Losses, gains, and brains: Neuroeconomics can help to answer open questions about loss aversion

Scott Rick

Department of Marketing, Ross School of Business, University of Michigan, 701 Tappan Street, Ann Arbor, MI 48109, USA

Received 19 February 2010; accepted 27 April 2010
Available online 7 June 2010

Abstract

Much is understood about loss aversion (the tendency for losses to have greater hedonic impact than comparable gains), but open questions remain. First, there is debate about whether loss aversion is best understood as the byproduct of a single system within the brain that treats losses and gains asymmetrically or the interaction of separate deliberative and emotional systems. Second, some have questioned whether loss aversion alone is the best account for the endowment effect. Alternative accounts, based on the differential focus induced by buying versus selling, the order in which buyers and sellers consider positive and negative aspects of the good, the extent to which ownership induces liking, and the desire to avoid making a bad deal, have been proposed. Third, it is unclear whether losses are actually experienced more intensely than comparable gains, or whether people simply behave as if they were. Some have argued that loss aversion is nothing more than an affective forecasting error, while others have argued that there are many situations in which losses are actually more impactful than comparable gains. This review synthesizes the insights that behavioral researchers and neuroeconomists have contributed to each debate, and highlights potential avenues for future research.

© 2010 Society for Consumer Psychology. Published by Elsevier Inc. All rights reserved.

Keywords: Loss aversion; Neuroeconomics; Decision-making; Endowment effect; Prospect theory; fMRI

The young field of neuroeconomics is evolving quickly. Its birth, in the late 1990s, was arguably inauspicious from a theoretical standpoint. Many of the early neuroeconomic studies were unapologetically atheoretical and exploratory, searching for any neural activation differences across conditions. However, as insights from exploratory studies cumulate, and methodological sophistication and diversity increase, neuroeconomics is beginning to produce findings of theoretical interest. Additionally, neuroeconomics has begun to abandon its near-exclusive reliance on functional magnetic resonance imaging (fMRI) data (which are correlational) in favor of utilizing multiple methods to better determine causality (e.g., Hsu, Bhatt, Adolphs, Tranel, & Camerer, 2005).

The potential payoff of neuroeconomics has been debated in several different fields (see Glimcher, Camerer, Fehr & Poldrack (2008) for an excellent review of the brief history of neuroeconomics). Perhaps the most visible and heated debate has been within economics, with some economists arguing for the transformative potential of neuroeconomics (Camerer, 2007; Camerer, Loewenstein, & Prelec, 2005) and others arguing to keep economics “mindless” (Gul & Pesendorfer, 2008; cf. Harrison, 2008). The spate of neuroeconomics-inspired models that have recently emerged in the economics literature (e.g., Benhabib & Bisin, 2005; Bernheim & Rangel, 2004; Brocas & Carrillo, 2008; Caplin & Dean, 2007; Fudenberg & Levine, 2006; Loewenstein & O’Donoghue, 2004) presumably argue in favor of the former perspective. The extent to which neuroeconomics can inform consumer research (Egidi, Nusbaum & Cacioppo, 2008), neuroscience (Camerer, 2008; Zak, 2004), psychology (Loewenstein, Rick & Cohen, 2008), public policy (Hoffman, 2004), and basic business practices (Shane, 2009) has also been considered.
The question of whether neuroscientific methods can illuminate topics of interest to consumer researchers has already been answered affirmatively to some extent, as illustrated by the fMRI-based studies recently published in marketing journals (e.g., Dietvorst et al., 2009; Hedgcock & Rao, 2009; Yoon, Gutchess, Feinberg, & Polk, 2006). Outside of marketing journals, neuroeconomists have investigated several topics of interest to consumer researchers, such as spending decisions (Knutson, Rick, Wimmer, Prelec, & Loewenstein, 2007), investment decisions (Kuhnen & Knutson, 2005), the placebo effect (Plassmann, O’Doherty, Shiv, & Rangel, 2008), charitable giving (Harbaugh, Mayr, & Burghart, 2007), time discounting (McClure, Laibson, Loewenstein, & Cohen, 2004), and self-control (Hare, Camerer, & Rangel, 2009).

One phenomenon that has been of particular interest to both consumer researchers (e.g., Ariely, Huber, & Wertenbroch, 2005; Camerer, 2005; Novemsky & Kahneman, 2005; Zhang & Fishbach, 2005) and neuroeconomists is loss aversion, which refers to the tendency for losses to have greater hedonic impact than comparable gains. Although much is known about loss aversion, open questions about the nature of loss aversion remain (Johnson, Gachter, & Herrmann, 2006). This review considers questions regarding the nature of loss aversion that have recently been raised (and, to varying degrees, answered) in the behavioral and neuroeconomic literatures.

Loss aversion: a brief history

The notion that losses loom larger than gains, originally formalized by Kahneman and Tversky (1979; Tversky & Kahneman, 1991; cf. Markowitz, 1952, p. 155), has proven to have tremendous explanatory power. In risky contexts, loss aversion can help to explain widespread risk aversion (e.g., Kahneman & Lovallo, 1993), as well as the St. Petersburg Paradox (Camerer, 2005). In riskless contexts, loss aversion is routinely invoked to explain the “endowment effect” (Thaler, 1980), which refers to the tendency for people to value an object more highly when they possess it than they would value the same object if they did not possess it. In addition to these basic examples, loss aversion can help to explain a wide range of phenomena, including the sunk cost fallacy, the attraction effect, the compromise effect, anticipated and experienced regret (Karides, 1994), the status quo bias (Samuelson & Zeckhauser, 1988), brand choice (Hardie, Johnson, & Fader, 1993), labor supply (Camerer, Babcock, Loewenstein, & Thaler, 1997), the equity premium puzzle (Benartzi & Thaler, 1995), organ donation decisions (Johnson & Goldstein, 2003), incumbency biases in elections (Quattrone & Tversky, 1988), and the greater tendency to cheat on one’s taxes when money is owed than when one is due for a refund (Schevanski & Shearer, 1995).

Although loss aversion has proven to be a remarkably robust phenomenon, several important moderators have been identified. Loss aversion for goods increases with duration of ownership (Strahilevitz & Loewenstein, 1998), though loss aversion has been observed even in the absence of ownership (Carmon, Wertenbroch, & Zeelenberg, 2003). Some evidence indicates that loss aversion is attenuated or reversed for money or goods already earmarked for exchange (Novemsky & Kahneman, 2005), gambles involving small amounts of money (Harnick, Van Dijk, Van Beest, & Mersmann, 2007), gambles whose probabilities are not stated explicitly, but instead inferred over time from outcome feedback (Erev, Ert, & Yechiam, 2008), unattractive goods (Brenner, Rottenstreich, Sood, & Bilgin, 2007), and goods consisting of several units (Burson, Faro, & Rottenstreich, 2009). Sellers with extensive marketplace experience (e.g., List, 2003) or influenced by incidental sadness (Lerner, Small, & Loewenstein, 2004) or positive mood (Zhang & Fishbach, 2005) also tend to exhibit reduced loss aversion. Although situational factors clearly influence loss aversion, Gachter, Johnson and Herrmann (2007) found that individuals chronically differ in the extent to which they are loss-averse and that these individual differences are fairly stable across risky and riskless contexts.

Clearly, much is understood about loss aversion. Still, open questions remain. For example, there is debate (primarily among neuroeconomists) about whether loss aversion, as well as several other decision-making phenomena, is the byproduct of a single system or the interaction of multiple systems within the brain. Some propose that losses evoke an over-learned fear response that overrides deliberative assessments, while others argue that loss aversion is adequately explained by a single system that treats gains and losses asymmetrically.

Additionally, behavioral researchers have begun to question whether loss aversion is the best account for the endowment effect. Several alternative accounts, based on the differential focus induced by buying versus selling, the order in which buyers and sellers consider positive and negative aspects of the good, the extent to which ownership induces liking, and the desire to avoid making a bad deal, have been proposed. Neuroeconomists, by contrast, have only recently begun to explore the neural correlates of the endowment effect.

Finally, both behavioral researchers and neuroeconomists have investigated whether losses are actually experienced more intensely than comparable gains, or whether people simply behave as if they were. Loss-averse decisions alone do not shed light on whether losses are actually more impactful than comparable gains, so researchers have begun to examine the experienced (dis)utility of losses and gains.

I organize this review around these ongoing debates, synthesizing the complementary contributions from behavioral and neuroeconomic research, and highlighting possible avenues for future research.

Is loss aversion a byproduct of multiple systems interacting in the brain?

Perhaps the most hotly debated topic within neuroeconomics is whether decision-making phenomena such as loss aversion

---

2 Erev and colleagues (e.g., Erev et al., 2008; Ert & Erev, 2010) question the explanatory power of loss aversion and the robustness of the phenomenon itself. For example, Ert and Erev (2010, p. 5) argue that “participants do not exhibit loss aversion in short studies, or at the beginning of long studies, but, in certain studies, they act as if they become loss averse over time.” It is difficult to reconcile this claim with previous research. Endowment effect experiments, for example, typically only elicit a single decision from owners and non-owners.
are best explained by the operation of single or multiple systems within the brain. Dual-system theories have been invoked to explain loss aversion (Ashraf, Camerer, & Loewenstein, 2005; Camerer, 2005), with the general idea being that the aversive response to loss stemming from the “hot” affective system tends to override the more objective evaluation performed by the “cool” deliberative system. Beyond loss aversion, distinct systems have been invoked to explain a variety of phenomena in consumer research (e.g., Lee, Amir, & Ariely, 2009; Strack, Werth, & Deutsch, 2005), psychology (e.g., Metcalfe & Mischel, 1999; Stanovich & West, 2000), and economics (e.g., Fudenberg & Levine, 2006; Loewenstein & O’Donoghue, 2004).3

The extent to which such theories can be reconciled with the way the brain actually operates is debatable (e.g., Johnson, 2008). Keren and Schul (2009) argue that the existing evidence from neuroscience is too ambiguous to support either the single or multiple systems perspective. Neuroeconomists have rapidly contributed much new evidence to the debate, though this evidence is not devoid of interpretational ambiguity.4

In one of the seminal papers in neuroeconomics, McClure et al. (2004) provided fMRI evidence suggestive of separate systems interacting to influence impatience. Participants in their study made a series of choices between small proximal rewards and larger delayed rewards (e.g., receive a $5 gift certificate immediately or a $10 gift certificate in 2 weeks). In some trials the proximal reward was available immediately, and in other trials it was not (i.e., participants chose between two delayed rewards). McClure et al. (2004) found that limbic and paralimbic cortical structures known to be rich in dopaminergic innervation and highly sensitive to the anticipation and experience of reward (e.g., the ventral striatum, the medial prefrontal cortex (MPFC), and the posterior cingulate cortex; see Fig. 1) were significantly more active in trials involving an immediate reward than in trials where both rewards were delayed. By contrast, fronto-parietal regions that support higher cognitive functions (e.g., the right dorsolateral prefrontal cortex) were activated about equally for both types of decisions. Moreover, McClure et al. (2004) found that when decisions involved an immediate reward, greater activity in fronto-parietal regions than in limbic regions predicted the selection of larger, delayed rewards. This difference disappeared (and flipped slightly) when participants chose immediate rewards. McClure et al. (2004) interpreted the results as a neural reflection of Laibson’s (1997) beta–delta model of quasi-hyperbolic discounting: when the patient delta (fronto-parietal) regions exerted greater influence than the impulsive beta (limbic) regions, participants tended to select the larger, delayed reward.5

However, the McClure et al. (2004) results were not universally interpreted as reflecting the interaction of separate systems in the brain (e.g., Ainslie & Monterosso, 2004; Glimcher, Kable, & Louie, 2007). Kable and Glimcher (2007) most forcefully argued against such a perspective. In their fMRI study, participants made a series of choices between receiving $20 immediately and receiving a larger delayed amount of money. Only the length of the delay and amount of the larger reward varied across trials; the immediate $20 option was held constant across trials. Kable and Glimcher (2007) found that activation in the beta regions identified by McClure et al. (2004) correlated negatively with the length of the delay and positively with both the size of the delayed reward and choices of it instead of the immediate reward. Kable and Glimcher (2007) thus proposed that activation in these beta regions encodes the

---

3 Keren and Schul (2009) utilize a computer analogy to highlight the distinction between systems and processes, concepts that are often (mis)used interchangeably. They propose that processes resemble the operations performed by software and that systems more closely resemble the hardware. Accordingly, I omit from the above list work that has proposed dual-process accounts of behavior (e.g., Chaiken & Trope, 1999).

4 A discussion of how neural activity is measured and transformed into interpretable data is beyond the scope of this review. Many relevant references are available for interested readers (e.g., Buxton, 2002; Egidi et al., 2008; Huettel, Song, & McCarthy, 2004; Logothetis, 2008; the Appendix in Yoon et al., 2006).

5 McClure, Ericson, Laibson, Loewenstein and Cohen (2007) later replicated these results with rewards that, unlike gift certificates, could actually be delivered and consumed immediately in the scanner (juice or water). See also Hariri et al. (2006).
subjective value of rewards at all delays. They reconciled their findings with the McClure et al. (2004) findings by noting that the subjective value of immediate rewards is particularly high, which could explain why McClure et al. (2004) observed preferential activation in beta regions when immediate rewards were available. The Kable and Glimcher (2007) findings strongly challenge the multiple systems perspective, though it is worth highlighting that their experimental design did not manipulate whether or not immediacy was an option (participants could always select an immediate reward). Because participants never chose between two delayed rewards, it is difficult to definitively rule out the hypothesis that the valuation process is qualitatively different when immediacy is an option.

Tom, Fox, Trepel and Poldrack (2007) also argued against a multiple systems perspective in their investigation of the neural basis of loss aversion. In their fMRI study, participants made a series of accept/reject decisions for mixed gambles (offering an equal probability of winning or losing money). If loss aversion is driven by negative affect, as some have speculated (e.g., Camerer, 2005), positive correlations between the size of the potential loss and activation in brain regions commonly associated with the experience of negative affect (such as the amygdala or anterior insula; see Fig. 1) should be observed. However, Tom et al. (2007) found that no brain regions displayed significant positive correlations with the size of the potential loss. Rather, activation in a group of regions including the dorsal and ventral striatum and ventromedial prefrontal cortex (dopaminergic targets that overlap somewhat with the beta regions discussed above), correlated positively with the size of the potential gain and negatively with the size of the potential loss. Critically, these regions demonstrated neural loss aversion: The reduction in activation in response to losses was greater than the increase in activation in response to comparable gains for nearly all participants. Additionally, the asymmetric responses to losses and gains in several regions, including the ventral striatum, strongly correlated with behavioral loss aversion. Although some question whether the current results rule out negative affect as a driver of loss aversion because of the asymmetric difficulty in picking up loss as opposed to gain signals with fMRI (Knutson & Greer, 2008), the results are more readily interpreted as evidence against the multiple systems perspective.6

De Martino, Camerer and Adolphs (2010) recently offered contradictory evidence. Their study utilized two adult participants who, due to an extremely rare genetic disease known as Urbach–Wiethe disease, had bilateral focal (selective) lesions of the amygdala. Both adults have difficulty processing fear, but otherwise have normal cognitive skills and IQ. Each participant was compared to a control group of six healthy adults matched on age, gender, income, and education. As in Tom et al. (2007), participants made a series of accept/reject decisions for mixed gambles. This was a strictly behavioral study: decisions were made outside the scanner. Non-lesion control participants exhibited typical levels of behavioral loss aversion, but neither amygdala-damaged participant exhibited loss aversion. A supplemental study examining risk attitudes over gains (e.g., choices between receiving $x$ for sure and flipping a coin to win nothing or 2x) revealed that the amygdala-damaged participants had similar levels of risk aversion as the non-lesion controls, suggesting that the difference between the two groups is restricted to their reactions to losses.

The results thus conflict with those of Tom et al. (2007), despite utilization of a similar paradigm. One possible source of the divergence is the range of gamble outcomes in the two studies. In Tom et al. (2007), possible gains ranged from +$10 to +$40, and possible losses ranged from −$5 to −$20; most gambles had a positive expected value. In De Martino et al. (2010), possible gains ranged from +$20 to +$50 and possible losses ranged from −$20 to −$50. Thus, it is possible that the losses were not severe enough in Tom et al. to elicit a detectable amygdala reaction. Unfortunately, De Martino et al. (2010) did not include gambles with small losses, which could have helped to reconcile the two studies (if the amygdala-damaged participants and non-lesion controls behaved more similarly when losses are small). De Martino et al. (2010) speculated that the amygdala may initially generate a (difficult-to-measure) aversive response that is then conveyed to the value-sensitive brain structures that correlated asymmetrically with gains and losses in Tom et al. (2007).

Behavioral research could contribute to this debate. Examining whether methods commonly used to manipulate the strength or influence of deliberative versus emotional responses, such as cognitive load (e.g., Shiv & Fedorikhin, 1999) or priming feelings versus deliberation (e.g., Hsee & Rottenstreich, 2004), moderate loss aversion would be informative.

**Is loss aversion the best explanation for the endowment effect?**

The endowment effect is most commonly defined as the difference between the minimum amount of money one is willing to accept to part with an owned good (selling price) and the maximum amount of money one is willing to pay to buy the good (buying price). Because buyers and sellers are in slightly different wealth positions, Kahneman, Knetsch and Thaler (1990) introduced a “choice” condition, whereby people make a series of choices between receiving a good and receiving an amount of money. The amount of money at which participants are indifferent between receiving it and receiving the good is the choice price. Because sellers and choosers are in identical positions (both are deciding whether to leave the lab with the money or the good), any difference between selling prices and choice prices is a pure framing effect. Typically, selling prices are significantly greater than choice prices, which tend to be only slightly and non-significantly greater than buying prices (Kahneman et al., 1990; but see Bateman, Kahneman, Munro, Starmer, & Sugden, 2005). In this section, I will use the term...
“endowment effect” to describe both the difference between selling and buying prices and the difference between selling and choice prices.

The endowment effect is typically interpreted as a manifestation of loss aversion. However, a variety of alternative explanations have also been proposed. Carmon and Ariely (2000), for example, proposed that people tend to focus on what is foregone in a potential transaction. Thus, buying prices are based largely on the pain of parting with the money, and selling prices are based largely on the pain of parting with the good. Indeed, Carmon and Ariely (2000) found that selling prices for basketball tickets were more sensitive to aspects of the game (e.g., whether it was a championship game) and that buying prices were more sensitive to monetary aspects (e.g., the retail price of the tickets).

Later, query theory (Johnson, Häubl, & Keinan, 2007) proposed that value is constructed based on a series of queries whose order is role-dependent. Sellers begin by calling to mind “value-increasing” aspects: positive aspects of the good and negative aspects of acquiring money (e.g., it would not be enough to buy something better). Buyers (or choosers) begin by calling to mind “value-decreasing” aspects: positive aspects of the money and negative aspects of the good (cf. Nayakanlkuppam & Mishra, 2005). Because aspects that are initially recalled tend to interfere with the retrieval of subsequent aspects, the order in which the queries are executed matters. Thus, owners tend to arrive at higher valuations of the good than non-owners. Indeed, Johnson et al. (2007) eliminated the endowment effect by reversing the typical order of queries (i.e., forcing sellers to begin by listing value-decreasing aspects and choosers to begin by listing value-increasing aspects). It is worth noting that query theory is not necessarily incompatible with loss aversion, but instead proposes a process that may give rise to it.

Morewedge, Shu, Gilbert and Wilson (2009), by contrast, abandon loss aversion as an explanation completely, instead proposing that it is ownership that causes the endowment effect (cf. Beggin, 1992; Beggin & Allison, 1997; Peck & Shu, 2009; Reb & Connolly, 2007; but see Barone, Shimp, & Sprott, 1997a,b). The argument is that people like what they own, regardless of how they came to own it, because owned items become associated with the self. Morewedge et al. (2009) therefore attempted to unconfound ownership and the prospect of loss, which are typically confounded in the real world and in the lab. For example, in one experiment, selling and choice prices for a mug were elicited in typical fashion. Additionally, in a novel “owner–buyer” condition, choosers who had already been endowed with a mug stated their choice price for a second, identical mug. Sellers stated significantly greater prices than non-endowed choosers, replicating the endowment effect. The choice prices of owner–buyers were statistically indistinguishable from selling prices. Although it is worth highlighting that sellers and owner–buyers were not making identical decisions (sellers were deciding whether to leave with one mug or money; owner–buyers were deciding whether to leave with two mugs or one mug plus money), the results do suggest that ownership-induced liking contributes to the endowment effect.

Another account of the endowment effect is based on transaction disutility, or the desire to avoid making a bad deal (e.g., Brown, 2005). Weaver and Frederick (2008) argue that because reference prices typically exceed the extent to which consumers value products, sellers who price goods at their subjective valuation face the prospect of selling at a loss. For example, if a mug retails for $8, but is only valued at $2, selling it at one’s valuation results in a $6 loss relative to the reference price. Buyers, by contrast, could generally offer their valuation without making a bad deal. Consistent with this account, Weaver and Frederick (2008) found that selling prices were particularly sensitive to information about reference (retail) prices (cf. Simonson & Drolet, 2004; but see Carmon & Ariely, 2000).

More exotic accounts have also been proposed. Gal (2006), for example, proposed that “a propensity to remain at the status quo” is sufficient to account for the endowment effect (i.e., sellers wish to maintain ownership and buyers wish to maintain non-ownership). However, this proposal cannot account for the basic observation that choosers (who choose between receiving the good and receiving money, and thus have no default option) routinely set choice prices that are significantly lower than selling prices. Plott and Zeiler (2005) attributed the endowment effect to “subject misconceptions” about experimental procedures. Subsequent research has revealed that dropping participants who misunderstand the task actually strengthens the endowment effect (Johnson et al., 2007). Later, Plott and Zeiler (2007) proposed a number of additional artifactual sources, such as “the regard subjects have for the experimenter” (p. 1462).

Neuroeconomists have only recently begun to examine the endowment effect, and these early studies were primarily designed to identify basic neural correlates, rather than to distinguish between different competing explanations. However, the exploratory results are still informative. Weber et al. (2007) conducted the first fMRI study of the endowment effect. Participants were endowed with digital copies of songs. On some trials, they were asked to state selling prices for their endowed songs, and on other trials, they were asked to state buying prices for different songs. Weber et al. (2007) observed significantly greater amygdala activation during the selling trials than during the buying trials. At first blush, the results appear consistent with the argument that loss aversion reflects an over-learned fear response to losses (e.g., Camerer, 2005). However, the authors did not examine whether activation in the amygdala predicted buying and selling prices, so it is unclear whether amygdala activation actually played any role in the decision-making process.

Knutson et al. (2008) conducted an fMRI study of the endowment effect, using high-value consumer goods (e.g., an iPod Shuffle). Participants were initially endowed with two goods and with money. There were three types of trials: selling, buying, and choosing. In the selling trials, participants were offered prices for their owned goods and made a series of binary selling (sell or keep) decisions. In the buying trials, participants were offered two other goods and made a series of purchase (buy or do not buy) decisions. In the choosing trials, participants were presented with two other goods and made a series of choices between receiving the good and receiving money. One choice for each of the six goods counted for real, meaning the
experiment was quite lucrative for participants. Consistent with Kahneman et al. (1990), selling prices were significantly greater than choice prices, which were slightly greater than buying prices. The behavioral results suggest that the endowment effect is not limited to small-stakes goods.

At the neural level, Knutson et al. (2008) found that activation in the MPFC in response to prices varied across conditions. Specifically, MPFC activation correlated negatively with prices in the buying and choosing conditions (though the negative correlation was particularly strong in the buying condition), and positively with prices in the selling conditions. The MPFC activation differences across choosing and selling conditions reflects a pure framing effect in the brain. The differences are also consistent with prior work demonstrating that MPFC activation is positively correlated with gains and negatively correlated with losses (e.g., Knutson, Fong, Bennett, Adams, & Hommer, 2003).

Knutson et al. (2008) also found some support for the notion that distress contributes to the endowment effect. Because participants never bought and sold the same product, two estimates of the endowment effect were computed. A “scanned” (between-subjects) estimate was equal to the difference between the selling price elicited in the scanner and the average buying and choosing price for the same product among the other subjects. A “postscan” (within-subject) estimate was the difference between hypothetical buying and selling prices for the same product elicited after the scanning session. Knutson et al. (2008) found that activation in right insula while contemplanting selling owned goods correlated positively and significantly with both estimates of the endowment effect. The insula tends to be active when anticipating distress and actually experiencing distressing events, such as unfair ultimatum game offers (Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003) or social exclusion (Eisenberger, Lieberman, & Williams, 2003). Insula activation also correlates positively with subjective ratings of distress (Masten et al., 2009). Thus, the positive correlations between insula activation and estimates of the endowment effect arguably suggest that the more distress participants feel when contemplating parting with owned goods, the more susceptible they are to the endowment effect.

De Martino, Kumaran, Holt and Dolan (2009) recently conducted an fMRI study of the endowment effect in which subjects bought and sold identical lottery tickets. The task was complicated (one-fourth of the sample had to be dropped for misunderstanding the instructions), but addressed a limitation of Knutson et al. (2008), who did not elicit within-subject endowment effects in the scanner. Consistent with Knutson et al. (2008), De Martino et al. (2009) found that insula activation correlated with the size of the endowment effect. However, it is worth noting that not all results were consistent with Knutson et al. (2008): for example, De Martino et al. also found that bilateral ventral striatum activation correlated with the size of the endowment effect, which Knutson et al. did not find.

Despite important methodological and empirical differences across papers, these exploratory fMRI studies generally support the premise that distress contributes to the endowment effect. Clearly, however, behavioral researchers have made most of the progress toward understanding what truly drives the endowment effect (or at least toward articulating the endowment effect’s multiple determinants). Thus, there is a need for more focused fMRI designs that pit some of the nuanced explanations for the endowment effect against one another.

**Are losses actually experienced more intensely than gains?**

Behavior in both risky and riskless contexts routinely reveals that individuals treat losses as if they were more impactful than comparable gains. However, loss-averse decisions alone do not shed light on whether losses are actually experienced more intensely than comparable gains or whether losses are simply predicted to be experienced more intensely than comparable gains (cf. Kahneman, Wakker, & Sarin, 1997). If losses are experienced no more intensely than comparable gains, then loss-averse behavior is misguided on both monetary and affective grounds (Kermer, Driver-Linn, Wilson, & Gilbert, 2006).

Diverse methods have been brought to bear on the question of whether losses are experienced more intensely than gains. Several studies have elicited self-reported emotional reactions to hypothetical gains and losses, with mixed results. Galanter (1990) asked subjects to fix a utility value to a focal event (e.g., assume winning a bicycle delivers 100 units of happiness) and then to assign utilities to other events (e.g., winning money in a lottery or paying a speeding ticket) based on that focal benchmark. Consistent with experienced loss aversion, the implied utility function was generally steeper for losses than for equivalent gains, though the absolute difference in slopes was small. However, Mellers, Schwartz, Ho and Ritov (1997) did not observe such an asymmetry when eliciting reactions to comparable gains and losses on a standard bipolar scale (ranging from extremely disappointed to extremely elated). McGraw et al. (in press) recently argued that losses are likely to be rated as more impactful than comparable gains only when losses and gains are compared and judged on a common scale. For example, they asked participants to rate the hedonic impact of a $200 gain and a $200 loss on a unipolar intensity scale (ranging from no effect to very large effect). When participants evaluated only the gain or only the loss, there was no evidence of experienced loss aversion. However, when participants evaluated both the gain and the loss, the loss was rated as significantly more impactful, consistent with experienced loss aversion.

One limitation of this prior work is that the “experienced” outcomes were purely hypothetical. In an article titled “Loss aversion is an affective forecasting error,” Kermer et al. (2006) addressed this limitation with two experiments examining whether real losses elicited larger changes in happiness than real gains, and whether forecasts of hedonic responses correspond with actual experience. Because the claim that loss aversion is a forecasting error is provocative, and potentially important, the data warrant heightened scrutiny. Kermer et al.’s (2006) first experiment exposed subjects to a sequence of 44 intermixed gains and losses that either resulted in a total net gain of $4 or a total net loss of $4. Participants either experienced the outcomes for real and then rated how they felt (experiencers) or simply
watched the outcomes unfold and imagined how they would feel if the outcomes were consequential (forecasters). Consistent with loss aversion, forecasters predicted that the net loss sequence would hurt about twice as much as the net gain sequence would feel good. Experiencers, however, rated the hedonic impact of the net loss sequence as about equal to the hedonic impact of the net gain sequence, which supports the argument that loss aversion is a forecasting error.

However, it is worth emphasizing that participants were not reacting to a pure gain or pure loss, but rather to a sequence of intermixed gains and losses. Thus, some people in the net loss sequence probably ended on a positive note, and some people in the net gain sequence probably ended on a negative note. (Details about the sequences were not provided, so it is unclear what percentage of net loss sequences concluded with a gain and what percentage of net gain sequences concluded with a loss.) If the next-to-last trial is the reference point by which the last trial is judged, then the appropriate comparison is arguably net loss sequences that concluded with a loss versus net gain sequences that concluded with a gain.

In Kermer et al.’s (2006) second experiment, participants were endowed with $5 and asked to imagine how their happiness would be impacted by a coin flip that would either result in a loss of $3 or a gain of an additional $5. Consistent with loss aversion, participants predicted that the loss would hurt nearly twice as much as the gain would feel good. The coin was then flipped, and participants evaluated how happy they were with the outcome. Although gains were as pleasurable as predicted, losses were less painful than predicted. Losses and gains had approximately equivalent effects on happiness (in terms of magnitude). However, given that the loss was of smaller magnitude than the gain (|$-3|<$5), the symmetric hedonic impact of the gain and loss appears to be perfectly consistent with loss aversion. Although both experiments provide novel evidence that people are too pessimistic when forecasting how much they would be bothered by losses (but see Andrade & Iyer, 2009), this experiment did not dispute the existence of experienced loss aversion. Thus, the data provide only equivocal support for their title.

Neuroeconomists have complemented this research by bringing physiological methods to bear on the question of whether losses are experienced more intensely than equivalent gains. Sokol-Hessner et al. (2009) conducted a study in which participants made a series of choices between mixed (gain/loss) gambles and sure amounts of money, and gambles were resolved immediately after they were selected. Skin conductance, a common measure of autonomic arousal, was measured immediately after each gamble was resolved. Consistent with experienced loss aversion, the average skin conductance response (SCR) per dollar lost was significantly greater than the average SCR per dollar won. Other subjects were encouraged to bracket their decisions broadly, focusing on “how your investments are performing as a whole,” similar to a method previously shown to reduce behavioral loss aversion (Gneezy & Potters, 1997; Thaler, Tversky, Kahneman, & Schwartz, 1997). Consistent with this prior work, Sokol-Hessner et al. (2009) found that “thinking like a trader” significantly reduced behavioral loss aversion, and eliminated the average SCR difference between dollars lost and dollars won.

Similarly, Hochman and Yechiam (2011) found that pupil dilation and heart rate was significantly greater in response to losses than to equivalent gains, though, in contrast with Sokol-Hessner et al. (2009), these autonomic responses did not predict gamble choices. However, it is worth highlighting that subjects in Hochman and Yechiam’s paradigm chose between two mixed gambles (e.g., gain 2x or lose 2x versus gain x or lose x). The lack of correlation between autonomic responses and gamble choices is thus less surprising, given that subjects could not protect themselves from the possibility of loss.

Only a small amount of neuroscientific research has investigated the relative hedonic impact of comparable gains and losses, and the experimental designs often did not permit definitive conclusions. For example, Gehring and Willoughby (2002) used electroencephalography (which involves attaching electrodes to the scalp) to study neural reactions to gains and losses, but soon after the experiment began subjects learned that only the gains would actually influence their earnings (p. 2282). Breiter, Aharon, Kahneman, Dale and Shizgal (2001) used fMRI to measure responses to actual gains and losses, but none of the gains and losses were of equal magnitude. Rutledge and Glimcher (2009) recently conducted an fMRI study in which participants experienced a series of gains and losses, and found greater activation in several regions (e.g., amygdala) in response to losses than to gains of equal magnitude, consistent with experienced loss aversion.

**General discussion**

There is much we know about loss aversion, but questions regarding its emotional underpinnings, its explanatory power, and whether it reflects actual experience, remain open. Neuroeconomics can potentially contribute important insights to these ongoing debates. The relevant neuroeconomic research conducted thus far is inconclusive regarding whether loss aversion is the byproduct of a single system or multiple systems, confirms the important role of distress in generating the endowment effect, and suggests that losses are actually experienced more intensely than comparable gains.

More research is needed to address each debate, but determining the underlying cause(s) of behavioral loss aversion is perhaps the most promising avenue for future research. The debate over the role of fear in loss aversion is at an early stage, but fairly simmering already (e.g., De Martino et al., 2010; Tom et al., 2007). Several methods could be brought to bear on this debate. For example, it would be useful to examine whether drugs that reduce amygdala activation (e.g., lorazepam; Paulus, Feinstein, Castillo, Simmons, & Stein, 2005) moderate behavioral loss aversion. Additionally, diffusion tensor imaging, a relatively new technology that can reveal the connectivity between neural regions, could offer insight into De Martino et al.’s (2010) hypothesis that the amygdala conveys aversive responses to the regions that demonstrated neural loss aversion in Tom et al. (2007). For behavioral researchers, it would be valuable to examine whether methods that tend to increase...
the influence of emotion, such as cognitive load (Shiv & Fedorikhin, 1999), moderate behavioral loss aversion.

Of course, beyond loss aversion, neuroeconomics has great potential to shed light on the processes underlying many phenomena of interest to consumer researchers (cf. Shiv, 2007). The cost of fMRI is high, but that cost is most justifiable when underlying processes are difficult to articulate, either because they occur outside of awareness or because people are reluctant to report the truth. Consider, for example, the moderately positive relationship between price and perceived quality (e.g., Rao & Monroe, 1989). Imagine that an experimenter manipulates only the price that participants are required to pay for a consumable and then demonstrates a positive relationship between price and self-reported pleasure of consumption. This price placebo effect could be driven by actual treatment differences in experienced pleasure. Alternatively, participants may simply use prices to help make sense of an ambiguous consumption experience. Dissonance reduction may also play a role: the more people pay, the more they may be motivated to report that what they paid for was of the utmost quality. It is difficult to distinguish between these possibilities on the basis of self-reports alone.

Plassmann et al. (2008) utilized fMRI to help shed light on the price placebo effect. In their study, participants sampled small amounts of Cabernet Sauvignon immediately after learning that its retail price was either low or high (e.g., $10 or $90). The actual wine (delivered via a tube while participants lay in the scanner) was held constant across low and high price points. Plassmann et al. (2008) found that the wines were reported as more pleasant when they were expensive. They also found that activation in the medial orbitofrontal cortex (mOFC) correlated positively with prices and with self-reported pleasantness ratings. This was consistent with previous work demonstrating a positive correlation between mOFC activation and self-reported pleasantness ratings of odors, consumables, and music. Most importantly, Plassmann et al. (2008) found that prices did not predict activation in primary taste areas that encode the actual sensory experience of consumption. The results suggest that prices influence self-reported pleasure by changing how the consumption experience is represented in the brain, but not by directly influencing the basic, bottom-up sensory components of pleasure.

Neuroeconomics promises to shed light on other underlying processes that appear murky when viewed solely through the lens of self-reports. However, when considering the promise of neuroeconomics, it is also important to be aware of its limitations. It is often difficult to draw causal conclusions from any particular neuroeconomics study. fMRI data are correlational, and it is often unclear whether the identified correlates are actually driving the behavior in question or whether the neural activity is merely a byproduct of the process that is actually driving behavior. Repetitive transcranial magnetic stimulation, which temporarily (and non-invasively) disrupts activation in a particular region (creating a “virtual lesion”), is increasingly used to examine the causal influence of the disrupted region (e.g., Knoch, Pascual-Leone, Meyer, Treyer, & Fehr, 2006; Knoch, Schneider, Schunk, Hohmann, & Fehr, 2009; Figner et al., 2010), but there is still uncertainty about how other regions respond when the targeted region is disrupted, complicating interpretation of the data. It is also difficult to draw causal conclusions from traditional brain lesion studies, because no two lesions are exactly alike, and it is typically unclear how the rest of the brain has compensated for the damage.

Additionally, fMRI researchers are often forced to draw “reverse inferences” from the data (Poldrack, 2006). Specific cognitive or emotional processes are often inferred from activation in a particular region, which can be misleading when that region is not selectively activated by a particular task or process (Huettel & Payne, 2009; Yoon, Gonzalez, & Bettman, 2009). That is, when a region is unselectively activated by many different processes, extreme caution is required when using activation in that region to infer the engagement of a particular process. Indeed, researchers are only beginning to be able to use neural activation to predict which of several different tasks (e.g., gambling decisions versus semantic judgments) participants are performing in the scanner (Poldrack, Halchenko, & Hanson, 2009).

Nevertheless, neural data seem to have a powerful placebo effect on the persuasiveness of scientific arguments. For example, Weisberg, Keil, Goodstein, Rawson and Gray (2008) asked students to read descriptions of several phenomena (e.g., curse of knowledge). Half of the participants received “good” descriptions, which explained why the phenomenon occurs, and half received “bad” descriptions, which simply restated the phenomenon and provided no psychological insight. This manipulation was crossed with a neuroscience information manipulation: half of the descriptions included minimal information about the neural correlates of the phenomenon, which did not alter the underlying logic of the explanation itself (e.g., “brain scans indicate that this “curse” happens because of the frontal lobe brain circuitry known to be involved in self-knowledge”), and half included no neuroscience information. Participants rated the extent to which each description was intellectually satisfying. Good explanations were equally satisfying with and without neuroscience information. Bad explanations, however, were significantly more satisfying when neuroscience information was present than when it was not. The minimally informative neuroscience data essentially rescued the bad explanations.7

Despite the limitations inherent in neuroscientific methods, the fairly exploratory nature of the early neuroeconomics research, and the tendency to overweight neuroscientific findings, it is difficult to be anything but optimistic about the potential for neuroeconomics to shed light on phenomena of interest to consumer researchers. In the domain of loss aversion,

---

7 Weisberg et al. (2008) found that neuroscience experts had the opposite reaction: bad explanations were equally dissatisfying with and without neuroscience information, but good explanations with neuroscience information were significantly less satisfying than good explanations without neuroscience information. The experts apparently realized that the neuroscience information added nothing to the explanation, and were thus somewhat annoyed with its presence. As familiarity with neuroscientific methods increases in other fields, neuroscientific data that do not substantively add to our understanding of phenomena may begin to face such boomerang reactions.
behavioral researchers and neuroeconomists are already generating complementary insights. However, the open questions about loss aversion considered in this review have yet to be answered definitively. Conclusive answers will require consumer researchers to be aware of the open questions and controversies stemming from neuroeconomics, and vice versa. This review is intended as a step in that direction.

Author's note
I thank the editor, the associate editor and two anonymous reviewers, Katherine Burson, and Carolyn Yoon for valuable feedback. Any errors are my own.

References


