed as seemingly unrelated regressions) with the dependent variables being the Comedies and Dramas rented as proportions of all movies rented. Interestingly, some of the controls are significant in these equations, although a priori it is not necessary, that the proportion of, say Comedies, should differ between days of week. The relative preference for Comedies and Dramas increases on the weekends and Tuesdays, which probably means, that when people rent more movies, the extra movies are more likely to be in the Comedy or Drama categories, rather than in other categories. Turning now to the variables of interest, on 9/11 the proportions of Comedies or Dramas remained normal, however on 9/12, there was a significant increase in the proportion of Comedies rented by 5.6% in absolute value (which is 37% increase over the usual 14%). The following Friday, there was also a 3.7% absolute increase in the proportion of Comedies rented (which is a 24% increase over a normal Friday). At the same time there was no influence on the Dramas rented as a proportion of total.

Discussion of the preliminary analysis. The preliminary analysis shows that on 9/11 and the following day, people rented much fewer movies than usual, but they "compensated" and probably even "overcompensated" on the Friday and Saturday to follow. We also find that in the absolute values Comedies and Dramas did not exhibit the initial fall, although they did exhibit an increase on the weekend, especially the Comedies. Finally, we find that in terms of proportions of movies rented, Sep 12 and Sep 14 exhibited a big increase in Comedies.

Thus the findings are mixed. On the one hand, because neither Comedies nor Dramas, exhibited any significant changes in the absolute numbers on and right after 9/11, we might conclude that there was no mood regulation going all. On the other hand, the results on the total movies watched reveals that perhaps people were watching TV instead of renting on 9/11 and the following day, or simply found it inappropriate to seek entertainment during a disaster. And at the same time, if movies were rented they were more likely than usual to be Comedies, so for those who did get to the rental store, some mood repair seems to be going on. As for the weekend, it could simply be exhibiting a compensation pattern that has nothing to do with mood regulation. In order to get a better picture of the data as well as achieve higher statistical power we now turn to panel data analysis, which takes advantage of the fact that customer IDs are tracked in the data.

Panel data. For the panel data analysis we reorganize the data, so that there is one observation per day per customer. Each customer who rented at least one movie in the 49 days of observation gets into the reorganized set; as a result there are a total of 10,312 customers in the data set. Each observation contains date, customer ID, the total number of movies rented, and number of Comedies and Dramas rented (by that customer, on that day). We also span the data set, so that there is one observation for each customer for each of the 49 days. If there is no such data in the original data set (where an observation is a transaction), we assign zeros to the three variables of interest (number of movies, comedies and dramas rented) in the newly created observation. As a result we get a completely balanced panel data set with a total of 505,288 observations.

Panel data analysis. On average, a customer rents .07 movies per day (SD=. 36, Max=6), out of which 0.01 (14%) are Comedies (SD=.11, Max=5), and 0.01 (14%) are

Dramas (SD=.11, Max=5). The regression analysis is summed up in Table 4. The independent variables are the same as in preliminary analysis.[3] In the first regression (with clusters for customers and robust standard errors), the dependent variable is the total number of movies that the customer rents on the day of observation.[4] The controls behave as expected. For example, on average, a customer rents 0.06 movies (or 91%) more on Saturday than on other days. As for the variables of interest, on 9/11 there was a decrease in the number of movies rented by about .015 movies (86% relative to a normal Tuesday) and was still low on 9/12 with a rebound on the following Friday and Saturday.

It seems natural to assume that the effects are coming mainly from changes in the probability to rent anything rather than in the number of movies rented per visit. To make sure, we run a second regression where the dependent variable is a dummy reflecting whether or not the customer rents any movie on the day of observation (the regression is a logit with clusters for customers and robust standard errors). For the average customer, the probability to rent a movie on any given day is 0.05 (SD = .21). The marginal effects of the independent variables on the probability to rent and the z-values are presented in the next column.[5] The results follow the pattern of the previous regression but with smaller relative effects. For example, the probability to rent dropped by about 0.7% on 9/11, which is a 9% drop relative to a normal Tuesday. This value is far from the 86% drop in the number of movies rented, so we conclude that both the probability to rent and the number of movies rented (in case there was any rent) dropped.

---

[3] Other control variables that are appropriate in this case are those reflecting the consumer's behavior in the past, for example, the number of days since last rental, type of movies rented last time, etc. We performed additional analysis with these variables included in the regressions; the results were very similar.
[4] This dependent variable is a count variable, so a poisson or a negative binomial regression is more appropriate. We performed this type of analysis as well with virtually identical results.
[5] Here and in all other cases, where the marginal effect instead of coefficients are presented, the intercept is not included. In all regressions, the intercept is significant at p <.001.

Next we look at the differential effects on the "Comedy" and "Drama" categories. The third regression is a multinomial logit (with clusters for customers and robust standard errors). The dependent variable is the type of the movie rented by the customer. It is not obvious, however, how to categorize the choices of the customers, because if a customer has at least one rental, he or she often rents more than one movie on the same occasion (M=1.5, SD=.8), and we have a problem if a person rents in more than one category on the same day. We choose to use the following categories for the customer choice:

- Comedy – on the day of observation, the customer rents in category "Comedy," but not in "Drama"
- Drama – on the day of observation, the customer rents in category "Drama," but not in "Comedy."
- Other – on the day of observation, the customer rents either in both categories "Comedy" and "Drama", or in other categories
- Nothing - on the day of observation, the customer does not rent movies[6]

The next four columns of Table 4 show the marginal effects of the independent variables on the probability to choose in each of the four categories (with z-values in parentheses). On 9/11 the probability to rent any movie (i.e. not to choose "Nothing") goes up by 0.8% (13% over a normal Tuesday). The increase in the "Nothing" category is mirrored by a decrease in the "Other" category, whereas the probabilities to rent in the "Comedy" or "Drama" do not change. A similar pattern is observed on 9/12, with a rebound on 9/14 and 9/15 in the probability to rent in general and the probability to rent

---

[6] We also performed the analysis with an additional category accounting separately for the cases of renting in both "Comedy" and "Drama." The results are virtually the same. In addition we performed a separate analysis where only one, randomly chosen, rental per customer per day was left in the data set. The results are also very close.

in comedies and dramas, with the effect for comedies being more pronounced than for dramas.

As in preliminary analysis we now turn to the relative preferences for the categories, conditional on renting any movie. Given that a consumer rents a movie, the probability that at least one of the rented movies is a Comedy is 0.19 (SD = 0.4), and the probability to rent a Drama is 0.18 (SD = 0.4). We now run a multinomial logit regression identical to the previous one, except that all the observations where DV is equal to "Nothing" are excluded from the analysis. The last three columns of Table 4 demonstrate the marginal effects of the independent variables on the probability to choose in each of the categories with the z-values in parentheses. In this analysis, 9/11 has no effect on the composition of choices. However, on 9/12 there is an increase of 6.7% to choose a Comedy (a 21% increase over a normal Tuesday). In addition, on 9/14 and 9/15 the probability to rent in this category is also elevated, whereas the Drama rentals are not influenced at all.

Discussion. Generally both the preliminary the panel data analysis follows the same pattern demonstrating that the effects of 9/11 on the movie rentals at the store chain; the pattern can be split into three components. First, on 9/11 and 9/12 fewer customers chose to rent at all, but reduction was followed by a rebound on the 9/14 and 9/15. Second, this general drop did not involve a drop in the rentals of Comedies or Dramas. Third, ignoring the non-rentals, the proportion of Comedies rented was higher than usual on 9/12, 9/14, and 9/15, whereas the proportion of Dramas rented stayed normal throughout the period.

The initial drop and the later rebound can be explained in several ways. First, it is possible that renting *any* movie is seen as an entertainment, mood-uplifting, technique and that initially customers chose to abstain from renting in line with H_3. The latter rebound can be either viewed as a switch to a mood repair strategy in line with H_2 like in Josephson *et al.* (1996), or a simple compensatory demand. At the same time, it could be the case that the initial drop is not related to mood regulation and is a result of appropriateness concerns or just being preoccupied with other media (e.g. news on TV or internet). The rebound on 9/14 and 9/15 could be viewed as a "compensating" measure, but it could also be a mood regulating measure in line with H_2, if we assume that by the weekend, it was already appropriate to rent (but not earlier).

The fact that the proportion of Comedies rented was elevated on three of the days in question seems to have a more straightforward mood regulation interpretation. Looking only at those who are renting, we are already looking at those for who the appropriateness concerns are less of an issue. These people seem to be acting in line with H_2.

General discussion

The results of the first two studies clearly point in the direction of H_3; people choose or predict others to choose movies in line with the current mood. At the same time the evidence from Study 3 is more complicated and could be interpreted as evidence for all three hypotheses. No change in the absolute number of comedies and dramas rented on 9/11-9/13 is evidence for H_1; increase in proportion of comedies rented on 9/12 and 9/14 but no changes for the proportion of dramas rented plus a larger and earlier increase

in the absolute number of comedies rented on 9/14-9/15 than the corresponding increase in absolute number of dramas rented are evidence for H_2; finally general reduction of movies rented on 9/11-9/13 is evidence for H_3 if we assume that a movie rental is a positive entertaining experience in general.

The increase in the number of movies rented on the weekend is hard to interpret because it could be a compensation for the initial reduction, and is likely to be driven by mood regulation than the initial changes in the rental pattern. For this reason, let us focus the discussion to the initial changes. We can summarize them in three simple facts; the total number of movies rented decreased, the proportion of comedies increased, the proportion of dramas did not change. As mentioned above, we could interpret the general drop as a combination of appropriateness concerns, substitution with other media and mood congruence (H_3). Let us now think of the behavior of those who did come to rent on 9/12. For these people all the three reasons for not coming into the store are not valid, or they would not have come. Thus, these participants self-select to behave in the opposite direction (H_2) and that could be why the proportion of comedies rented increased. One question that remains unanswered, what kind of movie would be chosen by those who did not rent on 9/11-9/12 had they found themselves in the store for whatever reason.

One other factor to consider is an important timing distinction between going to a movie theater and renting a movie. In the former case, you usually choose the movie you will see before arriving to the theater, whereas when renting you often decide on the movie when already in the store. Arguably, in the rental case, consumption is closer to the time of choosing, and thus, the rental choice is a somewhat "hotter" decision than

choosing the movie. Comparing the results of studies 2 and 3, we might argue that when you have to plan ahead in a more rational or cold state, you are more likely to act in line with H_3, perhaps in part for appropriateness reasons; however when the decision is hotter, mood repair (H_2) might seem as a more desirable option.

Another way to interpret the findings is to use the static and dynamic approach distinction on how affect influences behavior introduced in (Andrade, 2005). The static approach states that the affect people experience will influence their actions, but without any regulatory component. That is, "positive affect is expected to lead to a more favorable evaluation of the environment, which stimulates proactive behavior (e.g., increased consumption), whereas negative affect is expected to lead to a less favorable evaluation of the environment, which inhibits action (e.g., decreased consumption)." The dynamic approach is a mood regulation approach, stating that "people are likely to move toward the goal of a more positive affective state when they feel bad but also that they would try to protect a current affective state when they feel good"

Interestingly, Andrade (2005) finds support for both approaches. In his first experiment (as described in the introduction), participants who think that chocolate is good for good mood (mostly women) are more likely to hypothetically sample a chocolate when they are either in good or bad mood condition, than in a neutral condition, evidence of dynamic approach (trying to increase a bad mood or maintain a good one). At the same time, people who do not think chocolate is related to mood (mostly men) are less likely to sample in a bad mood than in a neutral mood, and more likely to sample in a good mood, evidence of static approach (acting when feeling good and abstaining when feeling bad).

It is quite possible that we also observe evidence of both dynamic and static behavior in our Study 3. People who do not consider movie rental to be mood related, abstain from renting on and right after 9/11; this is static approach, that is negative affect makes you less likely to take an action. At the same time, people for who movie rental is mood related prefer comedies over dramas on 9/12; this is dynamic approach, that is, people are trying to use the movie to improve mood. Of course, these arguments are largely post-hoc. In future research this possibility can be tested experimentally if we somehow influence the decision makers' believes abut whether, how, or by how much the movies to select from will influence their mood.

It is also important to point out that whichever way we interpret the findings of studies 2 and 3, any generalizations should be maid quite cautiously. On the one hand there is no doubt in how profound the effect of 9/11 was on mood and behavior. On the other hand, it was a very special event, most probably quite different from other mood influences. It was a national tragedy, widely covered by media and otherwise attention drawing – everybody knew about it. Thus there is no way to separate appropriateness issues from mood regulation. It seems quite plausible that the differential effects right after 9/11 (general decline in rentals) and on the weekend (increase in rentals and in particular in comedies) in the rental data can be attributed to the fact that early on, mood repair was overruled by appropriateness, but later mood repair prevailed. The lack of a similar pattern in the movie theater data can then be explained by the fact that movie going is a more public experience and thus the appropriateness issues are harder to overcome.

One last aspect that we would like to speculate about is what the results would be if the event were a happy one, rather than a said one, say the local baseball team won the World Series. The only study in this paper that can address this question is Study 1 (favoring H_3 – mood congruence), but of course, it has all problems of intuitions rather than real moods and behaviors as both dependent and independent measure. Study 1's results suggest that after such a happy event people will try hard to maintain their good mood and would be more likely to go for a happy movie. It is still possible, however, that they would be in less need of consolation and more tolerant of unhappy activities and their preferences for drama might increase.

Conclusion and Future Directions

Mood regulation is an important part of the market; alcohol, food, television, movies, theatrical plays, books, amusement parks, shopping in general can all be all seen as mood-relevant activities in the sense that consumer's choices within them and of them can be influenced by her mood. The literature on mood congruence and mood regulation overviewed in the beginning of this paper demonstrates that the experimental evidence is inconclusive on the direction of this influence. If a marketer wants to base her decisions about offering or highlighting a particular product on the expected predominant mood of the customer (e.g. predicted by weather or news) she cannot use the experimental findings, because they do not provide a clear answer to the question of whether a sad person is more, less or equally likely than a happy person to engage in a sad or happy activity (e.g. movie). We conducted a short survey and two field studies to resolve the

issue. The findings are not quite conclusive either. Two out of three studies (Studies 1 and 2) point into the direction of H_3 (mood matching), the third study points into the direction of H_2 (mood repair). Because the third study has more internal validity than Study 2 and definitely more external validity than Study 1, we are inclined to conclude, that absent of the appropriateness issues, when decisions are rather private, people are inclined to try to improve their bad mood with positive material rather than avoid such material for consistency reasons.

In retrospect, we think that the design of Study 1 could be easily improved by implementing two separate versions, one about movie going and the other about movie rental, to see how the privacy of the decision influences the outcome. Indeed, the results of Studies 1 and 2 can be consistent (and inconsistent with Study 3) because they both tackle the more public setting of movie going. To take the idea one step further, we can even have a third condition, where the decision at question is the choice of the movies that the decision maker already has available at home or a pay-per-view selection. If privacy is indeed a factor in the decision, then the outcome for the rental should be between those for a theater and the at-home decision.

However, as the literature review shows, the main question of this research can hardly be resolved in a lab setting. This paper should be seen as one of the first steps of bringing it into the wild. The next steps should consist of finding and better utilizing other dependent measures (theater, library borrowings, TV-viewing patterns, etc.) and independent measures (outcome of sports events, weather, etc.) to answer the question of whether mood congruency or mood repair will prevail and under what circumstances. One way to continue is to explore data sets, similar to the ones used in the current article.

For example, one could analyze the movie going and movie rental pattern changes right after the start of the wars in Afghanistan and Iraq. In addition, one could analyze the movie rental patterns depending on local sport events (e.g. how the rentals in Cambridge MA change when Red Sox win and when they loose), election outcomes (of course assuming that Cambridge is pro-democratic), or local weather.

As far as experimental and quasi-experimental studies are concerned, it makes sense to move them out of the lab as well. For example, one could manipulate the customer's mood at a rental store or in a movie theater (or in the library, or in a mall) by offering her a chocolate, or having a confederate making some pleasant or unpleasant remark about her; and see how these mood-enhancing manipulations influence mood-relevant behavior.

To sum up, mood regulation is a topic in need of more research, especially outside the lab, and the privacy of the decision should be treated as a factor in future mood regulation studies.

References:

Andrade, E. B. (2005). Behavioral consequences of affect: Combining evaluative and regulatory mechanisms. *Journal of Consumer Research, 32,* 355-362.

Andrade, E. B., & Cohen, J. B. (2007, forthcoming). On the consumption of negative feelings. *Journal of Consumer Research.*

Bower, G. (1981). Mood and memory. *American Psychologist, 36,* 129–148.

Caruso, E. M., & Shafir, E. (2006). Now that i think about it, i'm in the mood for laughs: Decisions focused on mood. *Journal of Behavioral Decision Making, 19*(2), 155-169.

Galea, S., Ahern, J., Resnick, H., Kilpatrick, D., Bucuvalas, M., Gold, J., et al. (2002). Psychological sequelae of the september 11 terrorist attacks in new york city. *The New England Journal of Medicine, 346*(13), 982-987.

Handley, I. M., Lassiter, G. D., Nickell, E. F., & Herchenroeder, L. M. (2004). Affect and automatic mood maintenance. *Journal of Experimental Social Psychology, 40*(1), 106-112.

Isen, A. M. (1985). Asymmetry of happiness and sadness in effects on memory in normal college students: Comment on hasher, rose, zacks, sanft and doren. *Journal of Experimental Social Psychology, 114*(3), 388-391.

Isen, A. M., Shalker, T. E., Clark, M., & Karp, L. (1978). Affect, accessibility of material in memory, and behavior: A cognitive loop? *Journal of Personality and Social Psychology, 36*, 1-12.

Josephson, B. R., Singer, J. A., & Salovey, P. (1996). Mood regulation and memory: Repairing sad moods with happy memories. *Cognition and Emotion, 10*(4), 437-444.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*(2), 263-292.

Leventhal, H. (1980). Toward a comprehensive theory of emotion. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 13, pp. 139–207). New York: Academic Press.

Loewenstein, G. (2005). Hot–cold empathy gaps and medical decision making. *Health Psychology, 24*(4 (Suppl.)), S49–S56.

Loewenstein, G., & Schkade, D. (1999). Wouldn't it be nice? Predicting future feelings. In D. Kahneman, E. Diener & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology* (pp. 85-105). New York: Russell Sage Foundation.

Parrott, W. G. (1993). Beyond hedonis: Motives for inhibitng good moods and for maintaining bad moods. In D. M. Wegner & J. W. Pennebaker (Eds.), *Handbook of mental control*. Englewood Cliffs, NJ: Prentice Hall.

Van Boven, L., & Loewenstein, G. (2003). Social projection of transient visceral feelings. *Personality and Social Psychology Bulletin, 29*, 1159-1168.

Wegener, D. T., & Petty, R. E. (1994). Mood management across affective states: The hedonic contingency hypothesis. *Journal of Personality and Social Psychology. Washington, 66*(6), 1034-1048.

Zillmann, D. (1988). Mood management: Using entertainment to full advantage. In L. Donohew, H. Sypher & E. T. Higgins (Eds.), *Communication, social cognition, and affect.* (pp. 147-172). Hillsdale, NJ: Erlbaum.

Zillmann, D., Hezel, R. T., & Medoff, N. J. (1980). The effects of affective states on selective exposure to televised enrertainment fare. *Journal of Applied Social Psychology, 10*, 323-339.
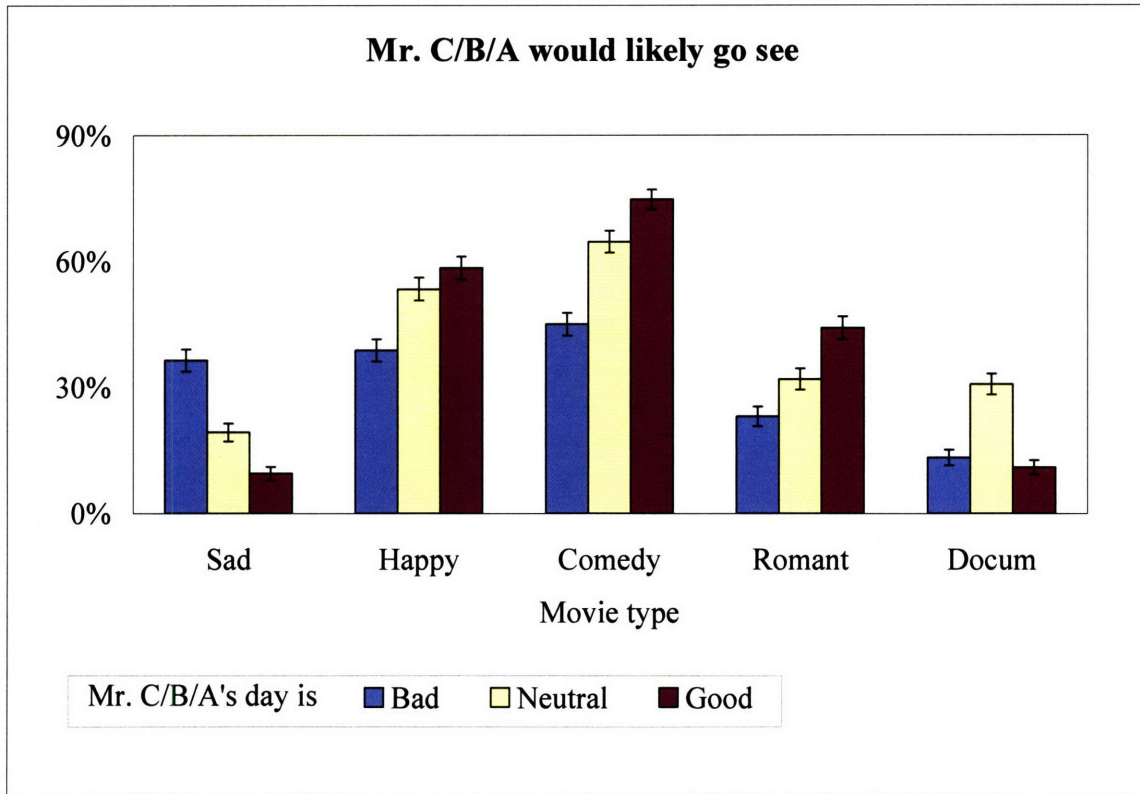
Figure 1

Mr. C/B/A would likely go see

Table 1

Indep. Vars	Sad Movie	Happy Movie	Comedy	Romantic comedy	Documentary
			Dependent Variables		
Bad Day	0.88***	-0.59***	-0.8***	-0.45***	-1.06***
	(0.17)	(0.14)	(0.14)	(0.14)	(0.18)
Good Day	-0.82***	0.2	0.48**	0.52**	-1.29***
	(0.22)	(0.14)	(0.18)	(0.12)	(0.19)

Note: z-values in parentheses; * - *p* < .05, ** - *p* < .01, *** - *p* < .001

Table 2

	Dependent Variables					
Indep. Vars	log (Gross)	log (PerTheater)	log (Gross)	log (PerTheater)	log (Gross)	log (PerTheater)
Days	-0.053***	-0.047***	-0.053***	-0.046***	-0.053***	-0.046***
	(-34.86)	(-31.6)	(-33.96)	(-30.87)	(-34.2)	(-31.04)
Tuesday	0.235***	0.22***	0.234***	0.219***	0.164**	0.149**
	(4.53)	(4.35)	(4.49)	(4.32)	(3.12)	(2.9)
Wednesday	0.038***	0.018	0.037	0.016	0.045	0.024
	(0.77)	(0.36)	(0.74)	(0.33)	(0.94)	(0.52)
Thursday	0.07	0.05	0.067	0.048	0.068	0.048
	(1.37)	(1.02)	(1.31)	(0.95)	(1.4)	(1.01)
Friday	1.257***	1.263***	1.256***	1.262***	1.276***	1.282***
	(23.8)	(23.96)	(23.61)	(23.77)	(25.28)	(25.5)
Saturday	1.682***	1.681***	1.681***	1.68***	1.713***	1.712***
	(31.79)	(31.99)	(31.56)	(31.78)	(33.77)	(34.03)
Sunday	1.243***	1.239***	1.242***	1.238***	1.26***	1.257***
	(25.68)	(26.34)	(25.54)	(26.21)	(27.1)	(27.72)
Sep2	0.28***	0.249***	0.278***	0.247***	0.265***	0.234***
	(4.74)	(4.04)	(4.45)	(3.8)	(4.31)	(3.61)
Sep3	1.231***	1.225***	1.228***	1.222***	1.233***	1.226***
	(17.78)	(14.7)	(17.1)	(14.26)	(17.36)	(14.46)
Sep11	-0.705***	-0.703***	-0.233	-0.277		
	(-9.58)	(-9.32)	(-1.34)	(-1.33)		
Sep12	-0.159*	-0.156*	0.186	0.143		
	(-2.55)	(-2.28)	(1.51)	(0.82)		
Sep13	-0.206**	-0.211**	0.176	0.126		
	(-3.12)	(-2.99)	(1.72)	(0.82)		
Sep14	-0.038	-0.032	0.207	0.295*		
	(-0.64)	(-0.49)	(1.47)	(2.03)		
Sep15	0.047	0.053	0.198	0.286*		
	(0.85)	(0.87)	(1.61)	(2.2)		
Sep16	-0.066	-0.063	0.15	0.234*		
	(-1.36)	(-1.15)	(1.42)	(2.16)		
Sep11to16					0.193	0.216*
					(1.93)	(2.09)
Sep11_Hap			-0.105**	-0.095*		
			(-3.19)	(-2.21)		
Sep12_Hap			-0.077**	-0.067*		
			(-3.21)	(-2.03)		
Sep13_Hap			-0.085***	-0.075*		
			(-3.74)	(-2.36)		
Sep14_Hap			-0.063*	-0.084**		
			(-1.99)	(-2.61)		
Sep15_Hap			-0.038	-0.059		
			(-1.15)	(-1.66)		
Sep16_Hap			-0.055*	-0.076**		
			(-2.03)	(-2.61)		
Sep11to16_Hap					-0.091***	-0.096***
					(-3.68)	(-3.84)
Constant	14.004***	6.155***	13.993***	6.144***	13.989***	6.139***
	(271.21)	(122.46)	(268.78)	(121.3)	(271.36)	(122.32)
R-sqared	0.55	0.67	0.55	0.67	0.55	0.66

Note: t-values in parentheses; * - $p < .05$, ** - $p < .01$, *** - $p < .001$

Table 3

	Regr1	Reg 2 (SUR)		Reg 3 (SUR)	
	TotalNumber	Comedy	Drama	PropComedy	PropDrama
Tuesday	448.83***	69.17***	69.33***	0.009	0.019**
	(6.09)	(6.75)	(8.07)	(1.13)	(2.87)
Wednesday	-25	-1.67	-4.33	0.004	-0.002
	(-0.51)	(-0.16)	(-0.5)	(0.49)	(-0.28)
Thursday	17.33	1.33	7.5	-0.002	0.012
	(0.35)	(0.13)	(0.87)	(-0.2)	(1.8)
Friday	655***	106.83***	106.17***	0.018*	0.025***
	(8.51)	(10.42)	(12.35)	(2.19)	(3.83)
Saturday	648.33***	98.17***	99.67***	0.012	0.021**
	(9.25)	(9.58)	(11.6)	(1.43)	(3.19)
Sunday	227.67**	43.6***	40.87***	0.019*	0.023**
	(2.96)	(4.06)	(4.53)	(2.12)	(3.44)
Sep2	327***	27.4	51.8**	-0.021	0.004
	(5.18)	(1.41)	(3.18)	(-1.35)	(0.3)
Sep3	239.67***	37	51.67**	0.008	0.035**
	(5.45)	(1.93)	(3.21)	(0.52)	(2.85)
Sep11	-150.17*	-11.17	-9.67	0.016	0.015
	(-2.54)	(-0.58)	(-0.6)	(1.02)	(1.2)
Sep12	-60.33**	10.67	-8	0.056***	-0.004
	(-2.79)	(0.56)	(-0.5)	(3.61)	(-0.35)
Sep13	17.33	13.67	12.17	0.025	0.022
	(0.75)	(0.71)	(0.76)	(1.6)	(1.81)
Sep14	277.67***	92.17***	30.5	0.037*	-0.006
	(4.4)	(4.81)	(1.9)	(2.34)	(-0.5)
Sep15	263.33***	76.83***	34*	0.028	-0.002
	(4.83)	(4.01)	(2.11)	(1.79)	(-0.16)
Sep16	-9	-5.6	8.8	-0.004	0.013
	(-0.14)	(-0.29)	(0.54)	(-0.27)	(1.07)
Constant	430.33***	58***	51.33***	0.134***	0.119***
	(9.78)	(8)	(8.45)	(22.65)	(25.99)
R-squared	0.92	0.89	0.90	0.45	0.51

Note: t-values in parentheses in the first column, z- in the rest; * - $p < .05$, ** - $p < .01$, *** - $p < .001$

Table 4

	Reg. 1	Reg. 2 (logit)	Reg. 3 (multinomial logit)			
	Number of movies rented	Rented or not	Category			
			Comedy	Drama	Other	Nothing
Tuesday	0.044***	0.029***	0.007***	0.007***	0.016***	-0.029***
	(18.75)	(14.7)	(7.57)	(7.21)	(10.72)	(-14.71)
Wednesday	-0.002	-0.004**	0.000	-0.001	-0.003**	0.004**
	(-1.56)	(-2.84)	(-0.33)	(-1.4)	(-2.58)	(2.8)
Thursday	0.002	0.000	0.000	0.001	-0.001	0.000
	(1.05)	(0.2)	(0.7)	(1.25)	(-0.76)	(-0.31)
Friday	0.064***	0.051***	0.011***	0.011***	0.029***	-0.052***
	(28.23)	(23.3)	(10.26)	(10.22)	(17.54)	(-23.19)
Saturday	0.063***	0.052***	0.011***	0.011***	0.03***	-0.052***
	(28.58)	(23.65)	(10.19)	(10.19)	(18.04)	(-23.55)
Sunday	0.022***	0.022***	0.005***	0.005***	0.012***	-0.022***
	(11.6)	(11.8)	(5.92)	(5.78)	(8.4)	(-11.89)
Sep2	0.032***	0.016***	0.003*	0.004**	0.01***	-0.016***
	(7.23)	(6.44)	(2.35)	(3.44)	(4.85)	(-6.44)
Sep3	0.023***	0.026***	0.005**	0.006***	0.014***	-0.026***
	(6.83)	(7.77)	(3.31)	(3.87)	(5.59)	(-7.81)
Sep11	-0.015***	-0.007***	-0.001	0.000	-0.006***	0.008***
	(-3.85)	(-4.28)	(-1.18)	(-0.44)	(-4.41)	(4.35)
Sep12	-0.006*	-0.006*	0.001	-0.001	-0.006**	0.006*
	(-2.29)	(-2.52)	(1.08)	(-1.24)	(-3.23)	(2.43)
Sep13	0.002	0.003	0.002	0.001	0.000	-0.003
	(0.56)	(1.09)	(1.24)	(0.9)	(0.18)	(-1.13)
Sep14	0.027***	0.012***	0.004***	0.001	0.006***	-0.012***
	(5.37)	(6.08)	(4.34)	(1.84)	(4.07)	(-6.04)
Sep15	0.026***	0.012***	0.004***	0.002*	0.007***	-0.012***
	(5.25)	(6.22)	(4.21)	(1.98)	(4.34)	(-6.21)
Sep16	-0.001	-0.003	0.000	0.001	-0.003	0.003
	(-0.24)	(-1.28)	(-0.29)	(0.59)	(-1.77)	(1.29)
Constant	0.042***					
	(34.93)					
(Pseudo) R-sq	0.01	0.02	0.02			

Note: * - $p < .05$, ** - $p < .01$, *** - $p < .001$

Reported values:

Reg. 1: coefficients with t-values in parentheses

Reg. 2: marginal effects (on the probability to rent) with z-values in parentheses

Reg. 3&4: marginal effects (on the probability to choose the category) with z-values in parentheses

Table 4 *(cont.)*

| | Reg. 4 (multinomial logit) | | |
| | Category | | |
	Comedy	Drama	Other
Tuesday	0.032*	0.029*	-0.06***
	(2.46)	(2.26)	(-4)
Wednesday	0.011	-0.005	-0.006
	(0.8)	(-0.37)	(-0.39)
Thursday	0.009	0.017	-0.026
	(0.63)	(1.27)	(-1.58)
Friday	0.029*	0.032**	-0.061***
	(2.44)	(2.72)	(-4.32)
Saturday	0.023	0.029*	-0.052***
	(1.95)	(2.49)	(-3.7)
Sunday	0.023	0.02	-0.043**
	(1.73)	(1.61)	(-2.78)
Sep2	-0.006	0.02	-0.014
	(-0.35)	(1.14)	(-0.67)
Sep3	0.01	0.03	-0.041
	(0.48)	(1.4)	(-1.57)
Sep11	0.011	0.025	-0.036
	(0.54)	(1.28)	(-1.46)
Sep12	0.067*	-0.01	-0.058
	(2.17)	(-0.36)	(-1.66)
Sep13	0.023	0.012	-0.035
	(0.91)	(0.51)	(-1.18)
Sep14	0.033*	-0.01	-0.023
	(2.27)	(-0.78)	(-1.33)
Sep15	0.031*	-0.009	-0.022
	(2.09)	(-0.67)	(-1.27)
Sep16	0.005	0.024	-0.029
	(0.24)	(1.18)	(-1.16)
Constant			
(Pseudo) R-sq	0.002		

Essay 3
Measuring Liking and Wanting

Abstract

Recently neuroscientists have gathered a vast body of evidence that wanting (motivated preferences) and liking (non-motivated prferences) are not one and the same. We explore the possibility of measuring the two types of preferences uintrusiveley, in a behavioral lab. In particular we find that wanting and liking for viewing pictures of attractive people are not perfectly aligned and especially for men.

Measuring Liking and Wanting

Introduction

Measuring preferences can be seen as one of the main tasks for researchers in both Economics and Psychology, but it has turned out to be far from an easy one. All kinds of manipulations that should have no effect on preferences from the perspective of microeconomics, in reality, lead to differing elicited preferences.

For example, matching procedures and choice procedures often lead to different (actual and inferred) choices between the same two options (Fischer *et al.*, 1999; Tversky *et al.*, 1988). In a classical example from Tversky et al. (1988) participants were asked to choose between two programs reducing life casualties in traffic accidents. One program would result in 500 casualties and cost $55M, the other would result in 570 casualties and cost $12M. Some of the participants were simply asked to choose between the two programs, whereas the others did not get the $12M number and were asked to generate *a* number that would make them indifferent between the two programs. Whereas in the simple choice the majority chose the more expensive program saving more lives, in the matching task the participants suggested a price higher than $12M for the second program, thus implicitly siding with the 570/$12M program, rather than the 500/$55M one.

It has since been shown that in the simple choice more weight is usually given to the more prominent (more important or more salient) attribute, than in matching; and the effect has been labeled prominence effect. In the example above, lives saved is a definitely more important and thus more prominent attribute.

Numerous other preference reversal studies show that people reveal different preferences in joint versus separate evaluations (Hsee *et al.*, 1999), in particular, in pricing versus choice (Lichtenstein & Slovic, 1971); other studies have demonstrated preference reversals in selling versus buying procedures (Kahneman *et al.*, 1990), and numerous other cases.

Several paths have been traveled in search of the explanation of what seems to be inconsistent preferences within one individual. For example, the lack of incentive compatibility in many procedures used to elicit preferences has been blamed (especially by economists) for preference reversals. However, at least two arguments suggest the lack of incentive compatibility does not completely explain the whole set of reversal phenomena. First, although participants in many studies do not have any incentives to reveal their true preferences, in most they do not have any incentives to lie either. The effort saving strategy of random answers would not result in the particular replicable (biased) patterns observed in the literature. Second, at least some of the preference reversal phenomena have been replicated with incentive compatible procedures (Kahneman et al., 1990).

The most accepted view on preferences reversals, however, is that of constructed preferences (e.g. Payne *et al.*, 1992; Payne *et al.*, 1999) . This approach suggests that in many domains, people do not have articulate known preferences, and that in such cases people construct preferences "on the spot," and the way these preferences are elicited plays a major role in what the input for creating these preferences will be, leading to procedural and descriptive invariance failures.

The choice between the two life-saving programs is an example of procedural invariance failure; two elicitation procedures that are theoretically equivalent lead to contradictory preferences. Another famous preference reversal – resulting from framing the options in terms of either lives saved or lives lost (Tversky & Kahneman, 1981) – is an example of violation of descriptive invariance; two identical situations described differently theoretically should still lead to the same preferences.

The evidence for constructed preferences is usually based on the findings that people seem to have contradictory preferences about the same thing depending on the elicitation task and/or mindset. The constructed preferences view does not suggest, however, that one person holds different preferences about the same choice. Rather, it suggests that the person has one preference which is very fuzzy, and depending on the starting point (framing and elicitation method) we will arrive to the same core preference, but we will observe a different aspect of these preferences. That, is we approach the same preference from a different direction, which is determined by the method and mindset (see Figure 1, panel A).

For example, when participants face a simple choice between two life saving programs in Tversky *et al.* (1988), they use a simple heuristic that life is more important than money; al long as they do not have to say how much more important and what the exchange rate is, they don't. However, when participants *are* faced with the need to create an exchange rate between lives saved and money spent (matching task), as much as morally demanding it is, they do come up with such a rate. In doing so, they do not let go of their believe that life is more important than money. It is just that this preference

has a boundary, and this boundary only is revealed in the matching task. That is, in the two tasks, we observe the same preference but at different level of complexity.

However, an alternative way to interpret the evidence for constructed preferences is simply to admit that there are multiple preferences within one individual and that various elicitation procedures make different preferences more or less accessible. This view does not assume arrival at different sides of the same preference depending on the starting point and the path taken; rather it suggests arriving at different destinations; in this view, people do not construct preferences on site, but rather reveal different preferences depending on the procedure (see Figure 1, panel B).[1]

Recent advances in neurobiology suggest that multiple preferences within one individual is indeed a very likely possibility. It has been shown that there are at least two coexisting types of preferences in animals and humans, which have been dubbed "liking" and "wanting." Although traditionally liking and wanting were considered to be two sides of the same coin, that is, people were expected to want everything that they liked (or thought they would like) and they were expected to like whatever they wanted; recently neuroscientists started challenging this simple outlook (Pecina *et al.*, 2003).

Originally, the dopamine system of the brain (the "reward center") was considered to reward "animals for doing things with survival value - eating or having sex, for example," (Phillips, 2003). However, recently Kent Berridge and his colleagues have suggested that the dopamine system is responsible for desire rather than for pleasure (i.e. wanting rather than liking). For example, dopamine depleted rats refuse to eat, but when force fed they still display the proper taste reaction (Berridge & Robinson, 1998; Berridge et al., 1989). That is, these rats like the food but do not want it. Conversely,

---

[1] Of course, some mix of constructed and multiple preferences is also possible.

mutant hyperdopaminergic mice exhibit higher than control mice wanting of food but no more liking of the food (Pecina *et al.*, 2003).

Based on these findings, the dopamine system is being relabeled "desire centre," and the scientists are searching for the "real" reward center. And they do find various other circuits in the brain that they deem responsible for liking rather than wanting (Francis *et al.*, 1999; Kelley *et al.*, 2002). For review of the reward study findings see Berridge & Robinson (2003); for a "lay" overview on the pleasure and desire research in neuroscience, see Phillips (2003).

In humans, the most dramatic evidence of the dissociation of wanting and liking comes from the studies of drug addicts (Robinson & Berridge, 2000). Originally it was considered that the reasons why addicts keep seeking drugs were the pleasure of administration and the negative effects of abstinence. This view basically states that liking promotes wanting. However Robinson & Berridge (2000) provide compelling overview of research that demonstrates that neither the negative withdrawal syndromes nor the pleasure of administering drugs can fully account for drug addiction. Here is what they write about the non-existent causality between withdrawal syndromes and addiction.

> "[D]rugs that do not produce strong withdrawal syndromes, such as psychostimulants, can be highly addictive. Conversely, there are drugs that produce tolerance and withdrawal syndromes but do not support compulsive patterns of use. The latter compounds include some tricyclic antidepressants, anticholinergics and kappa opioid agonists (Jaffe, 1992). Thus, as put by Jaffe (1992) "there is little correlation between the visibility or physiological seriousness of withdrawal signs and their motivational force" (p. 9). Another problem for withdrawal-based explanations is that drug craving is often elicited by drug administration itself, in association with euphorigenic effects, at the moments when withdrawal symptoms should be at their weakest. Similarly, in animals trained to self-administer heroin, reinstatement

of drug-taking behavior following extinction is more potently elicited by a priming injection of heroin, which elicits a drug-like effect, than by the injection of an opioid antagonist, which induces withdrawal signs (Sharam *et al.*, 1996; Stewart & Wise, 1992). For human addicts, the prolonged cessation of drug use, during which time withdrawal symptoms decay, is by no means a guarantee of a cure, as relapse to compulsive use even long after withdrawal is over remains a major problem in addiction (O'Brein, 1997)."

The authors then move on to the evidence that the pleasure of administering drugs cannot account for addiction either.

"Perhaps the most compelling evidence against the idea that the ability of drugs to promote drug-taking is directly attributable to their subjective pleasurable effects comes from studies showing that subjective states are often poorly correlated with drug-taking. First, drug-taking may increase dramatically over time as an addiction develops, but the pleasure induced by a given dose of a drug is not reported to increase (see note 5 in Robinson & Berridge (1993) for a discussion of this point). If addicts craved drugs in proportion to their ability to produce pleasure, then craving late in addiction ought not to be stronger than craving after the initial drug experience – but of course that is not the case.

Secondly, after pharmacological manipulations there is often a dissociation between the reported subjective effects of cocaine and cocaine-taking behavior (Fischman, 1989; Fischman *et al.*, 1990; Fischman & Foltin, 1992; Foltin & Fischman, 1991; Haney *et al.*, 1999; Haney *et al.*, 1998). For example, Haney *et al.*, (1998) reported recently that pergolide decreased cocaine's cardiovascular effects, decreased ratings of its subjective effects ("high", "stimulated"), increased ratings of "I want cocaine", while having no effect whatsoever on cocaine self-administration behavior. Similarly, Haney *et al.* (1999) reported: "that even a 50% decrease in certain of cocaine's subjective effects by ABT-431 … did not shift cocaine self-administration" (p. 108). Along the same lines, Comer *et al.*, (1999) reported that doses of intranasal and intravenous heroin that maintained the same "breakpoint" on a progressive ratio schedule resulted in very different subjective ratings of "high".

Thirdly, it has been reported that people will work for low doses of morphine or cocaine that produce no subjective pleasure at all; doses that indeed produce no reported subjective effects of any kind (Fischman, 1989; Fischman & Foltin, 1992; Lamb *et al.*, 1991). In summarizing their findings Lamb *et al.* (1991) concluded: "The reinforcing effects of morphine can occur in the absence of self-reported subjective effects and thus, do not appear to be causally related to drug-liking or euphoria" (p. 1172). These kinds of data are very important because they establish that the motivation to take drugs *(drug wanting)* is not always directly attributable to the subjective pleasurable effects of drugs *(drug liking)*, and it is possible this is especially true in addicts. That is, one must consider the possibility that in addicts the subjective pleasurable effects and the motivational effects of drugs are merely *correlated* effects. They occur together most of the time, but they can be dissociated and there is no necessary causal relationship between them."

Thus, Robinson & Berridge (2000) provide ample support that at least in drug addicts, liking and wanting are often severely dissociated. The wanting of drug in an addict might be completely dissociated from either the liking of administering it or the disliking of the withdrawal symptoms. But do these dissociations happen in more mundane settings? Well... Has it ever happened to you that you worked hard to get a good grade in a test to be mildly relived (rather than being ecstatic) upon receiving the grade? Have you ever spent an enormous amount of time choosing a piece of furniture or a vase to hardly even notice it once it settles in your living room? Have you ever been reluctant to spend time and money on buying a new gadget but immensely enjoyed the new toy upon receiving it as a gift? Of course, such occasions could be explained by mistaken predictions, but at the same time, we could interpret them as some divergence of motivated and non-motivated preferences, or liking and wanting.

In this paper we discuss measures that elicit two types of preferences that seem very much akin to wanting and liking, and at the same time, can be easily implemented in

a computer lab. The methods would also rate zero on intrusiveness in comparison with any neuroimaging study. In addition, the method that we suggest for eliciting wanting is incentive compatible, and both measures are continuous.

Method

In order to show that liking and wanting are distinct, it is first important to find a set of stimuli and a group of people for who liking and wanting for the stimuli might diverge. The main difference between liking and wanting is motivation, thus we need to come up either with a set of stimuli that are very motivating for some people but less for others, or two sets of stimuli such that one set is very motivating and the other is not.

One of the strongest human motivations, sexual drive, suits our purposes very well because there are two natural groups of people with different levels of motivation in respect to sexual drive, men and women. According to Baumeister & Vohs (2004), most heterosexual sexual activity is a marketplace where women are endowed with a valuable resource, sex, and trade it for money, and commitment, whereas male sexuality has no value. At the same time female sexual drive is relatively low and only interferes with the woman's bargaining power for the fair price. If this theory is too harsh to believe, here are some empirical findings from (Baumeister *et al.*, 2001, cited in Baumeister & Vohs, 2004)

> "On every measure, men were found to display greater sexual motivation than women. Specifically, men think about sex more often, have more frequent fantasies, are more frequently aroused, desire sex more often (both early and late in relationships and outside of relationships), desire a higher number of sex partners, masturbate more frequently, are less willing to forego sex and are less successful at celibacy (even when celibacy is supported by personal religious

commitments), enjoy a greater variety of sexual practices, take more risks and expend more resources to obtain sex, initiate more goal directed behavior to get sex, refuse sex less often, commence sexual activity sooner after puberty, have more permissive and positive attitudes toward most sexual behaviors, are less prone to report a lack of sexual desire, and rate their sex drives as stronger than women. No findings indicated that women had a stronger sex drive than men on any measure."

Thus, we should expect men to be more motivated than women to acquire even mild sexual content. At the same time on the motivation-free measures of liking, it is possible to find very little differences in men's and women's preferences for the same content. One such possible content is images of people. Whereas men should be highly motivated to watch pictures of women, women might be much less motivated to watch the pictures of men, whereas mere liking might not differ much between sexes.

Second, we need to establish two procedures that extract preferences for the same type of content such that one is rather passive and does not reflect motivation, and the other captures motivation, for example, makes one exert effort in order to get the content. We suggest that simple rating task is a natural measure for passive, non-motivated type of preferences (liking). As for the task high on motivation we will use the keypress task introduced in Aharon *et al.* (2001, also see Elman et al., 2005). This method requires participants to exert effort in order to experience more of the desired and less of the undesired stimuli. That is, participants view a set of pictures one by one for a fixed total amount of time, but they can prolong the viewing time of any given picture by alternating two keys on the keyboard and decrease viewing time by alternating two other keys. We speculate that the keypress task allows us to measure preferences that are motivated and

driven, rather than preferences that represent mere liking because without motivation why would one spend any effort on clicking?

We conduct four experiments to verify that the two measures actually measure liking and wanting, as well establishing some properties of wanting and liking for a certain type of stimuli, pictures of faces. In Experiment 1 we compare how men and women perform on the two measures. We find that even thought men can appreciate the beauty of both male and female faces, their driven preferences do not mimic mere liking. That is, men will only exert effort to see the pictures of beautiful women, but not of beautiful men. Contrary to our predictions women do not exert less effort to see beautiful male faces than men do for beautiful female faces; women are equally motivated to see beautiful female and male faces. We hypothesize that, for women, wanting has a social component together with the sexual one, whereas only the sexual one seems to be the driver for men with these stimuli. We test this conjecture in Experiment 2 and the data support it. In Experiment 3 we try to establish whether the results replicate with a set of stimuli that are devoid of any driving force (sexual or social), the images of buildings. We find that most of the distinction between male and female "wanting" for the stimuli is gone. Finally, in Experiment 4 we test some alternative explanations of the findings.

Experiment 1

Motivation. Because the main difference between wanting and liking is the presence and absence of motivation, we need to use two measures that differ in the same way; in addition we might want to use two sets of stimuli that differ on motivation or two populations that differ on motivation with respect to the stimuli. As mentioned above,

the task of finding two groups of people differing on motivation can be solved by using men and women and some type of (even mild) sexual content. According to (Baumeister et al., 2001) we should expect (heterosexual) men to be much more motivated to acquire such content than women. Following (Aharon et al., 2001) we use pictures of male and female faces as stimuli and predict that men will be much more driven to see pictures of attractive women than women will be motivated to see pictures of attractive men. As non-driving stimuli we use the pictures of unattractive people. In addition, for men pictures of men and for women pictures of women should have no motivation (or have negative motivation).

The task that does involve motivation is a task that requires effort to make a stimulus stay longer or shorter on screen. In particular, we use the keypress task from (Aharon et al., 2001) in which participants watch pictures one at a time for a total fixed duration of time but have some control over how long each particular picture stays on screen. Participants have to press two keys in alternation on the keyboard in order for a particular picture to stay longer on the screen and two other keys in order to reduce the viewing time of the picture. The motivation-free task that we use is a simple rating of attractiveness.

Thus we have a task that is motivational (the keypress task) and one that is not (ratings); we have a motivational stimuli (beautiful faces) and a stimuli with no or negative motivation (average faces); we also have a population-stimuli pair (men and pictures of beautiful women) that we predict to show more motivation than another pair (women and pictures of beautiful men). We predict that with the non-motivational task, participants will be rather "objective" and will rate beautiful faces higher than average

ones independent of the gender of either the participant or the target. With the motivational task, we predict that participants will exert effort to see the beautiful faces of opposite gender, and men will exert more effort to see faces of beautiful women than women to see faces of beautiful men.

Method. We asked 17 women and 18 men to rate 80 pictures of male and female faces on a scale from 1 (very unattractive) to 10 (very attractive). Of the 80 pictures, 40 were pictures of women's faces, and 40 of men's faces. Twenty of the 40 male pictures were pictures of attractive men (according to a pretest), and the other 20 – of average men, and similarly for female pictures. Before or after the ratings (counterbalanced), participants were asked to work for 10 minutes on the following task. They would see a picture (same pictures were used) that would stay on screen by default for 8 sec. Subjects could press keys "x" and "z" (in alternation) on the keyboard to increase the duration of the picture staying on the screen. Similarly, they could alternate between keys "n" and "m" to reduce the amount of time that the picture remained on the screen.[2] Once the picture disappeared, a new one would appear. A "slider" to the left of the picture indicated remaining viewing time and changed depending on the keypressing. In case

[2] Each keypress increased or decreased the total viewing time according to the following formula: NewTotalTime = OldTotalTime + (ExtremeTime – OldTotalTime)/K, where ExtremeTime was 0s for keypresses reducing the viewing time, ExtremeTime was 16 s for keypresses increasing the viewing time, and K was a scaling constant set to 40.

123
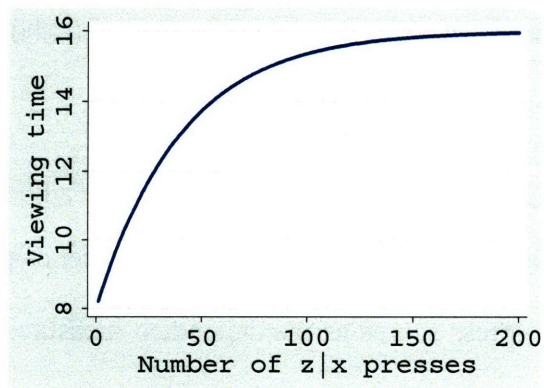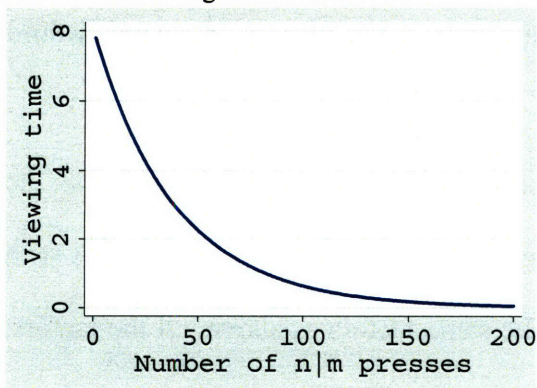
participants saw all the pictures before the 10 min expired, they started watching the pictures again. The pictures were appearing in a random order, created independently for each participant.

Dependent Measures. We use the following procedure to create each participant's keypress score and rating score for each type of pictures. For each viewing of a picture by a participant in the keypress task, we add a point to her score of that viewing for each key-press that a participants made in order to view the picture more, and we subtract one point for each key-press that she made in order to view the picture less. We then take a simple average of the scores of all viewings of the picture by the participant to create the participant's keypress score for the picture. Finally we collapse data within participants, so that there is only one rating and one keypress score observation for each combination of subject, target gender, and target beauty (i.e. four observations per participant).

Results. As predicted, participants rate beautiful pictures higher than the average pictures independent of the gender of the participant and the target (see

Figure 2). In a 2 [Subject Gender] x 2 [Target Gender] x 2[Target Beauty] ANOVA, with Subject Gender as a between subject factor, and the other two as within subject factors, and the ratings as the dependent variable, the only significant factor is Target Beauty [$F(1, 99) = 469.18$, $p<0.0001$], none of the other factors, or second or third order interactions is significant.

When clicking to make the pictures stay for longer or shorter times, the participants demonstrate a different set of preferences (see Figure 3). In an ANOVA with keypress scores as the dependent measure, and the same factors as above, all the factors,

second-order and third-order interactions are significant. To tease out the effects, we run separate ANOVAs for average and beautiful targets. For average faces, neither the gender of the target, nor the gender of the subject, nor the interaction is significant. However, for beautiful faces, all three are significant. To tease effects further apart, we perform t-tests and find that men exert more effort in order to see beautiful female pictures than beautiful male pictures ($t(17) = 5.88$, $p < 0.0001$), whereas women exert the same amount of effort to see the beautiful men and women ($t(16) = 1.18$, $p = 0.26$). Or alternatively, men and women exert the same amount of effort to see beautiful women ($t(33) = 0.02$, $p = 0.98$), but men spend much less effort to see beautiful men ($t(33) = 5.78$, $p < 0.0001$). Contrary to our predictions, there is no difference between the keypress scores of men clicking to see beautiful women and women clicking to see beautiful men ($t(33) = 0.51$, $p = 0.61$)

Given such statistics, women's liking and wanting patterns seemed to be in line, whereas for men, the two measures seemed to diverge. We compare men's and women's correlations of rankings and keypress scores, and we find that whereas for women the correlation is 0.79 ($p < 0.0001$), for men it is 0.21 ($p = 0.13$), and the latter is significantly lower ($z = 5.05$, $p < 0.0001$).[3]

Discussion. Participants rate beautiful faces higher than average faces independent of gender of the subject or target. However, when effort is required to prolong or reduce the duration of viewing the pictures, only women's actions are in line

---

[3] Additional analysis was performed with the original data (80 observations per participant, that is, one observation per picture per participant). We calculated rank-order (Spearman) correlations between rating and keypress for each participant. For women, mean correlation was 0.74 (SD =0.16, Min = 0.34, Max = 0. 93), all ps below 0.01. For men, the rank-order correlations were significantly lower (M = 0.47, SD = 0.25, Min = -0.08, Max = 0.79; compared to women $t(29) = 3.87$, p < 0.001), the mean p-value was 0.08 (SD = 0.23, Min < 0.0001, Max = 0.89). Thus, not only are men not interested in viewing beautiful men, even on the rank basis the two types of men's preferences are not aligned. It could imply, that men look at different aspects of the face when rating than when keypressing, rather than looking at the same aspects but valuing them differently.

with their ratings; they click more to see beautiful faces than average faces independent of the target gender. At the same time, men are only interested in seeing beautiful women, and click much less for all men's faces and average women's. Thus, even though there seems to be consensus on who is beautiful and who is not, independent of the gender of the viewer or the targets, women are interested in seeing all beautiful people, whereas men are only interested in seeing beautiful women.

The most important outcome of this experiment is that if we believe that the measures we use reflect liking and wanting, then liking and wanting for these simple stimuli – faces of people – diverge. In particular, the two measures diverge for men. We cannot interpret this finding in terms of constructed preferences, because if the two procedures influence the side of the preferences that is revealed by participants, there should not be any difference in the outcome for men and women. Similarly, lack of incentives cannot explain the results, because the keypress task is incentive compatible and but reveals different preferences from two subsamples. Thus, it seems quite likely that the two tasks measure two types of preferences for the same person.

The second finding is that women's driven preferences for beautiful faces are equally strong for beautiful faces of men and women. Although we expected this difference to be smaller than the corresponding difference for men, we did not expect it to be zero. There are at least three potential explanations for this finding.

One possibility is that the sample in the study included some nontrivial number of gay participants, and more gay women than gay men. Although in general, we have no

reason to believe that the proportion of gay women in our sample is higher than in general population,[4] this hypothesis is worth crossing out.[5]

Another possible explanation is that our stimuli elicit different drives from men and women. If men's reaction to the images is purely sexually driven, whereas women' reaction is socially driven (in addition to being sexual), the observed results should occur. If we think of the popular culture, what kind of pictures are men and women driven to? Anecdotal evidence suggests that pictures of beautiful females are just as likely to be seen in a female oriented as in a male oriented magazine. In fact, in a fashion magazine, whose audience is mostly women, most images could be those of women. On the contrary, any magazine with a predominantly male audience is unlikely to have even half of its images to be of men.

These observations suggest that men's desire for images of beautiful women rather than beautiful men, is brought about by the sexual drive; whereas in women there is a social driving force on top of the natural sex drive. As a result we might find that for women the desire to watch images of beautiful women is stronger than men's desire to watch images of beautiful men. As part of the future research, we can conduct an experiment where the keypress task is preceded by a priming stage, where participants are primed by some material related to either sexual or social interactions. If the

---

[4] For example (Smith, 2003) states that "the gay and lesbian communities have long adopted 10% as the portion of the population that is homosexual." His own estimates (and those of the studies he cites) are even lower.

[5] To see if there were indeed two types of women in the distribution, we regressed the keypresses and the ratings on target gender, beauty and interaction, separately for men and women, and analyzed the residuals. The histograms did not show any major split into two groups for either men or women, for either keypresses' or ratings' residuals. We also compared the variances of the residuals between genders. For the ratings, there was no significant difference ($F(67, 71) = 0.93$, $p = 0.77$), but the variance of women's keypress residuals was marginally higher than that for men ($F(67, 71) = 1.57$, $p = 0.06$). This later finding can be seen as evidence in support of a more diverse women's population as far as the motivated preferences are concerned. Thus, sexual orientation might be indeed a factor in the findings.

explanation above is true, it can be the case that socially primed men will become more interested in viewing pictures of beautiful male faces, and sexually primed women will become less interested in viewing beautiful women.

Finally it is possible that what we find is a very general pattern, that is, no matter what the stimuli are, for men their likings are less predictive of their wantings than for women.

In order to tease apart these three possibilities, we conduct two additional experiments. In order to pursue the possibility that social aspect is a part of women's wanting in relation to images of faces (as well as the minor goal of the possibility that the sexual orientation mix was biased in Experiment 1) we run Experiment 2, where we take record of the participants sexual orientation, and intentionally target half-gay-half-straight sample.

In Experiment 3 we use a set of stimuli that should have no drive for most people, images of buildings. If it is the case that for women, wanting and liking are better correlated than for men, the effect will be replicated even with these "driveless" stimuli; however it seems more plausible that for this type of stimuli liking and wanting will be perfectly aligned for all people.

Experiment 2

The results of Experiments 1 suggest that whereas for men the driving force in "wanting" is purely sexual, for women, it has a social aspect as well, because women are as motivated to see pictures of beautiful women as of beautiful men. However, it could be

the case that our sample in Experiment 1 includes some proportion of gay women (and no or fewer gay men). If this is the case, then the finding could be driven by the fact that women, just like men, are only motivated to see images of the gender of interest, and the social aspect has no weight in the drive. In order to see whether this is the case, in Experiment 2, we replicate the keypress part of Experiment 1, but this time, ask participants to identify their sexual preferences; we also target a half-straight-half-gay sample.

Another possibility is that it is not the gender of participants that drives the difference, but rather the preferred gender of sexual interest. It could be that it is not women who are equally motivated to view pictures of beautiful men and women, but it could be that those whose primary sexual interest is women (that is straight men and gay women) are equally motivated to watch the pictures of beautiful men and women.

Thus, we ask the question of whether the gender or sexual preferences drive the diversion of liking and wanting of watching images of faces. For example, when clicking to see the pictures, are gay men closer to straight men, because they share gender or are they closer to straight women because they share the gender of interest?

Method. Thirty men (15 gay and 15 straight) and 27 women (12 gay and 15 straight) participated in the keypressing task identical to the one in Experiment 1.

Results. Just as in Experiment 1, we collapse data, so that there is one observation of the keypress score for each subject, target gender, target beauty combination (in total, four observations per participant). The means of the keypress scores can be seen in Figure 4. We subject the scores to a 2 [Subject Gender] x 2 [Subject Sexual Preferences] x 2 [Target Gender] x 2 [Target Beauty] ANOVA, with the first two

factors between subject and the last two within subject. The main effect of target beauty is significant, participants exert more effort to see beautiful faces than to see average faces [F (1, 159) =81.48, p < 0.0001]; there is no main effect of subject gender [F(1, 159) = 0.98, p=0.33], or sexual preferences [F (1, 159) = 1. 60, p=0.21], or the target gender [F(1, 159) = 1.32, p=0.26]. Among the interaction effects, the following are at least marginally significant, Subject Gender x Subject Sexual Preferences [F (1, 159) =5.49, p = 0.02], Subject Gender x Target Beauty [F (1, 159) =3.30, p = 0.07], Subject Sexual Preferences x Target Gender [F (1, 159) =4.54, p = 0.03] Subject Sexual Preferences x Target Beauty [F (1, 159) =6.38, p = 0.01], Target Gender x Target Beauty [F (1, 159) =6.89, p < 0.01] Subject Gender x Subject Sexual Preferences x Target Gender [F (1, 159) =13.75, p < 0.001], Subject Gender x Subject Sexual Preferences x Target Gender x Target Beauty [F (2, 159) =11.57, p < 0.00001]

We then split the data by Target Beauty and run separate ANOVAs for beautiful and average targets. For average targets, there are no main or interaction affects. For the beautiful targets, there is no main effect of Subject Gender [F (1, 53) =3.30, p = 0.20], or Target Gender [F (1, 53) =1.42, p = 0.24], but there is a marginally significant effect of Subject Sexual Preferences [F (1, 53) = 2.98, p = 0.09], and the following significant interactions, Subject Gender x Subject Sexual Preferences [F (1, 53) = 6.03, p = 0.02], Subject Sexual Preferences x Target Gender [F (1, 53) = 10.26, p < 0.01], Subject Gender x Subject Sexual Preferences x Target Gender [F (1, 53) = 26.41, p < 0.00001].

In short, whereas all subjects exert equal effort to watch beautiful female images [F (3, 53) = 0.44, p = 0.73], there is a great variation on the effort spent watching beautiful male pictures [F (3, 53) = 11.81, p < 0.00001].

As there is small amount of interest among subjects in the average faces, and no difference between the amounts in the four types of participants, we next focus our attention on the beautiful faces only. We also construct a new dependent variable. For each participant we subtract his or her keypress score for beautiful targets of the "undesired" gender from the score for the "desired" gender. That is, for gay male and straight female participants, this is the difference between their scores for beautiful male and female targets; for straight male and gay female participants, it is the difference between their average scores for beautiful female and male targets. Figure 5 demonstrates the means of this variable across the four types of participants. We intentionally show the bars in an increasing order. This seems to be the order in which the weight of the sexual motivation is increased (and that of the social motivation is decreased) in the driving ("wanting") force of the pictures for the four types of participants.

We subject the new DV ("Difference") to an ANOVA with two between subject factors, Subject Gender and Subject Sexual Preferences. The main effect of Subject Gender is significant, for women the Difference is smaller than for men ($F(1, 53) = 10.26$, $p < 0.01$). The main effect of Subject Sexual Preferences is not significant ($F(1, 53) = 0.02$, $p = 0.88$). However the interaction is ($F(1, 53) = 4.26$, $p = 0.04$), that is, gay women and men are closer on this measure than straight men and women. To sum up, the difference in amounts of effort that men exert in order to see "desired" and "non-desired" images is significantly larger than that for women, and more so for straight than for gay participants.

It seems plausible that for men the most important driving force in this task is the sexual appeal, whereas for women the social drive is also present. This diversion is much smaller for gay people (in fact, it is insignificant, t (24) = 0.74, p = 0.47).

Discussion. The findings in Experiment 2 replicate those of Experiment 1 in that the straight women are equally interested in watching the pictures of attractive men and women, whereas straight men are only interested in watching pictures of beautiful women. Thus, it is not some unobservable mix of gay and straight participants driving the findings in Experiment 1. In addition, we find that the weights that the social and sexual drives have in the wantings of gay men and women are about the same, whereas straight women put the least weight on the sexual aspect, in fact non at all (for them the Difference is indistinguishable from zero, t(14) = 0.24, p = 0.82), and straight men put the most weight on the sexual drive.

Experiment 3

Experiment 1 shows that the wanting and liking for images of human faces are much more divergent for men than for women. One possible explanation is that men have more divergence on the two measures independent of the type of stimuli used. To test for this possibility in Experiment 3, we use a set of stimuli devoid of any driving force. For these stimuli we expect complete alignment of wanting and liking independent of the participant's gender. However, if it is indeed the case, that men are more divergent than women on liking and wanting in general, we expect to see it even with these motivation-free stimuli.

The main goal of Experiment 3, however, is to check that when there is no motivation in the stimuli, liking and wanting do not diverge. In Experiment 3, our participants perform the keypress and the rating tasks with stimuli that seem to be devoid of any kind of driving force, images of buildings. Because the main difference between liking and wanting is motivation, there should be no difference in the ratings and keypresses when the stimuli is void of any motivation.

Method. We asked 17 women and 17 men to rate 16 average and 16 beautiful (according to a pretest) buildings on a scale from 1 (very unattractive) to 10 (very attractive). Participants also performed the keypress task (before or after the rating), with the same pictures for the duration of 10 min[6]

Results. As in Experiment 1 we collapse the data so that there is one observation per building beauty per subject (i.e. two observations per subject). We first subject the ratings to a 2 [Gender] x [Building Beauty] ANOVA, with Gender as a between subject factor, and Beauty as within subject factors. Unsurprisingly, participants rate beautiful buildings as more beautiful than average buildings ($F(1, 32) = 125.6$, $p < 0.0001$), whereas Gender is insignificant ($F(1, 32) = .02$, $p = 0.89$), see Figure 6. The interaction is significant ($F(1, 32) = 6.78$, $p = 0.01$), with women showing larger difference in rating the two type of buildings than men. In a similar ANOVA with keypress scores as the dependent measure, the results are similar. Participants exert more effort to see beautiful buildings than average ones ($F(1, 32) = 55.2$, $p < 0.0001$); neither the main effect of

---

[6] The default time for the picture staying on screen was 6 sec; and participants could change according to the same formulas as in Experiment 1, except that the ExtremeTime for keypresses increasing the viewing time was equal to 12.

Gender ($F(1, 32) < 0.01$, $p = 0.97$) nor the interaction ($F(1, 32) = 0.53$, $p = 0.47$) is significant, see Figure 7.

The correlation coefficient between ratings and the keypress score is 0.75 ($p < 0.0001$) for women and 0.70 ($p < 0.0001$) for men, which are not significantly different ($z = 0.03$, $p = 0.62$).

Discussion. When the stimuli are devoid of any motivation, the liking and wanting measures are well aligned for both men and women. Thus, there is no inherent propensity to be more divergent on the two measures for men than for women. This finding supports the conclusion that the difference we observe in Experiment 1 is explained by the different patterns of wanting of the beautiful faces in men and women, such that men's wanting is represented by the sexual drive only, whereas women express both sexual and social drive.

The findings also provide additional support to the fact that the outcome of Experiment 1 cannot be interpreted in terms of constructed preferences. If the different patterns of outcomes for men and women in Experiment 1 stems from the fact that the two procedures elicit different aspects of preferences in men, and the same preference in women, then the same should be true with other stimuli, including the pictures of building. However, there is no such diversion for the buildings, and thus the constructed preferences explanation fails.

Experiment 4

So far, the data suggest that ratings and the keypress task measure two types of preferences that converge for some stimuli and some people (e.g. pictures of buildings for all people) and diverge for others (pictures of beautiful male faces for men). It also seems that ratings are measuring something akin to liking and keypresses measure something akin to wanting. This is so, because we associate wanting with motivation and exerting effort – and it is in the keypress task that participants have to exert effort, – whereas liking is a type of preferences that does not necessarily reflect motivation, and simple rating does not measure motivation. However, in addition to having no motivation, liking has other properties and aspects. In particular, liking should definitely reflect the internal enjoyment that the stimulus provides. And this is where our rating task might be faulty.

We do not ask participants to say how much they enjoy the pictures. Instead, we ask them how attractive the faces and the buildings are. When we say that straight men's liking and wanting for pictures of beautiful men are not aligned, *do we imply that the men in the Experiment 1 like the beautiful male faces and would enjoy watching them did they choose to click for them*? Or, is it that they simply acknowledged the beauty of the beautiful guys without enjoying those pictures? Thus, there seem to be two possible types of liking, one – an externally oriented feeling of admitting aesthetic qualities but not reflecting internal pleasure, and the other – reflecting internal enjoyment and more predictive of wanting. So far, the instrument we use to measure liking, rating of attractiveness, seems to be the external, aesthetics appreciation type of preferences rather than a more internal, enjoyment type of liking.

To see if we can tackle the enjoyment aspect of liking and to see how it relates to the two constructs that we have discussed so far, we conduct another experiment, where

we introduce a different rating task, measuring enjoyment, rather than acknowledgement of aesthetic quality. That is, instead of asking our participants to rate the faces on attractiveness, we ask them how much they are enjoying watching the face. If this new measure is better aligned with the attractiveness ratings than with the keypress task then we are at a better position to claim that the ratings (even on the "wrong" scale) are about liking and keypresses are about wanting. If, on the other hand, the new measure is better aligned with the keypresses than with the attractiveness ratings then our original interpretation is in trouble and keypresses actually capture some combination of liking and wanting as well, whereas the attractiveness measure captures aesthetic judgments.

Method. The experiment involved three tasks, "Keypress task," "Attractiveness Ratings," and "Enjoyment Ratings." In all three tasks the participant worked with the same 80 pictures as in Experiment 1.

In the keypress task, the participant watched the pictures in a random order for a default time of 4 sec per picture, and a total time of about 10 min. The participant could alternate "z" and "x" keys to reduce the viewing time of a picture or alternate "m" and "n" keys to increase the viewing time.[7] If the participant viewed all the 80 pictures before the 10 min elapsed, the new cycle with the new random order would start.

In attractiveness rating task, the participant viewed each picture once in a random order. Underneath the picture there was a scale with a probe. The participant could move the probe to the left by pressing the "z" key, or to the right by pressing the "n" key, and as a result answer the question "How attractive do you find this person?" with the anchors "Not at all attractive" and "Very attractive."

---

[7] The exact formula is the same as in Experiment 1, except that ExtremeTime was set at 8 sec.

The enjoyment rating task was identical to the attractiveness rating task, except that the question was "How much are you enjoying viewing this face?" and the anchors were "Not at all enjoying" and "Enjoying a lot."

Each participant performed two of the three tasks, randomly selected and in random order. Each participant worked on one task in the beginning of a 1.5 hour experimental session, and on the other task later in the session, with unrelated studies in between. After performing the second task, participants were also asked about their gender and sexual preferences. In the sexual preferences question, participants had an option not to respond; if they did respond, they did so by moving a pointer on a scale anchored at "Definitely straight" and "Definitely gay."

Results and Discussion. A total of 72 men and 54 women took part in the experiment. Three men and two women did not report their sexual orientation; 5 men and one woman reported to be definitely gay; 42 men and 24 women reported to be definitely straight. The histograms of the answers on a 0 (definitely straight) to 100 (definitely gay) scale are represented in Figure 8. Because there are too few people reporting to be significantly gay we will consider only those reporting to be definitely straight and those who score above 0 on the gay scale, but below 26 – a natural breaking point in the data. In the preliminary analysis in none of the three DVs do we find any difference between these two groups of participants (straight and slightly gay), nor do we find any significant interactions between this factor and other factors. Thus, we do not differentiate between these two groups in the consequent analysis. Our final sample consists of 62 men and 48 women all of whom we may consider to be straight. Each of the 3 tasks was performed by approximately two thirds of all participants.

In the same way as before, we create a score for each target gender - target beauty combination for each subject for each task she performed (the ratings are set to be on a scale from 0 to 100). The means of the keypress scores, the ratings of attractiveness and the ratings of enjoying split by the gender of subject and the gender and the beauty of the targets are shown in Figure 9, Figure 10, Figure 11. Eyeball analysis suggests that the keypress task and the attractiveness ratings replicate the findings in Experiment 1, with the rating of enjoyment placing in between the two measures.

We perform a 2 [Subject Gender] x 2 [Target Gender] x 2 [Target Beauty] ANOVA for each of the three dependent measures. In all three ANOVAs the three way interaction is significant (Keypress: $F(1, 213) = 9.88$, $p < 0.01$, Rating Attractiveness: $F(1, 207) = 5.35$, $p = 0.02$, Rating Enjoying: $F (1, 222) = 10.59$, $p < 0.01$), as well as the main effect of Target Beauty (Keypress: $F(1, 213) = 183.46$, $p < 0.0001$, Rating Attractiveness: $F(1, 207) = 871.54$, $p < 0.0001$, Rating Enjoying: $F (1, 222) = 338.36$, $p < 0.0001$). For simplicity we perform separate analysis for Average and Beautiful Target. For the average Targets, neither the gender of the subject nor of the target, nor the interaction was significant for any of the DVs, so we will limit our analysis to the Beautiful Targets only. For the beautiful targets, the participants' and the target gender and the interaction term are significant in all three cases (all ps <0.01); the subject gender is significant for ratings of attractiveness only ($F(1, 69) = 4.58$, $p = 0.04$). Thus on all three measures participants prefer faces of men to faces of women, but the effect is driven by male participants. Additional t-tests reveal that women's scores on all three DVs do not significantly differ between beautiful male and female faces, whereas men's scores do; on all three measures men preferred women to men.

Thus, the findings of Experiment 1 do not quite replicate. This time on both the Keypress task and the Rating of attractiveness men show a preference for beautiful female faces over beautiful male faces; in Experiment 1 men do not show such a pattern when rating. However, even in Experiment 1, this difference is true directionally (although insignificant), so the greater number of participants in Experiment 4 (as well as eliminating gay subjects) suggests that the pattern is true. [8]

Moving on to the main interest of Experiment 4, how do the three measures differ? In order to directly compare the three measures we need to convert them to one scale. We do so by going to the original data with 80 observations per subject per task performed. Then for each participant and each task that she performed we create a z-score for each observation by subtracting the mean and dividing by the standard deviation (for these subject and task). We then average the z-scores so that there is only one value for each subject, task performed, target gender, and target beauty. The means of these new values for beautiful target only are represented in Figure 12. It is easy to see that women show the same scores for beautiful male and female faces on all three measures, whereas men show a large difference, with the largest in the keypresses. As in Experiment 2, for each of the three measures, we subtract the value that a person has for the "desired" gender, and the "undesired." The mean differences are presented in Figure 13. We then subject the "Differences" to a 3[Measure type] x 2[Subject Gender] ANOVA. Not surprising the Difference is larger for men than for women ($F(1, 122) = 29.35$, $p < 0.0001$), but most importantly, the main effect of Measure type ($F(2, 122) = 2.99$, $p = 0.05$)) and the interaction effect ($F(2, 122) = 2.99$, $p = 0.06$)) are marginally

---

[8] Note, another difference between the procedures in Experiment 1 and Experiment 4 is that in the latter some of the subjects performed two very similar procedures. However, the pattern remains true if we only keep those who preformed the Rating of attractiveness first.

139

significant. Thus, we can conclude that the Rating of Enjoyment falls somewhere in between our two original measures.

Another question we might ask, is what type of liking is more predictive of wanting, external or internal, that is the attractiveness or the enjoyment rating. To do this, we need to regress the keypress scores on the two ratings. This is impossible, however if the unit of observation is participant, because none of the participants performed all three tasks. Thus, we go to the original data, and make a picture a unit of observation. For each picture we observe an average keypress score, two rating scores and the beauty and gender of the face in the picture. We regress (with robust standard errors) the keypress score on the two ratings and the dummies for gender (of the target) and beauty. In this analysis only two predictors are significant, Rating of Enjoyment (β = 0.65, t = 4.34, p < 0.001) and the dummy for Female Target (β = 2.18, t = 2.94, p < 0.01). If we exclude the Rating of Enjoyment from the regressors, then the Rating of Attractiveness becomes significant (β = 0.58, t = 9.99, p < 0.001). Thus, not surprisingly, both types of liking are predictive of wanting, but the internal type (measured by Rating of Enjoying) is more predictive. Thus we might argue, that there is a whole continuum between non-motivated and motivated preferences, as we move from no-motivation appreciation of, say, beauty, to a heavily motivated, driven, deeply experienced wanting.

General Discussion and Conclusion.

The fact that different elicitation procedures lead to different revealed preferences is not new. However, the fact has been largely interpreted in terms of constructed

preferences, suggesting that people construct exhibited preferences on the spot and that these exhibited preferences are just different sides of some core preference that is approached from different angels by different elicitation methods. Recently neuroscience has offered evidence on truly different preferences within one individual. In particular, neuroscientists have suggested that wanting and liking are not two sides of the same coin, but rather two largely independent preferences mostly differing in presence or lack there of motivation.

In this paper we discuss two methods of eliciting preferences, simple rating, and the keypress task. The findings from Experiments 1 -3 suggest that for certain stimuli and certain participants the two methods elicit two differing types of preferences. Data from the first three experiments support the supposition that the two types of preferences exhibited are liking, the non-motivated preferences (measured by rating) and the motivated, driven preferences, wanting (measured by the keypress task). In light of these findings, it is even possible that the earlier research on preference reversals might have tackled the same two types of preferences. For example, it is possible that in the prominence effect studies, choice reflects wanting and matching reflects liking.

The fourth experiment suggests however that the differentiation between wanting and liking is not a simple dichotomy, but rather a continuum, and constructs such as the one measured by ratings of enjoyment might be somewhere on that path. Experiment 4 differs from the other experiments in that it offers an additional rating task that reflects enjoyment of the stimuli rather than basic beauty appreciation. This measure seems to fall in between the original two and be better predictive of wanting than the attractiveness rating. Another way of interpreting the results of the fourth study is to conclude that not

only there is a dichotomy between wanting and liking, there is an additional dichotomy between internal and external liking. The external liking is mere acknowledgement of the fact that (some) other people will enjoy the stimulus, whereas the internal liking reflects real enjoyment and is a better predictor of wanting than the external liking.

One last issues to discuss is the implications of this research for the contingent evaluation (CV) method (see, for example, Carson et al., 2001). CV is usually used as an alternative to revealed preferences in estimating a particular public good's value to the public. For example, in estimating the (negative) value of noise, one might compare the rent prices close to and far from an airport (the revealed preferences approach), or on can run a survey and ask the respondents how much they are eager to pay for the airport not to be built next to their home, or how much they would accept for the airport to be built (CV approach). One can draw a parallel between the rating task and CV on the one hand, and the keypress task and revealed preferences on the other hand. Ratings and CV measure merely stated preferences, not substantiated by any kind of motivation, whereas keypresses and revealed preferences measures pertain to actual behavior and thus could be seen as more solid. For this very reason, economists prefer the revealed preferences approach to CV.

However, imagine a person, who on the one hand states in a survey, that she would pay a substantial amount of money for preservation of rainforests, but on the other hand, buys the furniture produced of rainforest wood because it is cheaper. Which of the two contradictory preferences should the policy makers take into the account? The constructed preferences approach suggests some composite of the two measures should be used. However, if the two measures (as we argue in this paper) actually measure two

different types of preferences, then taking an average does not make much sense. Instead, one might say that the preference revealed in the survey is morally superior to the one revealed on the store, because the former is not tainted by self-interest.

Here is some insight from the current research. Imagine that we were in the business of figuring out which men are beautiful and which are not, and we had the rating and the keypress task data discussed above. Which type of the men's preference for faces should we take into the account, rating or keypress. If we only measure keypresses we do not feel that the outcome would reflect what people actually consider beauty. Revealed preferences approach can be suffering from this drawback as well. We argue that because CV reflects no self-interest, it is morally superior and is the one that should use. Thus, the findings in this paper suggest that the non-motivated preferences are not necessarily inferior to the motivated ones and should not be completely discarded. Whereas economists suggest that CV should be a supplementary measure in general, or only used when there is no market for measuring revealed preferences (e.g. for value of a wild life habitat where public will not be admitted), we suggest that it could be measuring something completely different from revealed preferences and something which might in fact be morally superior to motivated preferences.

References

Aharon, I., Etcoff, N., Ariely, D., Chabris, C. F., O'Connor, E., & Breiter, H. C. (2001). Beautiful faces have variable reward value: Fmri and behavioral evidence. *Neuron, 32*, 537-551.

Baumeister, R. F., Catanese, K. R., & Vohs, K. D. (2001). Is there a gender difference in strength of sex drive? Theoretical views, conceptual distinctions, and a review of relevant evidence. *Personality and Social Psychology Review, 5*, 242-273.

Baumeister, R. F., & Vohs, K. D. (2004). Sexual economics: Sex as female resource for
social exchange in heterosexual interactions. *Personality and Social Psychology
Review, 8*(4), 339-363.

Berridge, K. C., & Robinson, T. E. (1998). What is the role of dopamine in reward:
Hedonic impact, reward learning, or incentive salience? *28*, 309-369.

Berridge, K. C., & Robinson, T. E. (2003). Parsing reward. *Trends in Neurosciences,
26*(9), 507–513.

Berridge, K. C., Venier, I. L., & Robinson, T. E. (1989). Taste reactivity analysis of 6-
hydroxydopamine-induced aphagia: Implications for arousal and anhedonia
hypotheses of dopamine function. *Behavioral Neuroscience, 103*(1), 36-45.

Carson, R. T., Flores, N. E., & Meade, N. F. (2001). Contingent valuation: Controversies
and evidence. *Environmental and Resource Economics, 19*(2), 173-210.

Comer, S. D., Collins, E. D., Macarther, R. B., & Fischman, M. W. (1999). Comparison
of intravenous and intranasal heroin selfadministration by morphine-maintained
humans. *Psychopharmacology, 143*, 327-338.

Elman, I., Ariely, D., Mazar, N., Aharon, I., Lasko, N. B., Macklin, M. L., et al. (2005).
Probing reward function in post-traumatic stress disorder with beautiful facial
images. *Psychiatry Research, 135*, 179– 183.

Fischer, G. W., Carmon, Z., Ariely, D., & Zauberman, G. (1999). Goal-based
construction of preferences: Task goals and the prominence effect. *Management
Science, 45*(8), 1057–1075.

Fischman, M. W. (1989). Relationship between self-reported drug effects and their
reinforcing effects: Studies with stimulant drugs. *NIDA Research Monographs,
92*, 211-230.

Fischman, M. W., Foltin, R. W., Nestadt, G., & Pearlson, G. D. (1990). Effects of
desipramine maintenance on cocaine self-administration by humans. *Journal of
Pharmacology and Experimental Therapeutics, 253*, 760-770.

Fischman, M. W., & Foltin, R. W. i. (1992). Selfadministration of cocaine by humans: A
laboratory perspective. In G. R. Bock & W. J. (Eds.), *Cocaine: Scienti. C and
social dimensions.* (166 ed., pp. 165-180). Chichester, UK: Wiley.

Foltin, R. W., & Fischman, M. W. (1991). Assessment of abuse liability of stimulant
drugs in humans: A methodological survey. *Drug and Alcohol Dependence, 28*, 3-
48.

Francis, S., Rolls, E. T., Bowtell, R., McGlone, F., O'Doherty, J., Browning, A., et al.
(1999). The representation of pleasant touch in the brain and its relationship with
taste and olfactory areas. *Cognitive Neuroscience, 10*(3), 453-459.

Haney, M., Collins, E. D., Ward, A. S., Foltin, R. W., & Fischman, M. W. (1999). Effect
of a selective dopamine d1 agonist (abt-431) on smoked cocaine self-
administration in humans. *Psychopharmacology, 143*, 102-110.

Haney, M., Foltin, R. W., & Fischman, M. W. (1998). Effects of pergolide on
intravenous cocaine self-administration in men and women.
Psychopharmacology, 137, 15-24.

Hsee, C. K., Loewenstein, G. F., Blount, S., & Bazerman, M. H. (1999). Preference
reversals between joint and separate evaluation of options: A review and
theoretical analysis. *Psychological Bulletin, 125*, 576-590.

Jaffe, J. H. (1992). Current concepts of addiction. In C. P. J. O'Brein, J. H. (Ed.), *Addictive states* (pp. 1-21). New York: Raven Press.

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1990). Experimental tests of the endowment effect and the coase theorem. *The Journal of Political Economy, 98*(6), 1325-1348.

Kahneman, D., & Krueger, A. B. (2006). Developments in the measurement of subjective well-being. *Journal of Economic Perspectives, 20*(1), 3–24.

Kahneman, D., & Thaler, R. H. (2006). Anomalies. Utility maximization and experienced utility. *Journal of Economic Perspectives, 20*(1), 221–234.

Kahneman, D., Wakker, P. P., & Sarin, R. (1997). Back to bentham? Explorations of experienced utility. *The Quarterly Journal of Economics, 112*(2), 375-405.

Kelley, A. E., Bakshi, V. P., Haber, S. N., Steininger, T. L., Will, M. J., & M., Z. (2002). Opioid modulation of taste hedonics within the ventral striatum. *Physiology & Behavior, 76*(3), 365-377.

Lamb, R. J., Preston, K. L., Schindler, C. W., Meisch, R. A., Davis, F., Katz, J. L., et al. (1991). The reinforcing and subjective effects of morphine in post-addicts: A dose-response study. *Journal of Pharmacology and Experimental Therapeutics, 259*, 1165-1173.

Lichtenstein, S., & Slovic, P. (1971). Reversals of preference between bids and choices in gambling decisions. *Journal of Experimental Psychology, 89*, 46-55.

O'Brein, C. P. (1997). A range of research-based pharmacotherapies for addiction. *Science, 278*, 66-70.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1992). Behavioral decision research: A constructive processing perspective. *Annual Review of Psychology, 43*, 87-131.

Payne, J. W., Bettman, J. R., & Schkade, D. A. (1999). Measuring constructed preferences: Towards a building code. *Journal of Risk and Uncertainty, 19*(1), 243-270.

Pecina, S., Cagniard, B., Berridge, K. C., Aldridge, J. W., & Zhuang, X. (2003). Hyperdopaminergic mutant mice have higher "wanting" but not "liking" for sweet rewards. *Journal of Neuroscience, 23*(28), 9395-9402.

Phillips, H. (2003, October 11). The pleasure seekers. *New Scientist*.

Robinson, T. E., & Berridge, K. C. (1993). The neural basis of drug craving: An incentive-sensitization theory of addiction. *Brain Research Reviews*(18), 247-291.

Robinson, T. E., & Berridge, K. C. (2000). The psychology and neurobiology of addiction: An incentive–sensitization view. *Addiction, 95* (Supplement 2), S91-S117.

Schreiber, C. A., & Kahneman, D. (2000). Determinants of the remembered utility of aversive sounds. *Journal of Experimental Psychology: General, 129*(1), 27-42.

Sharam, Y., Rajabi, H., & Stewart, J. (1996). Relapse to heroin-seeking in rats under opioid maintenance: The effects of stress, heroin priming and withdrawal. *Journal of Neuroscience, 16*, 1957-1963.

Smith, T. W. (2003). American sexual behavior: Trends, socio-demographic differences, and risk behavior.

Stewart, J., & Wise, R. A. (1992). Reinstatement of heroin self-administration habits: Morphine prompts and naltrexone discourages renewed responding after extinction. *Psychopharmacology, 108*, 79-84.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*, 453-458.

Tversky, A., Slovic, P., & Sattath, S. (1988). Contingent weighting in judgment and choice. *Psychological Review, 95*(3), 371-384.
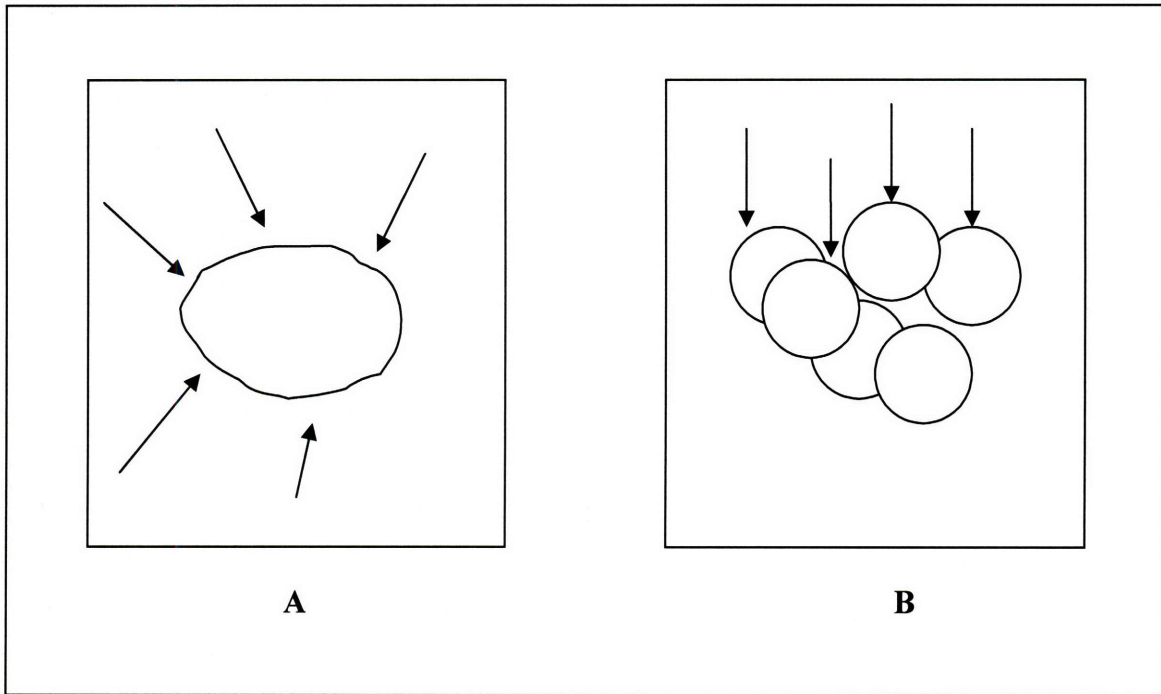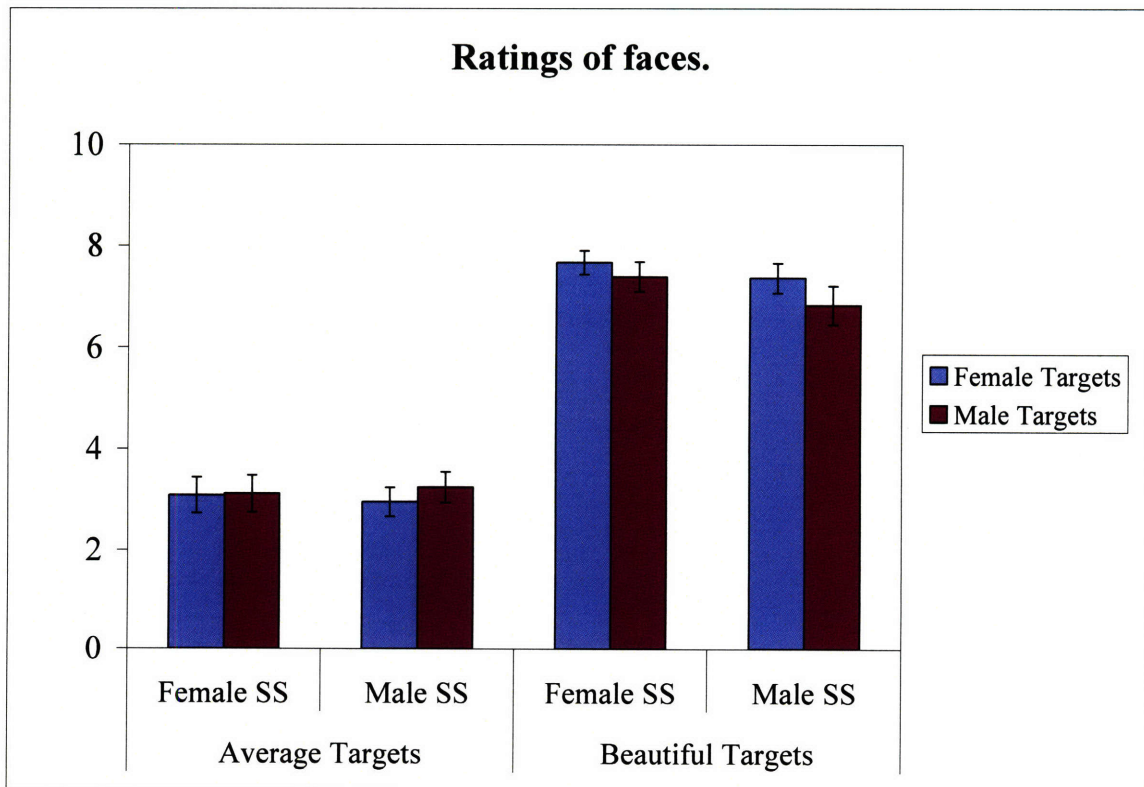
Figure 1

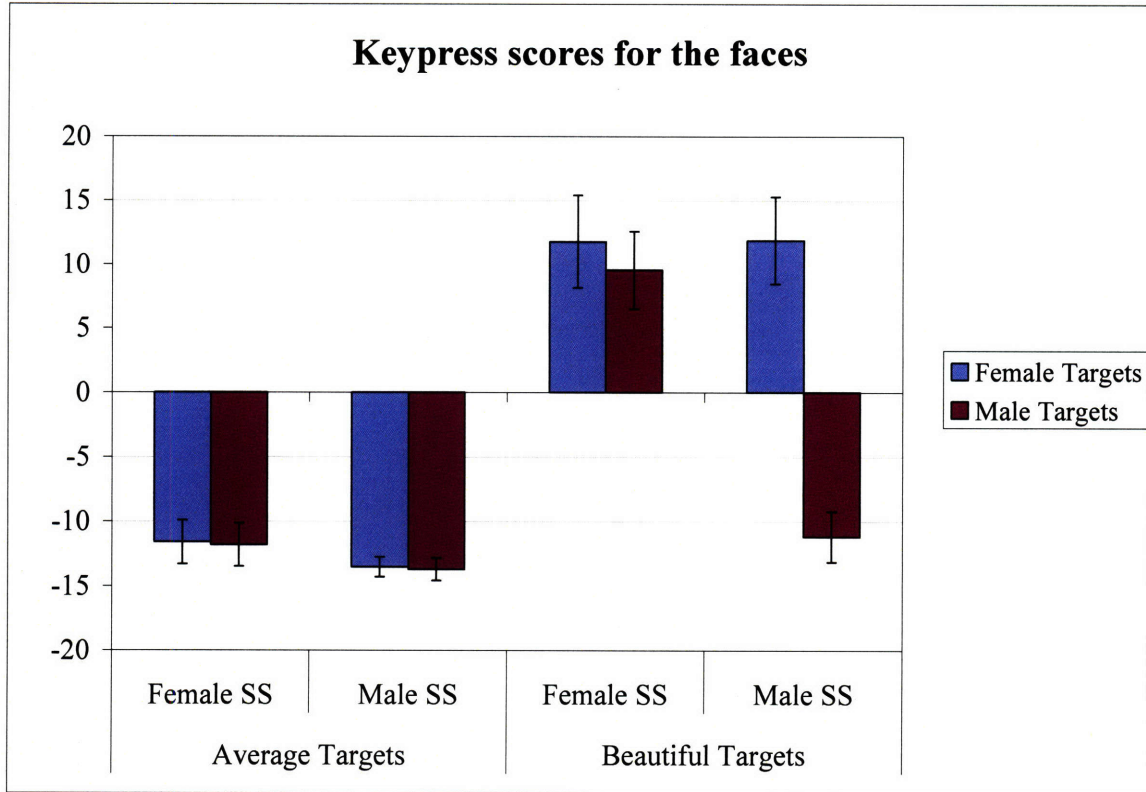A   B

Figure 2

Ratings of faces.

Figure 3

Keypress scores for the faces

Figure 4

Keypress scores for the faces

Figure 5

Difference in keypress scores between Desired and Non-desired gender targets, beautiful targets only

Figure 6

Ratings of the buildings.

151

Figure 7

Keypress scores for buildings.

Figure 8

Figure 9

Key Presses

Figure 10

Ratings of Attractiveness

Figure 11

Ratings of Enjoying Viewing

Figure 12

Z-scores, Beautiful targets only

Figure 13

Z-scores Differences between beautiful desired and undesired targets