Word count

Abstract: 181
Text (excluding tables and figure captions): 10,892

Differential Discounting and Present Impact of Past Information

ABSTRACT

How does information about a person's past, accessed now, affect individuals' impressions of that person? In two survey experiments and two experiments with actual incentives we compare whether, when evaluating a person, information about that person's past greedy or immoral behaviors is discounted similarly to information about her past generous or moral behaviors. We find that, no matter how far in the past a person behaved greedily or immorally, information about her negative behaviors is hardly discounted at all. In contrast, information about her past positive behaviors is discounted heavily: recent behaviors are much more influential than behaviors that occurred a long time ago. The lesser discounting of information about immoral and greedy behaviors is not caused by these behaviors being more influential, memorable, extreme, or attention-grabbing; rather, they are perceived as more diagnostic of a person's character than past moral or generous behaviors. The phenomenon of differential discounting of past information has particular relevance in the digital age, where information about people's past is easily retrieved. Our findings have significant implications for theories of impression formation and social information processing.

Keywords: Discounting of Information, Impression Formation, Diagnosticity, Experiments, Dictator Game

Differential Discounting and Present Impact of Past Information

“The evil that men do lives after them;
The good is oft interred with their bones.”
Shakespeare, Julius Caesar

The web, they say, means the end of forgetting (Rosen, 2010): Much of people’s fleeting good and bad (an article mentioning our professional successes; an embarrassing photo we uploaded on the spur of the moment) is recorded – preserved as in amber in the eternal present of a Google search – and easily accessed. How does the fact that information about people’s past can be so readily accessed now end up affecting how people are judged? Is information about past negative behaviors discounted as much as information about past positive behaviors? For instance, when others form a first impression about us, is the impact of embarrassing pictures uploaded five years ago discounted as much as the impact of a five year old article mentioning our professional success?

Previous literature in psychology and behavioral decision research has extensively considered the differential weighting of negative and positive information in general – that is, independent of time – as well as the differential weighting of negative and positive information over time. The ubiquity and effortless accessibility of personal information brought forward by modern information systems, however, has given relevance to a particular scenario: what is the present impact of past information – or, in other words, how information about the past is discounted, rather than how future information gets discounted. In this manuscript, we develop and test a differential discounting of past information hypothesis: when forming impressions about a target person, information about past negative (greedy or immoral) behaviors is discounted less than information about past positive (generous or moral) behaviors. In four experiments (two survey experiments and two experiments with monetary incentives) we find evidence supporting this hypothesis, and test for the mechanism explaining these results. We find that lesser discounting of past negative behaviors cannot be explained by negative behaviors being more influential, memorable, extreme, or attention-grabbing; rather they are perceived as more diagnostic of personality. Besides contributing to theories of impression formation, these results suggest consequences for public policy governing the storage and accessibility of personal data, as well as for market outcomes, such as in hiring decisions.

We first review the literature on the differential weighting of negative and positive information in general, and the literature on differential weighting of negative and positive information over time. We then develop and test our hypothesis that information about past negative (greedy or immoral) behaviors is discounted less than information about past positive (generous or moral) behaviors.

Background

Differential Weighting of Negative and Positive Information

Prior research suggests that ‘bad’ information has, in general, a stronger impact than ‘good’ information on human behavior (for two comprehensive reviews, see Rozin & Royzman, 2001; Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001). This is not to say that people are generally pessimistic. To the contrary, in many judgment and decision making tasks people have been found to be too optimistic. The literature distinguishes between two types of overoptimism: wishful thinking and overconfidence (Vosgerau, 2010). Overconfidence describes people’s overoptimism with respect to their own performance (for a review, see Moore & Healy, 2008). Wishful thinking, in contrast, denotes people’s overoptimism about the likelihood of future events that are unrelated to their performance (Krizan & Windschitl, 2007). In general, motivated reasoning causes people to be overoptimistic in their expectations and reasoning about their own performance and future events (Kunda, 1990).

In contrast to overoptimistic expectations and reasoning, the “bad is stronger than good” hypothesis concerns how information – for example, the behaviors of others or experiences of events – affects human behavior. Negative information in general may have a stronger impact than positive information because it can have broader and more influential consequences than positive information. This differential impact implies a longer lasting effect of bad as compared to good. In particular, many studies have found that negative events – such as behaviors by others, physical accidents, or monetary losses – have a stronger impact on judgment, information processing, and learning than their positive counterparts. Several theories have been advanced to explain why bad in general is stronger than good. These theories can be broadly divided into two classes: theories that focus on the affective consequences of positive versus negative information, and theories that emphasize the informational value of positive versus negative information.

The most prominent theory focusing on the affective consequences of information is an evolutionary account. According to this theory, it may be adaptive to weigh negative information more than positive information, because ignoring negative information may have more severe consequences than ignoring positive information (Baumeister et al., 2001). This can be expressed as an asymmetric loss-function (Weber, 1994). Given the potentially more severe consequences of ignoring negative information, animals and humans may possess an innate predisposition to give greater weight to negative entities (e.g., events, objects, personal traits; Rozin & Royzman, 2001). Indeed, Ito, Larsen, Smith, and Cacioppo (1998) provided neurological evidence that event-related brain potentials are more sensitive to negative than to equally probable, evaluatively extreme, and arousing positive stimuli. The risk-reward model by Peeters (1971; 1983; also see Peeters & Czapinski, 1990) is based on a similar premise. According to this theory, negative behaviors are perceived as more threatening and risky than positive behaviors are perceived as potentially rewarding.

In the judgment and decision making literature, a general negativity bias is described as loss aversion (Tversky & Kahneman, 1974). Numerous studies have shown that in gambles with monetary payoffs losses loom larger than gains. However, while loss aversion is easily demonstrated in simple gambles involving either gains or losses, it is surprisingly difficult to demonstrate with mixed gambles involving both gains and losses (Rozin & Royzman, 2001; Yechiam & Hochman, 2013). Yechiam and Hochman (2013) suggested that, rather than being weighted more than gains, losses increase attention being allocated to the task at hand. When people play simple gambles with either losses or gains, gambles involving losses increase attention allocated to the reinforcement structure of the gamble; as a consequence, people react more strongly to losses than to gains. When gambles involve both losses and gains (so-called mixed gambles), losses increase attention paid to both losses and gains, and consequently neither is over- or under-weighted. Yechiam and Hochman (2013) provided evidence for their attention model by showing that for mixed gambles most people are risk- and loss-neutral rather than risk- and loss-averse.

That negative stimuli trigger greater attention has also been proposed as an explanation for asymmetric weighting of inconsistent information in person impression formation (Fiske, 1980; Pratto & John, 1991). When presented with inconsistent information about a target person, for example the target person having shown positive and negative behaviors, negative and extreme behaviors are more influential in forming an overall impression of the target than positive behaviors. The greater impact of negative and extreme behaviors is explained by social negative information being more attention-grabbing because of the asymmetry of the loss-functions involved.

The earliest theory emphasizing the informational value of negative information is correspondent inference theory (Jones & Davis, 1965). Because people in general hold moderately positive views of the world and others, negative and extreme behaviors are perceived as non-normative. Non-normative behaviors bear social costs (Jones & McGillis, 1976), which is why people make stronger inferences about the personalities of those who display such non-normative behaviors. This leads to negative and extreme behaviors being overweighed compared to more normative positive and moderate behaviors.

Another informational explanation for a general negativity bias may be that negative information tends to be better remembered than positive information. For example, negative interpersonal events such as social exclusions are better recalled than their positive counterparts (Gardner, Pickett, & Brewer, 2000). However, this does not seem to hold for autobiographical memory: people in general have positively tainted recollections of their past, as most life events are perceived as positive and, more importantly, negative affect tends to fade faster than positive affect (Walker, Skowronski, & Thompson, 2003).

Finally, the most prominent theory focusing on informational value is the cue diagnosticity approach by Skowronski and Carlston (1987; 1989). Whereas all previous theories predict negative behaviors to have a stronger impact on impressions of a target person than her/his positive behaviors, Skowronski and Carlston argued that in certain domains positive behaviors can be more informative and hence carry more weight. For example, winning a chess tournament once is enough to judge a person as intelligent, whereas a person needs to show incompetence repeatedly to be judged as incompetent. So, in ability judgments competent (positive) behaviors are more diagnostic than incompetent (negative) behaviors; in judgments of morality, on the other hand, immoral (negative) behaviors are more diagnostic than moral (positive) behaviors (Reeder & Brewer, 1979). Stealing once is enough for a person to be labeled a thief, whereas a person needs to behave always morally to be called a person of high integrity. Skowronski and Carlston (1987) demonstrated empirically that in judgments of ability competent behaviors are more diagnostic than incompetent ones leading to a positivity bias, whereas in morality judgments immoral behaviors are more diagnostic than moral ones leading to a negativity bias. Note that affect-based theories positing a general negativity bias have difficulties explaining the existence of a positivity bias without invoking auxiliary assumptions (Skowronski & Carlston, 1989; Peeters & Czapinski, 1990; Skowronski, 2002). The cue-diagnosticity approach also makes an interesting prediction concerning judges' belief about the malleability of human traits. So-called entity-theorists believe traits to be fixed and fundamentally unchangeable, whereas incremental theorists believe intelligence and morality to be malleable and improvable through training, experience, and persuasion (Dweck, Chiu, & Hong, 1995; for a review, see Levy, Plaks, Hong, Chiu, & Dweck, 2001). If differential weighing of positive and negative information follows the diagnosticity of the behaviors in question, incremental theorists should show greater differential weighing than entity theorists. In contrast, in affect-based theories positivity and negativity biases should be independent of a judge's beliefs about trait malleability. Skowronsky (2002) empirically tested these opposing predictions and found support for the former, further bolstering support for positivity and negativity biases in person impression formation being driven by informational differences of the target persons' behaviors.

In sum, while there is an impressive amount of evidence across domains that 'bad' in general has a stronger impact than 'good,' there are also exceptions in each domain of human behavior. People tend to be overoptimistic in their expectations and reasoning about events as evidenced in motivated reasoning. Whether negative or positive information has a stronger impact depends on the consequences for the decision maker, on whether negative and positive information is experienced jointly or separately, and on the inferences that decision makers draw from the information.

Differential Weighting of Negative and Positive Information Over Time

Apart from the literature investigating whether negative or positive information has a stronger impact per se, two separate streams of literature have investigated how the impact of negative and positive information changes over time: research on time preferences and on hedonic adaptation.

The behavioral economics literature on time preferences has shown that the more an event lies in the future, the less it impacts current valuations and choices (this is called time discounting or discounting of the future). The impact of a future event on present judgments/choices is determined by multiplying the magnitude of the event (its negativity or positivity) with a discounting factor (e.g. exponential or hyperbolic) which varies over time. Many studies have shown that small amounts of money are discounted more than large amounts of money (the so-called ‘magnitude-effect’), and gains are discounted more than equivalent losses (the ‘sign-effect;’ cf., Benzion, Rapoport, & Yagil, 1989; Loewenstein, 1987; Thaler, 1981; for an overview, see Frederick, Loewenstein, & O’Donoghue, 2002).

Paralleling these findings, the literature on hedonic adaptation suggests that people adapt slower to negative events than to positive events, or equivalently, that negative events have a longer lasting impact than positive events. For example, Brickman, Coates, and Janoff-Bulman (1978) interviewed people who had won a lottery the year before, people who had been paralyzed in an accident the year before, and a control group. While the lottery winners were no happier than the control group, the accident victims were significantly less happy. A similar conclusion has been drawn from research on child abuse and sexual abuse.

Even if the abuse occurred only once or twice, it produced long-lasting harmful consequences, such as depression, relationship problems, re-victimization, and sexual dysfunction (Cahill, Llewelyn, & Pearson, 1991; Fleming, Mullen, Sibthorpe, & Bammer, 1999; Silver, Boon, & Stones, 1983; Styron & Janoff-Bulman, 1997; Weiss, Longhurst, & Mazure, 1999). In contrast, no positive experiences have been documented that have equally long-lasting beneficial effects (Frederick & Loewenstein, 1999; Baumeister et al., 2001).

While both streams of literature – time preference and hedonic adaptation – provide evidence that negative information has a longer-lasting impact than positive information, the studies in the two research streams differ in their time orientation and experience of time.

Regarding time orientation, studies on time discounting look at how future gains and losses influence current valuations and choices, whereas studies on hedonic adaptation investigate how past negative/positive events affect the current wellbeing of the person who experienced them.

Regarding the experience of time, in hedonic adaptation time is experienced, whereas in time discounting forecasts over periods of time are made. This is an important distinction, because experiencing time leads to very different outcomes than forecasting over time (cf., McGuire & Kable, 2015). For example, when experiencing time, people generally adapt quickly to most life events; however, when forecasting their reactions to future events, people generally overestimate how much and how long future events will affect them (Gilbert et al., 1998; Gilbert & Wilson, 2007). Similar differences for experienced versus forecasted time have been documented in discounting phenomena. In forecasting most people exhibit deep discounting of the future, but when people are explicitly encouraged to imagine their future selves – a form of making forecasted time more experiential – future events are less discounted (Ersner-Hershfield et al., 2009; Ersner-Hershfield, Wimmer, & Knutson, 2009).

A Differential Discounting of Past Behaviors Hypothesis

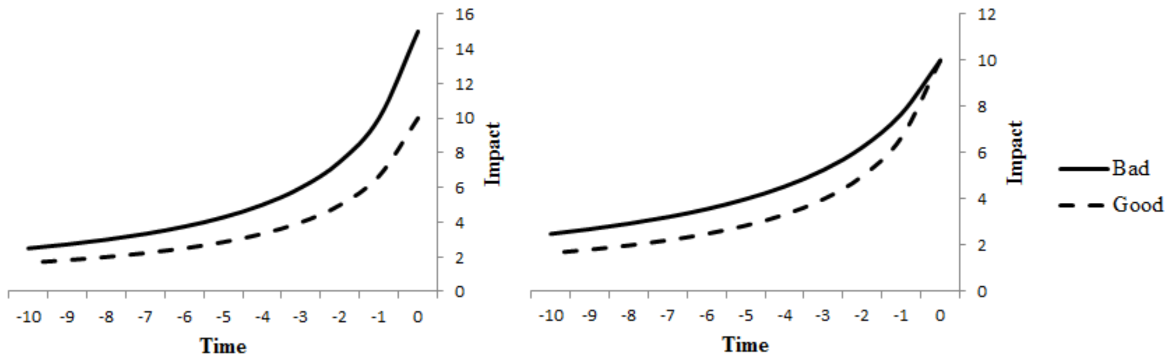
In this paper, we investigate how information about a person's past negative behaviors, compared to her past positive behaviors, affects others' impressions of, and their behavior towards that person, as a function of the point in time to which such information refers. Similar to studies of time discounting, our research looks at whether negative or positive information is discounted more the farther it lies in the past. Time in the tasks we use here, however, is not experienced; rather, judgments are made about the longevity of the impact of positive and negative information. Like research on hedonic adaptation, our research focuses on past rather than future information.

The longer lasting effect of bad found in the previous literature does not say much about differential discounting of bad and good, because a longer lasting effect of bad may simply result from the stronger impact of bad at the time it occurs (as hypothesized by Baumeister et al., 2001). For instance, imagine that one could measure the impact of good and bad with some objective index: Figure 1, Panel A depicts the value of such an impact measure over time for good and bad events, where discounting is modeled as hyperbolic discounting applied to time elapsed with the same discounting parameter applied to both events. The bad event has a stronger impact than the good event in the present (time 0), and thus its impact is discounted more over time (the slope of the bad-discounting curve is steeper than the slope of the good-discounting curve). Even though the impact of bad is discounted faster than the impact of good, the impact of bad lasts longer because bad has a stronger effect than good at the time it takes place.

In contrast, we investigate whether bad is discounted less than good in the past. Panel B in Figure 1 depicts such a case. Good and bad have an equal impact in the present, but are discounted differently over time, so the slope of the good-discounting curve is steeper than the slope of the bad-discounting curve. As a consequence, bad lasts longer than good, even though both had an equally strong effect at the time they took place. In this case the impact of bad is discounted less over time than the impact of good. In our experiments, we will manipulate elapsed (but not the experienced passage of) time at two levels: recent past and far past.

FIGURE 1: Hyperbolic past-time discounting ($f_H(T) = \frac{1}{1+kT}$) of good and bad with the same (Panel A, k = 0.5) or a different (Panel B, k = 0.5 for good and 0.3 for bad) discounting parameter.

Panel A Panel B



When presented with information about a person, people readily form an impression of her. First impressions not only include a general evaluation of the person (Zajonc, 1980), but also evoke trait concepts and stereotypes that activate corresponding behaviors (Fiske, Cuddy, & Glick, 2007; Todorov, Mandisodza, Goren, & Hall, 2005). Because immoral behaviors are more diagnostic about a person's character than moral behaviors (Skowrosnki and Carlston, 1987; 1989; Goodwin, Piazza, & Rozin, 2013; Pizarro & Tannenbaum, 2011), we hypothesize that information about past immoral behaviors is discounted less than information about past moral behaviors. The lesser discounting of past immoral behaviors should manifest itself in trusting and liking of the target person, because behaviors pertaining to morality and fairness/generosity are socially relevant and can be potentially harmful to others (Cacioppo, Gardner, & Berntson, 1997; Fiske et al., 2007; Fiske, Cuddy, Glick, & Xu, 2002; Peeters, 2002; Reeder, Kumar, Hesson-McInnis, & Trafimow, 2002).

We test this differential discounting of past behaviors hypothesis in four experiments. Experiment 1 tested whether, when evaluating a person, information about her past immoral behaviors is discounted less than information about her past moral behaviors. Experiment 2 tested whether discounting follows the diagnosticity of behaviors. If so, positive behaviors should be discounted more than negative behaviors in morality judgments, whereas in judgments of ability negative behaviors should be discounted more than positive behaviors. Experiment 3 and 4 explored the behavioral consequences of such judgments.

As in experiments on time discounting in which the time of future events is manipulated to test its impact on current behavior/decisions, across our experiments we manipulate how much time has passed since a target person showed a specific behavior by presenting information that refers to a closer or more distant past.

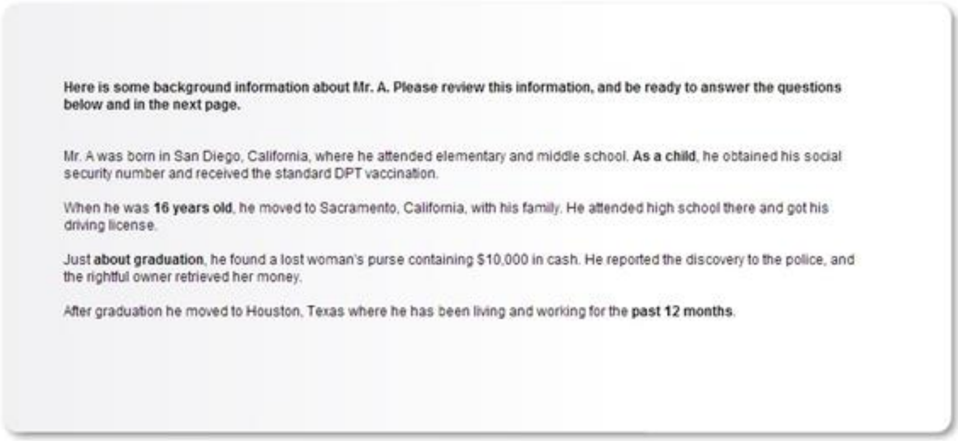
Experiment 1

Method

Participants. The experiment employed a 2 (moral vs. immoral behavior) x 2 (recent past vs. far past) + 1 control condition between-subjects design. Sample size was set to about 30 participants per condition; the data was analyzed only after data collection had been completed. One hundred sixty-eight students (80 females, M_{age} = 21.7, SD = 2.49) at a North Eastern American university were recruited. Students were approached by the experimenter in a popular indoor gathering place on campus and asked to complete a 5-minute online survey on impression formation in exchange for candies. Those who agreed were given a laptop on which they completed the study.

Procedure. Participants read background information about a target person (place of birth, schooling, where they got her driving license, where they went to college, their hometown and where they moved to, and their current job; for a screen shot, see Figure 2).

FIGURE 2. Background information in the positive valence recent behavior condition



Here is some background information about Mr. A. Please review this information, and be ready to answer the questions below and in the next page.

Mr. A was born in San Diego, California, where he attended elementary and middle school. **As a child**, he obtained his social security number and received the standard DPT vaccination.

When he was **16 years old**, he moved to Sacramento, California, with his family. He attended high school there and got his driving license.

Just **about graduation**, he found a lost woman's purse containing $10,000 in cash. He reported the discovery to the police, and the rightful owner retrieved her money.

After graduation he moved to Houston, Texas where he has been living and working for the **past 12 months**.

The background information was ordered by time of occurrence. In the negative valence condition, the background information included the additional information that the person had found a wallet with $10,000 in cash but did not report it to the police to be returned to the owner. In the positive valence condition, the person was reported as having returned the wallet with the cash to the police. The negative and positive behaviors were described to have occurred either recently (12 months ago) or in the past (5 years ago) and – to avoid spatial position effects – always appeared in the same position in the text. The control condition included only the neutral behaviors. Participants were then asked to answer three comprehension questions: What was the name of the person?, Where was that person born?, Did this person ever get a driver's license?, which all participants answered correctly. After the comprehension questions, participants were asked how much they liked the target person and whether they would like to work with this person on seven-point scales with endpoints “dislike very much”/“don’t want to work with this person” (1) and “like very much”/“want to work with this person” (7). Finally, the General Social Survey (GSS) / World Values Survey (WVS) trust measure was included (“Imagine that you met Mr. A for work: generally speaking, would you say that this person can be trusted, or that you can't be too careful when dealing with him?” Possible answers were: “You can’t be too careful when dealing with him”, “I don’t know”, and “This person can be trusted”, which we coded as 1 – Low Trust through 3 – High Trust). The study ended with demographic questions.

We standardized the three dependent variables – liking, wanting to work with the target person, and trust – and subjected them to a principal component analysis, which, using the eigenvalue greater than one criterion, yielded a single factor capturing 87.3% of the variance. Internal consistency proved to be satisfactory (Cronbach’s alpha = .93). We therefore averaged across all three standardized variables to obtain an overall impression index of the target person.

Results

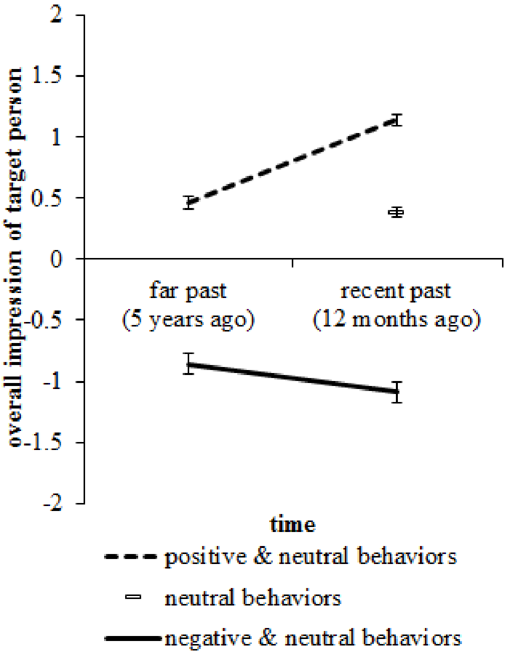
Manipulation Check. To verify whether our valence manipulation was successful, we tested the overall impression index scores across the three valence conditions for recent behaviors (positive recent behavior vs. control vs. negative recent behavior); $F(2, 96) = 308.47$, $p < .001$, $\eta_p^2 = .87$ (see Table 1). As expected, one recent positive behavior led to more positive impressions than only neutral behaviors, which in turn led to less negative impressions than one negative behavior.

Discounting of the Impact of Information about Moral/Immoral Behaviors. To test whether the impact of information on immoral behaviors is discounted less than information about moral behaviors (that is, whether discounting slopes differ), immoral and moral behaviors need to affect evaluation scores in the same direction. To this end, we multiplied evaluation scores in the immoral behavior conditions by -1. All subsequent analyses were performed on these transformed scores but we report the raw means for ease of exposition. A 2 (moral vs. immoral behavior) x 2 (recent vs. far past) analysis of variance (ANOVA) showed that information about recent behaviors had a stronger impact than information about behaviors in the past ($F(1, 132) = 41.55$, $p < .001$, $\eta_p^2 = .24$; see Figure 3). The predicted interaction of behavior and time was significant, $F(1, 132) = 9.69$, $p = .002$, $\eta_p^2 = .07$. Planned comparisons revealed that the target person having found and kept the wallet 5 years ago was evaluated less negatively ($M = -0.86$, $SD = 0.50$) than the target having found and kept the wallet 12 months ago ($M = -1.09$, $SD = 0.51$, $t(132) = 2.36$, $p = .020$, $d = 0.41$. In the positive domain, however, the differences in evaluations were more pronounced. Having found and returned the wallet 5 years ago led to much less positive evaluations ($M =0.46$, $SD = 0.30$) than having found and returned the wallet 12 months ago ($M = 1.14$, $SD = 0.30$); $t(132) = 6.76$, $p < .001$, $d = 1.18$. As shown by the difference in effect sizes, the immoral behavior was discounted less than the moral one.

TABLE 1. Standardized average impression of the target person as a function of her recent behaviors in Experiment 1. Standard deviations are in parentheses.

Recent behaviors by target person	Impression of target person (greater values indicate more positive impression)	*t*-test
One positive recent behavior, all others neutral	1.14 (0.30)	
		$t(63) = 11.03$, $p < .001$
Only neutral behaviors (control)	0.38 (0.26)	
		$t(64) = 14.67$, $p < .001$
One negative recent behavior, all others neutral	-1.09 (0.51)	

FIGURE 3. Evaluation of the target person as a function of behaviors displayed in the recent (12 months) or far past (5 years ago) in Experiment 1. Error bars represent +/- 1 SEM.



Discussion

In support of our hypothesis, information about a past immoral behavior was found to be discounted less than information about a past moral behavior. These results are consistent with the hypothesis that immoral behaviors are more diagnostic than moral behaviors (Skowronski & Carlston, 1987; 1989). If diagnosticity of behaviors drives their discounting, in judgments of ability one would expect positive (competent) behaviors to be less discounted than negative (incompetent) behaviors, because in this domain positive behaviors are more diagnostic than negative behaviors. To test this hypothesis, in Experiment 2 we manipulated the domain of judgments – morality and ability.

Experiment 2

Experiment 2 tested whether immoral behaviors are more diagnostic and hence less discounted than moral behaviors, and competent behaviors are more diagnostic and hence less discounted than incompetent behaviors.

In a pre-test, we first determined the diagnosticities of the moral/immoral and competent/incompetent behaviors to be used in the main experiment. In the main experiment, in addition to time elapsed (recent past vs. far past) and valence (positive vs. negative behaviors), we also manipulated the domain of judgments (morality versus intelligence). Participants were asked to judge a target person's morality/intelligence, and their overall liking and trustworthiness. We expected the lesser discounting of immoral behaviors to show in morality judgments and in person evaluations, because behaviors pertaining to morality and fairness/generosity are the primary determinants of liking and trust (Cacioppo, Gardner, & Berntson, 1997; Fiske et al., 2007; Fiske, Cuddy, Glick, & Xu, 2002; Peeters, 2002; Reeder, Kumar, Hesson-McInnis, & Trafimow, 2002; Skowronski, 2002; Wojciszke, 2005a; 2005b; Tausch, Kenworthy, & Hewstone, 2007). In contrast, the lesser discounting of intelligent behaviors was expected to show only in judgments of intelligence but not in person evaluations, because judgments of competence are less socially relevant than judgments of morality.

Pre-Test

The aim of the pretest was to test whether the immoral and intelligent behaviors to be used in Experiment 2 were more diagnostic than the moral and unintelligent behaviors. In the morality conditions, moral/immoral behaviors were the same as in Experiment 1. In the ability conditions, the critical positive/negative behaviors consisted of proving a hard math theorem or needing a calculator to add up two numbers of any kind (these behaviors were adapted from Skowronski & Carlston, 1987).

Diagnosticities of behaviors were measured following Skowronski and Carlston's (1987) cue diagnosticity approach. If immoral people are the only ones to display immoral (negative) behaviors, immoral behaviors should be more diagnostic than moral (positive) behaviors because moral behaviors can be displayed by both moral and immoral people. In other words, immoral behaviors should tell one more about the morality of a person than moral behaviors. The diagnosticity of a behavior is measured with conditional probabilities. The probability that a person is immoral (indicated by *IM*) given that s/he displayed an immoral behavior (indicated by *im*) should be greater than the probability that a person is moral (indicated by *M*) given that s/he displayed a moral behavior (indicated by *m*). Mathematically this is expressed as:

$$p(IM \mid im) > p(M \mid m) \tag{1}$$

Skowronski and Carlston call these conditional probabilities "cue-validity-indices" or simply "cue-validities." The higher the cue-validity of a behavior, the more diagnostic it is for the underlying trait. The computation of cue-validities follows the logic of Bayes' theorem:

$$p(M \mid m) = \frac{p(m \mid M)}{p(m \mid M) + P(m \mid IM)} \tag{2}$$

$$p(IM \mid im) = \frac{p(im \mid IM)}{p(im \mid IM) + p(im \mid M)} \tag{3}$$

For example, the cue-validity $p(M \mid m)$ of the moral (honest) behavior "finding and returning a lost purse containing $10,000" is the probability $p(m \mid M)$ that a moral (honest) person would do so, divided by the sum of this probability with the probability *p(m|IM)* that an immoral (dishonest) person would do so (formula 2). Likewise, the cue-validity *p(IM|im)* of the immoral (dishonest) behavior "finding but not returning a lost purse containing $10,000" is the probability *p(im|IM)* that an immoral (dishonest) person would do so, divided by the sum of this probability with the probability *p(im|M)* that a moral (honest) person would do so (formula 3). The computation of cue validities for intelligent and unintelligent behaviors follows the same logic.

Method

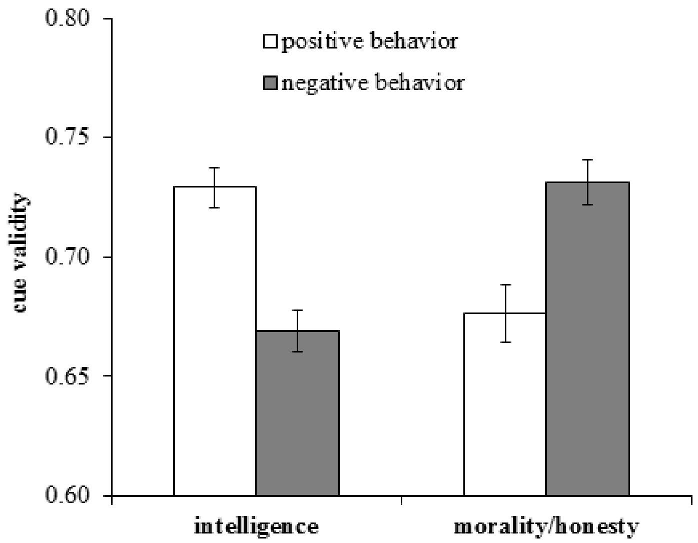
Participants. The pilot study employed a 2 (valence: positive vs. negative) x 2 (domain: intelligence vs. honesty) within-subject design. Sample size was set to 200 participants; data were analyzed only after data collection had been completed. Two hundred and one participants (115 male, M_{age} = 30.60, *SD* = 9.91) were recruited through Amazon Mechanical Turk (Paolacci, Chandler, & Ipeirotis, 2010) for a study on impression formation and were paid $.20 ($M_{\text{duration of study}}$ = 3.19mins, *SD* = 1.51).

Procedure. Following the procedure of Skowronski and Carlston (1987), participants were asked to indicate on 9-point scales with anchors, "Extremely Unlikely" (1), "Moderately Likely" (5), and "Extremely Likely" (9), how likely a person with a certain personality trait would be to engage in a certain behavior. The personality traits pertained to either the morality (honest/dishonest) or the intelligence (intelligent/unintelligent) of the target person. The behaviors consisted of reporting or not reporting a lost purse to the police containing $10,000 in cash, and proving a very hard math theorem or needing a calculator to add up two numbers of any kind. The order of the resulting eight questions was randomized.

Results

For each of the four behaviors (honest, dishonest, intelligent, and unintelligent), we calculated the resulting cue validities as described above (cf., Skowronski & Carlston, 1987). The four cue validities were subjected to a 2 (valence: positive vs. negative) x 2 (domain: intelligence vs. honesty) within-subject ANOVA. As expected, the valence x domain interaction was significant ($F(1, 200) = 48.36$, $p < .001$). The immoral (negative) behavior "finding and not returning a lost purse containing $10,000" was rated as more diagnostic ($M = 0.73$, $SD = 0.13$) than the moral (positive) behavior "finding and returning a lost purse containing $10,000" ($M = 0.68$, $SD = 0.17$; $t(200) = 4.32$, $p < .001$). In contrast, for judgments of intelligence, the intelligent (positive) behavior "proving a very hard math theorem" was rated as more diagnostic ($M = 0.73$, $SD = 0.12$) than the unintelligent (negative) behavior "needing a calculator to add up two numbers of any kind" ($M = 0.67$, $SD = 0.12$; $t(200) = 5.61$, $p < .001$; see Figure 4).

FIGURE 4: Cue-validities computed from rated probabilities for positive (moral and intelligent) and negative (immoral and unintelligent) behaviors in pretest for Experiment 2. Error bars represent +/- 1 SEM.



Main Experiment

Method

Participants. Experiment 2 employed a 2 (positive vs. negative behavior) x 2 (time: recent vs. far past) x 2 (domain: morality vs. ability) between-subjects design, with additional control conditions in both domains. According to our hypothesis, we expected a three-way interaction: positive moral and negative ability behaviors would be discounted more than negative moral and positive ability behaviors. In a similar pilot study conducted online, effect sizes were much smaller than in study 1, so we decided to increase sample size to about 100 participants per condition (as in study 1, the data were analyzed only after data collection had been completed). Nine hundred seventy-six participants were recruited through Amazon Mechanical Turk for a study on impression formation and were paid $.20 ($M_{\text{duration of study}}$ = 4.56mins, SD = 2.43).

Procedure. Following the findings of the pre-test, the moral/immoral behaviors were the same as in Experiment 1, and the positive/negative skill-related behaviors consisted of proving a hard math theorem or needing a calculator to add up two numbers of any kind. Participants were asked the same comprehension check questions as in Experiment 1 before being asked to rate the target person's honesty or intelligence, and how much they liked the target person, whether they would like to work with this person, and whether they would trust the target person. All measures used 7-point rating scales.

Results

Fifty-nine participants (6%) quit the study before completing it, and 102 participants (10.4%) failed to correctly answer the comprehension questions. Their responses were excluded, leaving 815 responses for further analysis (487 males, $M_{age} = 29.13$, $SD = 10.04$). Unless otherwise noted, results were virtually the same with those responses included.

Manipulation Check. The two dependent variables of liking and wanting to work with the target person showed high internal consistency (Cronbach's alpha = .90). We averaged the two variables to an overall impression index, and used this impression index, the level of trust, and participants' judgments of the target person's honesty/intelligence (morality/ability) as our dependent variables. Even though the impression index correlated highly with trust ($r = .79$, $p < .001$), we examined it separately to test whether similar patterns would be observed. While the correlation between the impression index and intelligence was high ($r = .60$, $p < .001$), trust did not correlate strongly with intelligence judgments ($r = .38$, $p < .001$; test of difference in correlations: $z = 6.01$, $p < .001$). In contrast, trust correlated highly with judgments of morality ($r = .92$, $p < .001$), as did the impression index ($r = .88$, $p < .001$). Not surprisingly, people trust those with high moral standards, but not necessarily those of high ability.

As in Experiment 1, we tested whether our valence manipulation was successful by comparing the honesty/intelligence ratings, overall impression index, and trust ratings across the three recent behavior valence conditions (honesty/intelligence: $F(2, 494) = 447.53$, $p < .001$, $\eta_p^2 = .64$; overall impression: $F(2, 494) = 148.19$, $p < .001$, $\eta_p^2 = .37$; trust: $F(2, 494) = 154.26$, $p < .001$, $\eta_p^2 = .38$; see Table 2).

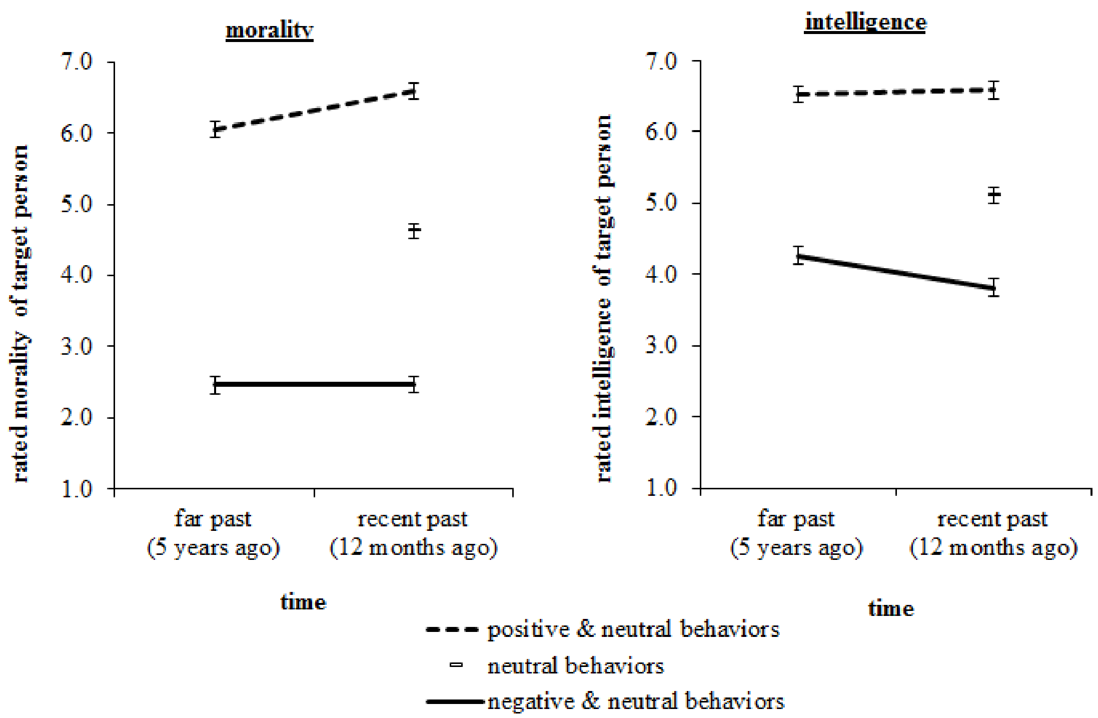
TABLE 2. Average morality, intelligence, and overall impression ratings of the target person as a function of her recent behaviors in Experiment 2. Standard deviations are in parentheses.

	Morality conditions		Ability conditions	
Recent behaviors by target person	Morality rating of target person	*t*-test	Intelligence rating of target person	*t*-test
One positive recent behavior, all others neutral	6.60 (0.79)		6.58 (0.73)	
		$t(164) = 14.77$, $p < .001$		$t(164) = 10.73$, $p < .001$
Only neutral behaviors	4.62 (0.92)		5.11 (1.02)	
		$t(165) = 12.88$, $p < .001$		$t(160) = 7.82$, $p < .001$
One negative recent behavior, all others neutral	2.46 (1.22)		3.81 (1.09)	
	Impression of target person	*t*-test	Impression of target person	*t*-test
One positive recent behavior, all others neutral	5.52 (0.96)		5.07 (0.87)	
		$t(164) = 7.44$, $p < .001$		$t(164) = 4.01$, $p < .001$
Only neutral behaviors	4.52 (0.75)		4.55 (0.81)	
		$t(165) = 10.26$, $p < .001$		$t(160) = 3.96$, $p < .001$
One negative recent behavior, all others neutral	2.99 (1.13)		4.09 (0.66)	

	Trust towards target person	*t*-test	Trust towards target person	*t*-test
One positive recent behavior, all others neutral	6.01 (0.99)	$t(164) = 12.71$ $p < .001$	4.71 (0.99)	$t(164) = 2.31$ $p = .022$
Only neutral behaviors	4.13 (0.91)		4.36 (0.95)	
One negative recent behavior, all others neutral	2.26 (1.21)	$t(165) = 11.29$ $p < .001$	4.00 (0.86)	$t(160) = 2.57$ $p = .011$

Morality/ability judgments. Similarly to Experiment 1, we transformed negative judgments, as well as trust and overall impression scores, in the negative behavior conditions. Since scores were on a scale from 1 to 7, we subtracted original scores from 8 in the negative behavior conditions. All subsequent analyses were performed on these transformed scores but we report the raw means for ease of exposition. A 2 (positive vs. negative behavior) x 2 (recent vs. far past) x 2 (ability vs. morality) ANOVA on morality/ability judgments showed that information about recent behaviors was more influential than information about past behaviors, $F(1, 643) = 10.34$, $p = .001$, $\eta_p^2 = .016$. The predicted behavior x time x domain interaction was significant, $F(1, 643) = 7.96$, $p = .005$, $\eta_p^2 = .012$ (not significant when including all participants; $F(1, 726) = 1.93$, $p = .165$, $\eta_p^2 = .003$). Planned contrasts revealed that in morality judgments, information about negative behaviors was discounted less than information about positive behaviors. It made little difference whether the target person was described as having found and kept the wallet 12 months ($M = 2.46$, $SD = 1.22$) or 5 years ago ($M = 2.47$, $SD = 1.38$), $t(643) = 0.41$, $p = .967$, $d = 0.01$, but the target person was judged to be more honest when they had found and returned the wallet 12 months ($M = 6.60$, $SD = 0.79$) rather than 5 years ago ($M = 6.05$, $SD = 1.14$), $t(643) = 3.38$, $p = .001$, $d = 0.55$. The opposite pattern was found in ability judgments: information about positive behaviors was discounted less than information about negative behaviors. The target person was rated equally intelligent whether they were described as having proven the math theorem 12 months (M = 6.58, SD = 0.73) or 5 years ago (M = 6.53, SD = 0.75), $t(643)$ = 0.36, p = .723, d = 0.07, but they were rated less intelligent when they had needed a calculator to add up two numbers of any kind 12 months (M = 3.81, SD = 1.10) than 5 years ago (M = 4.26, SD = 1.17), $t(643)$ = 2.70, p = .007, d = 0.40 (see Figure 5; no significant difference when including all participants; $t(726)$ = 0.63, p = .529, d = 0.087).

FIGURE 5. Raw means of morality and intelligence ratings of the target person as a function of behaviors displayed recently (12 months) or in the past (5 years ago) in Experiment 2. Error bars represent +/- 1 SEM.



Overall impressions and trust scores. Unlike the case of ability and morality judgments, for overall impressions of the target person we expected only information about moral behaviors to be discounted, because judgments of morality are more socially relevant than judgments of ability (Fiske et al., 2002).

A 2 (positive vs. negative behavior) x 2 (recent vs. far past) x 2 (ability vs. morality) ANOVA on overall impressions of the target person yielded a marginally significant interaction of behavior x time x domain, $F(1, 643) = 3.13$, $p = .077$, $\eta_p^2 = .005$ ($p = .113$ when including all participants). Replicating Experiment 1's results, planned contrasts revealed that it made little difference whether the target person had found and kept the wallet 12 months ($M = 2.99$, $SD = 1.13$) or 5 years ago ($M = 2.88$, $SD = 1.34$), $t(643) = 0.66$, $p = .506$, $d = 0.09$, but the target person was evaluated marginally more positively when they had found and returned the wallet 12 months ($M = 5.52$, $SD = 0.96$) rather than 5 years ago ($M = 5.25$, $SD = 1.04$), $t(643) = 1.79$, $p = .075$, $d = 0.26$. In the ability domain, in contrast, the target person was evaluated equally favorably, whether they had proven the math theorem 12 months ($M =$ 5.0.7, $SD = 0.87$) or 5 years ago ($M = 5.23$, $SD = 0.87$), $t(643) = 1.04$, $p = .300$, $d = 0.18$, and whether they had needed a calculator to add up two numbers of any kind 12 months ($M = 4.09$, $SD = 0.66$) or 5 years ago ($M = 4.10$, $SD = 0.89$), $t(643) = 0.08$, $p = .936$, $d = 0.01$ (see Figure 6 for pattern of raw means).

For trust scores, the interaction of behavior x time x domain was not significant, $F(1, 643) = 1.24$, $p = .267$, $\eta_p^2 = .002$. Planned contrasts revealed that it made little difference whether the target person had found and kept the wallet 12 months ($M = 2.26$, $SD = 1.21$) or 5 years ago ($M = 2.22$, $SD = 1.26$), $t(643) = 0.22$, $p = .823$, $d = 0.03$, but the target person was trusted more when they had found and returned the wallet 12 months ($M = 6.01$, $SD = 0.99$) rather than 5 years ago ($M = 5.61$, $SD = 1.18$), $t(643) = 2.34$, $p = .020$, $d = 0.37$. In the ability domain, the target person was equally trusted, whether they had proven the math theorem 12 months ($M = 4.71$, $SD = 0.99$) or 5 years ago ($M = 4.65$, $SD = 1.24$), $t(643) = 0.37$, $p = .711$, d

= 0.05, and whether they had needed a calculator to add up two numbers of any kind 12 months (*M* = 4.00, *SD* = 0.86) or 5 years ago (*M* = 4.01, *SD* = 1.12), $t(643) = 0.07$, $p = .943$, $d = 0.01$ (see Figure 7 for pattern of raw means).

FIGURE 6. Raw means of overall impression ratings of the target person as a function of behaviors displayed recently (12 months) or in the past (5 years ago) in Experiment 2. Error bars represent +/- 1 SEM.

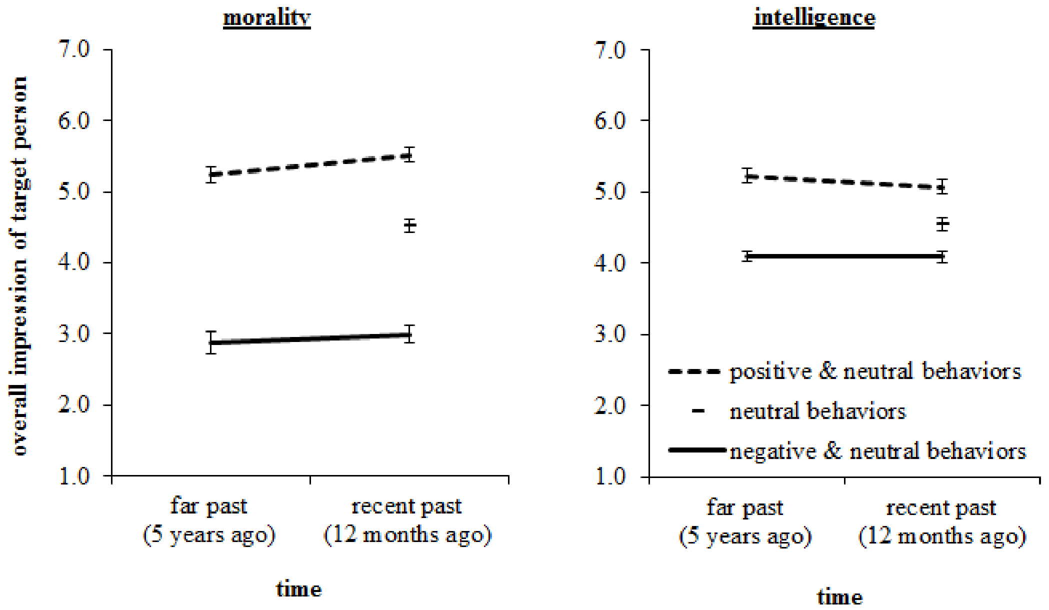
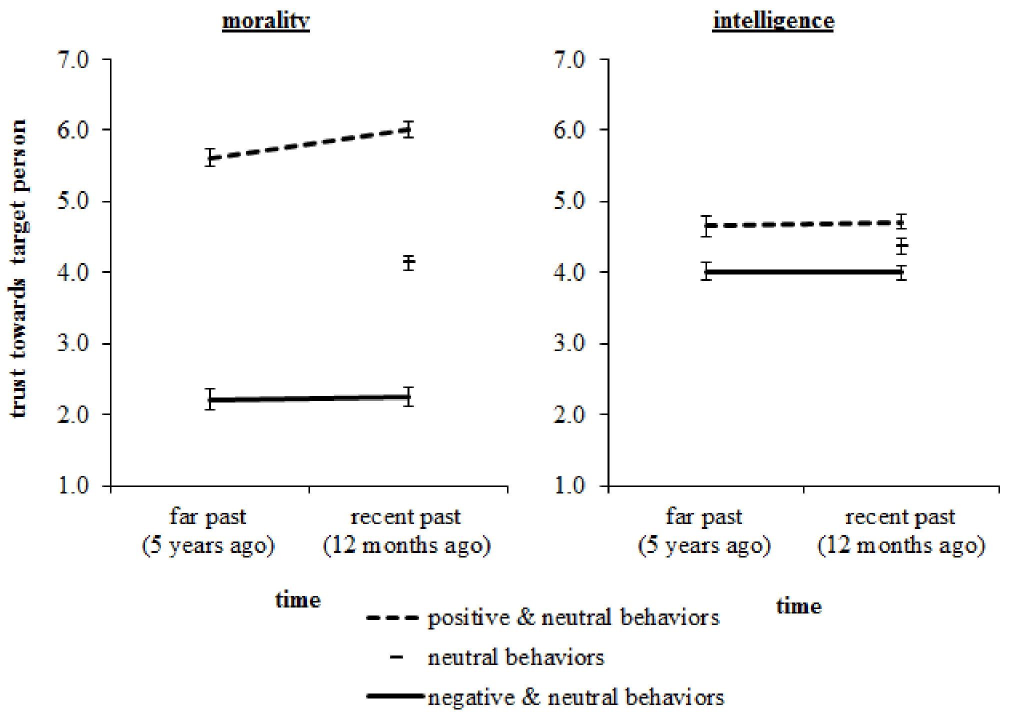


FIGURE 7. Raw means of trust ratings of the target person as a function of behaviors displayed recently (12 months) or in the past (5 years ago) in Experiment 2. Error bars represent +/- 1 SEM.



Discussion

While they may not conclusively demonstrate that differences in diagnosticity result in differential discounting, the results of Experiment 2 show that information about diagnostic (immoral and intelligent) behaviors is discounted less in judgments of a person's morality and ability than information about less diagnostic (moral and unintelligent) behaviors. This pattern of results cannot be explained by information about negative (immoral and unintelligent) behaviors being more attention-grabbing than information about positive (moral and intelligent) behaviors.

Information on behaviors concerning a person's intelligence only affected judgments of that person's intelligence. Information on behaviors concerning a person's morality, in contrast, affected both morality judgments and overall evaluations of (and trust towards) that person, because they are more socially relevant (Cacioppo et al., 1997; Fiske et al., 2007; Fiske, Cuddy, Glick, & Xu, 2002; Peeters, 2002; Reeder, Kumar, Hesson-McInnis, & Trafimow, 2002; Skowronski, 2002).

Experiment 3

While the previous two experiments focused on person evaluations, Experiment 3 and 4 were designed as a further test of why information about immoral behaviors is discounted less than information about moral behaviors in resulting behaviors toward the target person. Specifically, Experiment 3 tested whether resulting behavior towards the target person is indeed caused by people's inferences about the moral character of the target person (Goodwin et al., 2013; Pizarro & Tannenbaum, 2011). Behaviors perceived as intentional are treated as more indicative of an underlying trait or character than behaviors that are perceived as unintentional (Reeder et al., 2002). Consequently, we hypothesized information about immoral/greedy behaviors to be discounted when such behaviors were perceived as intentional but not when they were perceived as unintentional. The experiment consisted of a dictator game with actual cash incentives, and employed a 2 (valence: generous vs. greedy allocation) x 2 (time: recent vs. past) x 2 (intentional vs. unintentional) + 1 (control, i.e., fair allocation) between-subject design. Our hypothesis predicted a three-way interaction of valence, time, and intentionality.

Method

Participants. To achieve sufficient statistical power while keeping the cost of the study in check, sample size was set to about 40 participants per condition (as in the previous studies, data were analyzed only after data collection was completed). Three hundred seventy-three students (200 males, M_{age} = 22.8, SD = 1.32) at a North Eastern American university were recruited to play an "allocation game" to win money.

Procedure. Students were approached by the experimenter in a popular indoor gathering place on campus, and asked whether they would like to play a game with another player in a virtual game room. Participants logged onto the virtual game room (which we developed and hosted on a university server) on laptops provided by the experimenter.

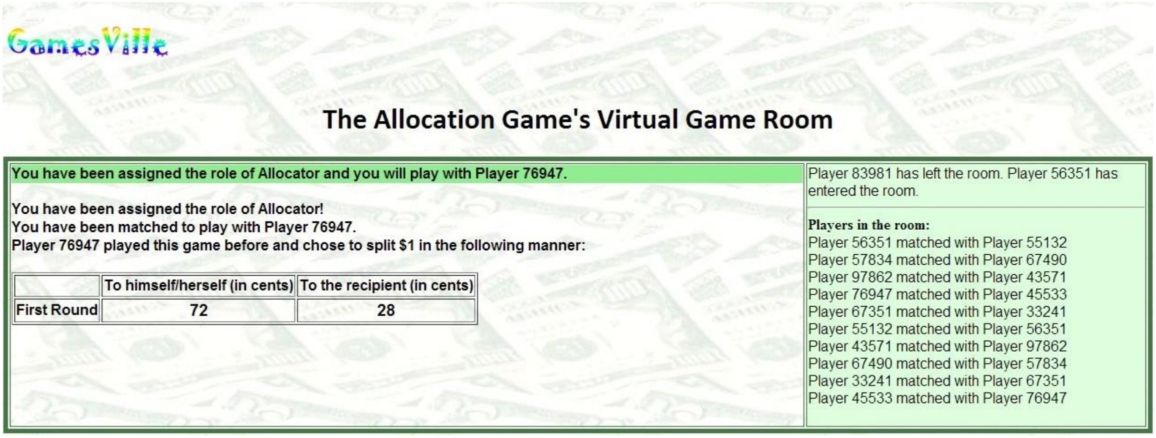
Prior to starting the game, each participant was told that $1 each was provisionally allocated to her and to all other players in the virtual game room. Participants would be randomly paired with another player (in reality, but unbeknownst to the participants, fictitious) and assigned the role of either "allocator" or "recipient." As allocator, they could decide whether and how to split the money; as recipient, they could only accept the decision of the allocator. In fact, all participants were assigned the role of allocator. Participants were further told that, on average, allocators keep 70% of the money for themselves and give 30% to the other player (the 70-30 split is the empirical average of dictator allocations; see Forsythe, Horowitz, Savin, & Sefton, 1994). Thus, we primed participants to consider a 70-30 division a fair allocation.

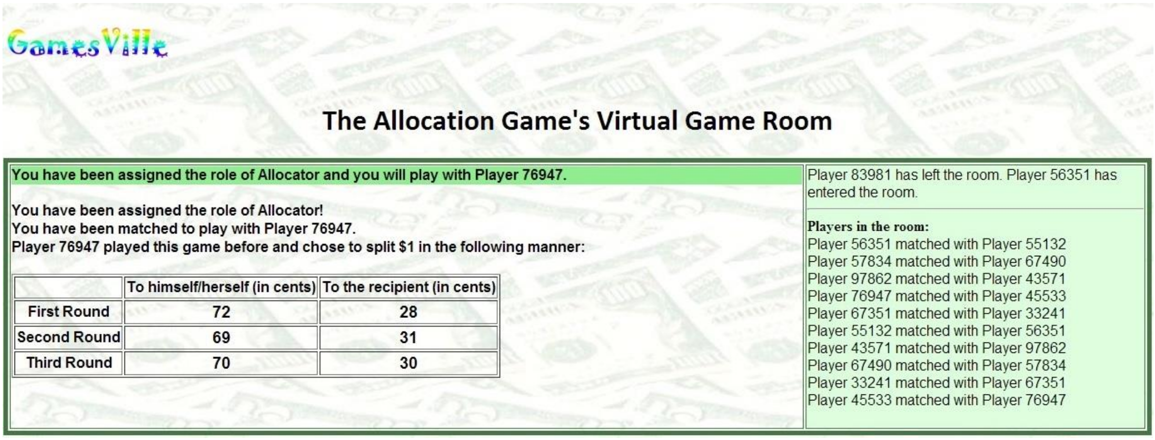
Before starting the game, participants were asked to describe the rules of the allocation game in their own words, and to indicate the average sum that allocators keep for themselves. All participants answered these two comprehension check questions correctly. Participants were then matched with the other (fictitious) player, and informed that said player had already played seven rounds of the game as an allocator. Before deciding how to allocate the money, participants were shown a table summarizing the allocations previously decided by the other player in her previous game. We systematically varied the content of the table to manipulate valence, time, and intentionality of the other player's previous allocations. In the control condition, all allocations were close to a fair 70-30 division. In the generous/greedy conditions, the other player had chosen a generous 50-50/greedy 100-0 allocation in one of the seven rounds, either in round 2 (past) or round 6 (recently). Finally, in the involuntary conditions, participants were informed that the generous/greedy allocation had been caused by a bug in the computer program, and the player had no chance to change that allocation.

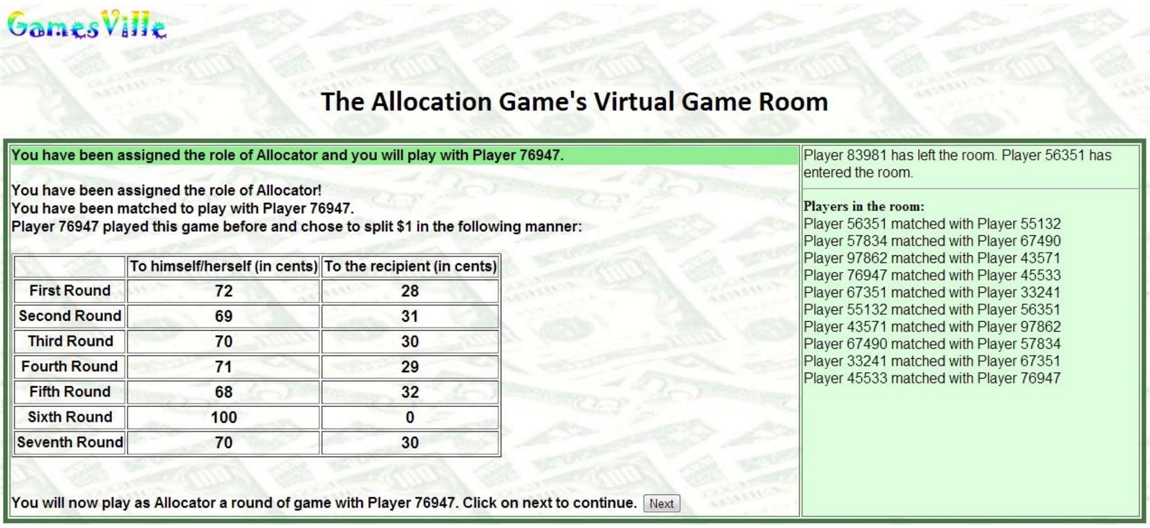
To provide a realistic feeling of a dynamic game taking place in a crowded virtual game room, rows of the table of previous allocations by the other player appeared sequentially on the computer screen, simulating the fictitious player making actual allocation decisions with other players before playing with the participant. This design also allowed us to make the temporal order of allocations more salient (for screen shots, see Figure 8).

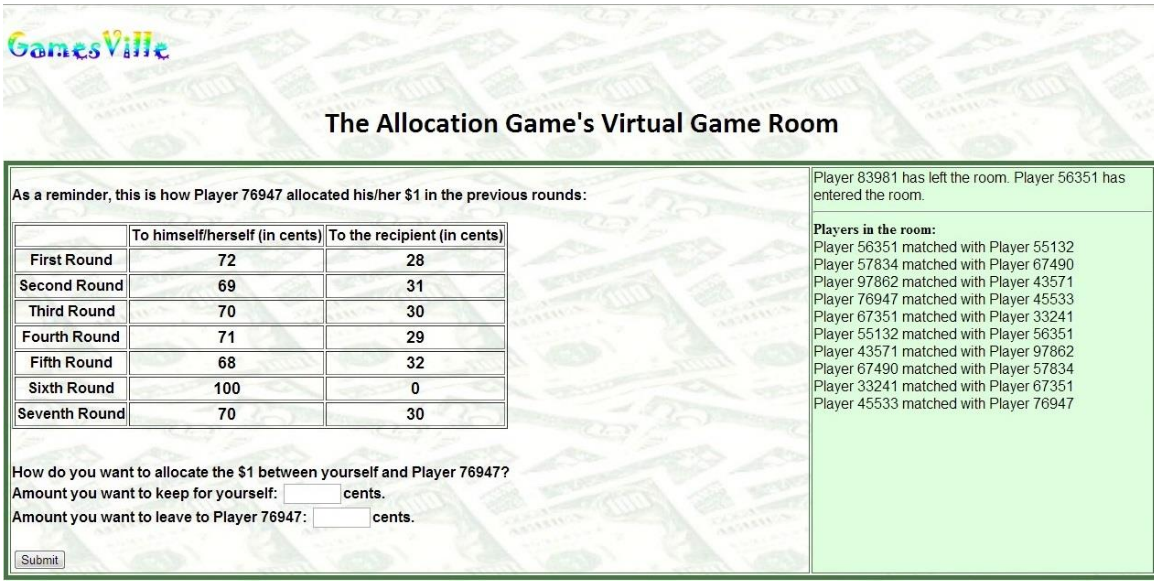
After the seventh round of previous allocations by the counterpart was displayed, participants were asked how they wanted to divide the $1 between themselves and the other player (the table with the previous allocations of the other player was still displayed). After making their allocation decision, participants were asked how fair they perceived the other player to be on a seven point scale, with endpoints "very unfair" (1) and "very fair" (7), and why they chose that allocation. Participants were also asked whether any technical error occurred during the game, and finally their demographics. The final thank-you screen summarized the payoffs, and participants were paid accordingly and debriefed by the experimenter.

FIGURE 8. Sequence of screenshots from the Allocation Game, Experiment 3.









GamesVille

The Allocation Game's Virtual Game Room

As a reminder, this is how Player 76947 allocated his/her $1 in the previous rounds:

	To himself/herself (in cents)	To the recipient (in cents)
First Round	72	28
Second Round	69	31
Third Round	70	30
Fourth Round	71	29
Fifth Round	68	32
Sixth Round	100	0
Seventh Round	70	30

How do you want to allocate the $1 between yourself and Player 76947?
Amount you want to keep for yourself: cents.
Amount you want to leave to Player 76947: cents.

Submit

Player 83981 has left the room. Player 56351 has entered the room.

Players in the room:
Player 56351 matched with Player 55132
Player 57834 matched with Player 67490
Player 97862 matched with Player 43571
Player 76947 matched with Player 45533
Player 67351 matched with Player 33241
Player 55132 matched with Player 56351
Player 43571 matched with Player 97862
Player 67490 matched with Player 57834
Player 33241 matched with Player 67351
Player 45533 matched with Player 76947

Results

Manipulation Check. As in the previous experiments, we tested whether our valence manipulation (greedy vs. generous allocation) was successful by comparing the amount of money that participants allocated to, and their fairness rating of the other player across the recent (round 6) generous, greedy, and control (fair allocation) conditions. The manipulation was successful (for allocations, $F(2, 204) = 37.53$, $p < .001$, $\eta_p^2 = .269$; for fairness ratings, $F(2, 204) = 68.76$, $p < .001$, $\eta_p^2 = .403$). For the average allocations to and fairness ratings of the other player as a function of her round 6 allocations, separated for the intentional and unintentional allocation conditions, see Table 3.

TABLE 3. Average allocations to and fairness ratings of the other player as a function of recent intentional/unintentional allocations by the other player in Experiment 3. Standard deviations are in parentheses.

	Intentional allocation conditions		Unintentional allocation conditions	
Recent allocations by other player	Allocation to other player	*t*-test	Allocation to other player	*t*-test
One generous allocation in round 6 (50-50)	42.75c (12.61)	$t(80) = 5.23$, $p < .001$	30.95c (11.85)	$t(82) = 0.5$, $p = .620$
All allocations fair (~70-30, control)	29.76c (9.75)	$t(83) = 8.65$, $p < .001$	29.76c (9.75)	$t(80) = 0.79$, $p = .430$
One greedy allocation in round 6 (100-0)	10.23c (11.02)		27.75c (13.1)	
	Fairness rating of other player	*t*-test	Fairness rating of other player	*t*-test
One generous allocation in round 6 (50-50)	5.25 (0.84)	$t(80) = 8.04$, $p < .001$	3.95 (0.80)	$t(82) = 0.15$, $p = .880$
All allocations fair (~70-30, control)	3.93 (0.64)	$t(83) = 16.43$, $p < .001$	3.93 (0.64)	$t(80) = 1.44$, $p = .150$
One greedy allocation in round 6 (100-0)	1.86 (0.52)		3.706 (0.79)	

Money Allocated to /Fairness Ratings of other Player. For the analyses, we standardized allocations to and fairness ratings of the counterpart and – as in the previous experiments – multiplied them by -1 in the greedy allocation (negative behavior) conditions. All subsequent analyses were performed on these transformed scores but we report the raw means for ease of exposition. A repeated measures ANOVA with the between-subject factors valence (generous vs. greedy allocation), time (round 6 vs. round 2), and intentionality (intentional vs. unintentional allocations) on participants' allocation to and fairness ratings of their counterpart showed that the counterparts' previous greedy/generous allocations were more influential when they had occurred recently in round 6 than in round 2, $F(1, 323) = 15.69$, $p < .001$, $\eta_p^2 = .046$. The predicted valence x time x intentionality interaction was significant, $F(1, 323) = 7.75$, $p = .006$, $\eta_p^2 = .023$.

Money Allocations. Planned contrasts revealed that in the intentional conditions, participants allocated more money to their counterpart when they knew she had made a generous allocation to another player in round 6 ($M = 42.75$ cents, $SD = 12.60$) than in round 2 ($M = 31.63$, $SD = 10.22$), $t(323) = 4.41$, $p < .001$, $d = .49$, but allocated the same amount of money to their counterpart when they knew she had made a greedy allocation in round 6 ($M = 10.23$, $SD = 11.02$) than in round 2 ($M = 12.93$, $SD = 11.01$), $t(323) = 1.08$, $p = .283$, $d = .12$. When the counterpart's previous allocations were unintentional, no differences were observed in money allocations (generous allocations round 6 ($M = 30.95$, $SD = 11.85$) vs. round 2 ($M = 31.19$, $SD = 11.52$), $t(323) = 0.10$, $p = .924$, $d = .01$; greedy allocations in round 6 ($M = 27.75$, $SD = 13.10$) vs. round 2 ($M = 28.25$c, $SD = 10.35$), $t(323) = 0.20$, $p = .846$, $d = .02$, see Figure 9).

Fairness Ratings. The same pattern was observed for fairness ratings: planned contrasts revealed that in the intentional conditions, participants rated their counterpart as fairer when they knew she had made a generous allocation to another player in round 6 ($M = 5.25$, $SD = 0.84$) than in round 2 ($M = 4.47$, $SD = 0.59$), $t(323) = 4.83$, $p < .001$, $d = .54$, but rated her equally unfair when they knew she had made a greedy allocation in round 6 ($M = 1.86$, $SD = 0.52$) than in round 2 ($M = 2.10$, $SD = 0.74$), $t(323) = 1.47$, $p = .143$, $d = .16$. When the counterpart's previous allocations were unintentional, no differences were observed in fairness ratings (generous allocations round 6 (*M* = 3.95, *SD* = 0.80) vs. round 2 (*M* = 4.07, *SD* = 0.95), *t*(323) = 0.74, *p* = .461, *d* = .08; greedy allocations round 6 (*M* = 3.7, *SD* = .79) vs. round 2 (*M* = 3.8, *SD* = .61), *t*(323) = 0.61, *p* = .546, *d* = .07, see Figure 10).

FIGURE 9. Participants' allocations to the other player as a function of the other player's previous generous/greedy/intentional/unintentional allocations in past (round 2) or recent (round 6) rounds in Experiment 3. Means of raw allocations are displayed, error bars represent +/- 1 SEM.

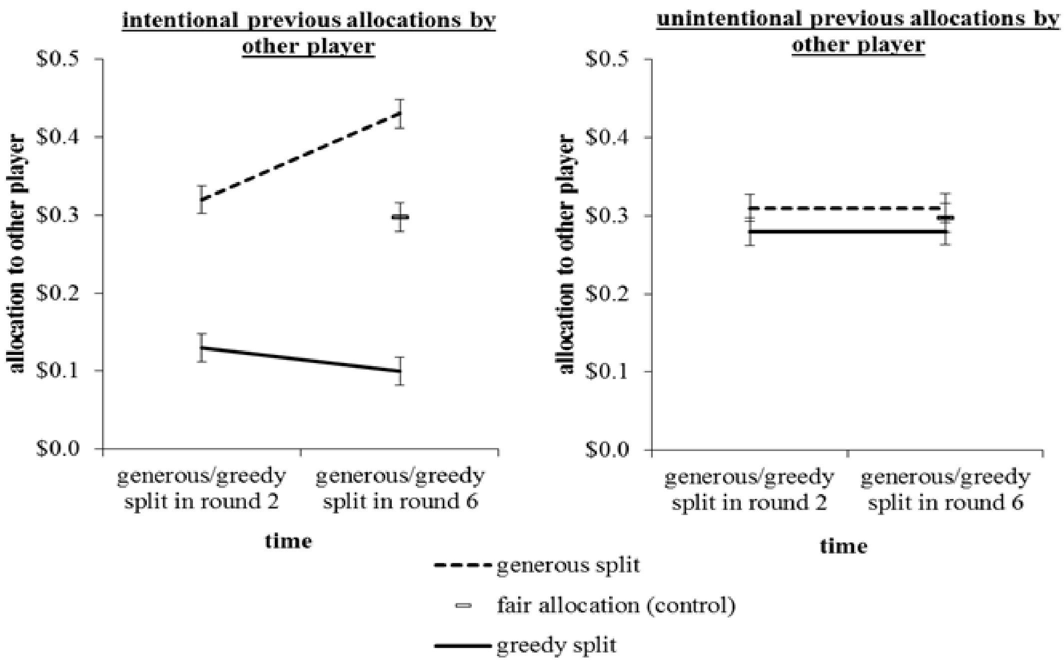
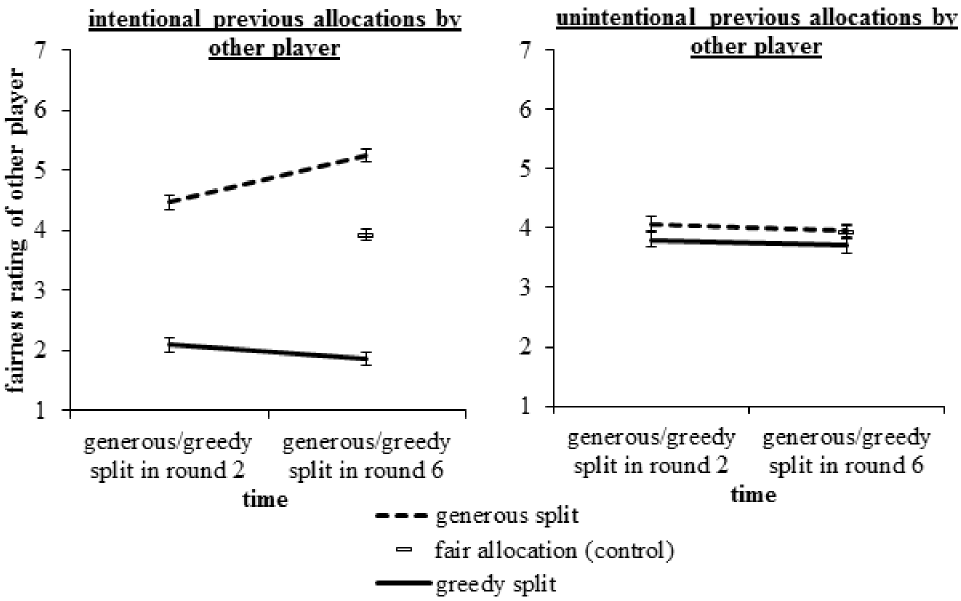


FIGURE 10. Fairness ratings of the other player as a function of the other player's previous generous/greedy/intentional/unintentional allocations in past (round 2) or recent (round 6) rounds in Experiment 3. Means of raw allocations are displayed, error bars represent +/- 1 SEM.



Discussion

A person being greedy once in a dictator game, even when it happened previously and with a complete stranger, significantly decreased participants' willingness to share money with that person. The effect of having been generous in the past, in contrast, quickly dissipated. Importantly, this happened only when the counterpart's allocations were perceived as intended.

Although Experiment 3 was not designed to test for the existence and magnitude of punishment utility (De Quervain et al. 2004), as in this experiment the punishment constitute per se a benefit for the punisher (higher payoff from the dictator game), the lower allocations to greedy counterparts in the dictator game may be interpreted as punishments for norm violations. A player who previously paid others less than the fairness norm dictates (i.e., a 70-30 split) was punished by equally low allocations by the dictator. While such punishments have been demonstrated in trust games, Experiment 3 makes an additional point: negative norm violations are punished no matter when they occurred, whereas positive norm violations are rewarded only if they occurred recently.

Two things are important to note, however. First, in the dictator game, the greedy behavior (a 100-0 split) was more extreme than the generous behavior (a 50-50 split) compared to the fair behavior (a 70-30 split). The so-called fair split of 70-30 was chosen because it represents the empirically found average in dictator games (Forsythe, Horowitz, Savin, & Sefton, 1994). This certainly contributed to information about greedy behaviors being discounted less than information about generous behaviors. Experiment 4 tests whether extremity alone is responsible for the pattern observed in Experiment 3. Second, the time period between the occurrence of an immoral/greedy behavior and the judgment of that behavior is perceived in relative terms. In Experiments 1 and 2, the described time difference was 4 years (5 years ago vs. 12 months ago); in Experiment 3 it was a few minutes (round 6 vs. round 2). We will come back to this point in the general discussion.

Experiment 4

Experiment 4 was similar to Experiment 3 but specifically tested whether resulting behavior towards the target person is affected by how extreme the allocations of the other player were. We hypothesized that extremity was not the leading factor determining participants' behavior. Rather, we expected inferred morality of the other player to be the main driver of participants' chosen money allocations and fairness ratings of the other player. The experiment consisted of a dictator game with actual cash incentives, and employed a 2 (valence: generous vs. greedy allocation) x 2 (time: recent vs. past) x 2 (extreme vs. moderate) + 1 (control, i.e., fair allocation) between-subject design. According to our theory, we would expect an interaction of valence and time, such that generous allocations are discounted more than greedy allocations, irrespective of their extremity. If discounting was caused by extremity rather than valence, an interaction of extremity and time would be expected such that extreme allocations are discounted less than moderate allocations, irrespective of whether they are generous or greedy.

Method

Participants. To achieve sufficient statistical power, sample size was set to about 100 participants per condition (as in the previous studies, data were analyzed only after data collection was completed). Nine hundred and seven students (515 males, M_{age} = 20.41, *SD* = 2.71) at a Southwestern American university were recruited to play an “allocation game” to win money.

Procedure. Procedures were identical to those in Experiment 3, except for the fact that we manipulated extremity of the other player’s allocation decisions instead of intentionality. The extreme allocations consisted of a 100/0 split in the greedy, and a 50/50 split in the generous condition (same as Experiment 3). The moderate allocations consisted of a 90/10 split in the greedy, and a 60/40 split in the generous condition (10 cents less extreme than Experiment 3).

Results

Manipulation Check. As in the previous experiments, we tested whether our valence manipulation (greedy vs. generous allocation) was successful by comparing the amount of money that participants allocated to, and their fairness rating of the other player across the recent (round 6) conditions. The manipulation was successful (for allocations, $F(4, 499) = 32.55$, $p < .001$, $\eta_p^2 = .207$; for fairness ratings, $F(4, 499) = 30.50$, $p < .001$, $\eta_p^2 = .196$). Allocations and fairness ratings differed for all extremeness levels except for the most generous behaviors. No significant differences were observed between a player's moderate and extreme generous behavior in terms of allocation to and fairness rating of that player. For the average allocations to and fairness ratings of the other player as a function of her round 6 allocations, separated for moderate and extreme allocations, see Table 4.

Money Allocated to /Fairness Ratings of other Player. For the analyses, we standardized allocations to and fairness ratings of the counterpart and – as in the previous experiments – multiplied them by -1 in the greedy allocation (negative behavior) conditions. All subsequent analyses were performed on these transformed scores but we report the raw means for ease of exposition. We ran a repeated measures ANOVA with the between-subject factors valence (generous vs. greedy allocation), time (round 6 vs. round 2), and extremity (extreme vs. moderate allocations) on participants' allocation to and fairness ratings of their counterpart. A main effect for time emerged showing that the counterparts' previous greedy/generous allocations were more influential when they had occurred recently in round 6 than in round 2, $F(1, 787) = 8.87$, $p = .003$, $\eta_p^2 = .011$. The main effect for extremity indicated that participants responded more to extreme than to moderate allocations of the counterpart, $F(1, 787) = 6.47$, $p = .011$, $\eta_p^2 = .011$. This main effect was qualified by a valence x extremity interaction, indicating that participants responded more strongly to extreme greedy than extreme generous allocations, irrespective of time, $F(1, 787) = 11.01$, $p = .001$, $\eta_p^2 = .014$. The valence x time interaction was marginally significant, showing that – as predicted by our own account – participants discounted greedy allocations less than generous ones, $F(1, 787) = 3.65$, $p = .056$, $\eta_p^2 = .005$. In contrast, the extremity x time interaction – which would indicate that discounting is driven by extremity rather than valence – was not significant, F(1, 787 = 0.17, $p = .68$, $\eta_p^2 < .001$.The triple interaction was not significant, $F(1, 787) = 2.23$, $p = .136$, $\eta_p^2 = .003$.

TABLE 4. Average allocations to and fairness ratings of the other player as a function of recent moderate/extreme generous/greedy allocations by the other player in Experiment 4. Standard deviations are in parentheses.

	Allocations to other player		Fairness rating of other player	
Recent allocations by other player	Allocation to other player	*t*-test	Fairness rating of other player	*t*-test
Extremely generous allocation in round 6 (50-50)	36.43c (13.67)		4.77 (1.23)	
		$t(204) = 0.41$, $p = .679$		$t(203) = 1.19$, $p = .236$
Moderately generous allocation in round 6 (60-40)	37.38c (18.87)		4.99 (1.45)	
		$t(206) = 3.76$, $p < .001$		$t(205) = 3.16$, $p = .002$
All allocations fair (~70-30, control)	28.61c (14.54)		4.38 (1.32)	
		$t(202) = 2.33$, $p = .021$		$t(202) = 3.63$, $p < .001$
Moderately greedy allocation in round 6 (90-10)	22.74c (21.10)		3.53 (2.00)	
		$t(193) = 3.35$, $p = .001$		$t(192) = 1.88$, $p = .062$
Extremely greedy allocation in round 6 (100-0)	13.75c (15.88)		3.07 (1.24)	

Money Allocations. A 2 (valence: generous vs. greedy allocation) x 2 (time: round 6 vs. round 2) x 2 (extremity: extreme vs. moderate allocations) ANOVA showed the same effects as described above: a main effect for time, $F(1, 794) = 9.79$, $p = .002$, $\eta_p^2 = .012$, and a main effect for extremity, $F(1, 794) = 5.41$, $p = .020$, $\eta_p^2 = .007$, which was qualified by a valence x extremity interaction, $F(1, 794) = 9.52$, $p = .002$, $\eta_p^2 = .012$. More importantly, the interaction between valence and time was significant, $F(1, 794) = 4.21$, $p = .040$, $\eta_p^2 = .005$, whereas the interaction between extremity and time was not, $F(1, 794) = 1.24$, $p = .27$, $\eta_p^2 = .002$, indicating that discounting is driven by valence rather than extremity. Planned contrasts revealed a pattern consistent with our hypotheses. For the moderate conditions, participants allocated more money to their counterpart when they knew the counterpart had made a generous allocation to another player in round 6 (M = 37.38 cents, SD = 18.87) than in round 2 (M = 31.26, SD = 14.63), $t(794) = 2.64$, $p = .009$, $d = .18$, but allocated the same amount of money to their counterpart when they knew the counterpart had made a greedy allocation in round 6 (M = 22.74, SD = 21.10) than in round 2 (M = 21.33, SD = 15.67), $t(794) = 0.596$, $p = .552$, $d = .04$. Similarly, when the counterpart's previous allocations were extreme, participants allocated more money to their counterpart when they knew the counterpart had made a generous allocation to another player in round 6 (M = 36.43, SD = 13.67) than in round 2 (M = 30.44, SD = 14.07), $t(794) = 2.60$, $p = .01$, $d = .18$, but allocated similar amounts of money when they knew the counterpart had made a greedy allocation in round 6 (M = 13.75, SD = 15.88) vs. round 2 (M = 17.67, SD = 17.11), $t(794) = 1.66$, $p = .096$, $d = .12$, see Figure 11).

The three-way interaction of valence, time, and extremity was not significant, $F(1, 794) = 1.36$, $p = .243$, $\eta_p^2 = .002$.

Fairness Ratings. A 2 (valence: generous vs. greedy allocation) x 2 (time: round 6 vs. round 2) x 2 (extremity: extreme vs. moderate allocations) ANOVA yielded only an interaction between valence and extremity, $F(1, 787) = 4.48$, $p = .035$, $\eta_p^2 = .006$. Neither the interaction between valence and time, $F(1, 787) = 1.00$, $p = .317$, $\eta_p^2 = .001$, nor the interaction between extremity and time were significant, $F(1, 787) = 0.16$, $p = .69$, $\eta_p^2 < .001$ (see Figure 12).

FIGURE 11. Participants' allocations to the other player as a function of the other player's previous generous/greedy/extreme/moderate allocations in past (round 2) or recent (round 6) rounds in Experiment 4. Means of raw allocations are displayed, error bars represent +/- 1 SE.

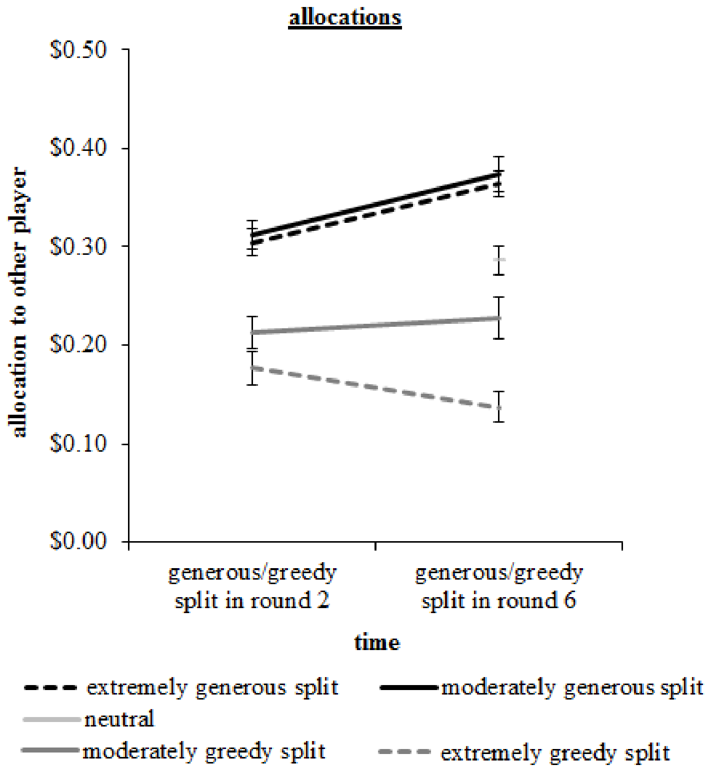
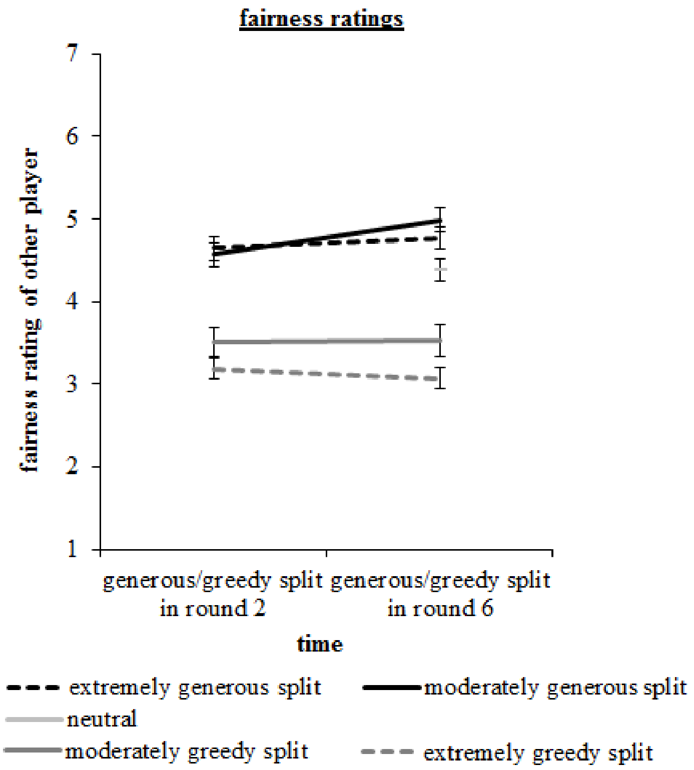


FIGURE 12. Fairness ratings of the other player as a function of the other player's previous generous/greedy/extreme/moderate allocations in past (round 2) or recent (round 6) rounds in Experiment 3. Means of raw allocations are displayed, error bars represent +/- 1 SE.



Discussion

Experiment 4 was designed to test whether extremity could explain lesser discounting of greedy allocations than generous allocations. To this end, allocations were varied at five levels: extremely generous 50/50, moderately generous 60/40, fair 70/30, moderately greedy 90/10, and extremely greedy 100/0. This experimental design allowed us to test whether generous allocations are discounted more than greedy ones (which would be denoted by an interaction of valence and time), or whether moderate allocations are discounted more than extreme ones (which would be denoted by an interaction of extremity and time). Participants' allocations to their counterparts provide support for the former hypothesis. Replicating the findings from Experiment 3, moderate and extreme generous previous allocations of the counterpart were discounted more than moderate and extreme greedy allocations. The same pattern was observed for fairness ratings, however, their valence x time interaction did not reach significance.

General Discussion

In four experiments, we observed that information about past immoral and greedy behaviors is discounted less than information about past moral and generous behaviors. A person having once shown an immoral or greedy behavior is judged and treated in kind, no matter when the negative behavior occurred. The impact of moral or generous behaviors, in contrast, gets quickly discounted. People believe that they can change for the better (Wilson & Ross, 2001), but when judging others they heed Shakespeare's advice: "For trust not him that hath once broken faith" (Shakespeare, 1591).

While the findings may not conclusively demonstrate its causal effect, diagnosticity seems the only theory capable of explaining both a negativity bias in morality judgments and impression formations, and a positivity bias in ability judgments. Immoral and competent behaviors are more diagnostic and are discounted less than moral and incompetent behaviors. Interestingly, neutral information elicits evaluations and behaviors that, while close to the scales' midpoints, are closer to the evaluations and behaviors elicited by positive than negative information. This is consistent with positivity offset accounts of the Evaluative Space Model (Cacioppo & Berntson, 1994; Cacioppo, Gardner & Berntson, 1999), suggesting that, in the dual systems that are responsible for positive and negative affect evaluations, absence of negativity per se is seen as positive.

Our findings contribute to the understanding of impression formation in adding an important factor that drives overall impressions. The early research in impression formation postulated and showed that negative and extreme behaviors have a stronger impact on impression formation than positive and moderate behaviors (e.g., Fiske, 1980). Skowronski and Carlston (1987, 1989) added diagnosticity as another critical factor: A negativity bias is observed for morality judgments because here negative behaviors are more diagnostic than positive behaviors, but a positivity bias is found for judgments of ability because here positive behaviors are more diagnostic than negative behaviors. Our research suggest a third critical factor: time. Diagnostic past behaviors (immoral, greedy, and intelligent, competent behaviors) are discounted less than less diagnostic past behaviors (e.g., moral, fair, and stupid, incompetent behaviors).

There are two noteworthy aspects about time. First, in our studies the time that had elapsed between the occurrence of an immoral/greedy behavior and its judgement was perceived in relative terms. In Experiments 1 and 2, the time difference was 4 years (5 years ago vs. 12 months ago); in Experiments 3 and 4 it was about a minute (round 6 vs. round 2). Therefore, what seems to matter is how much relevant information is provided in-between, not the absolute amount of time that has elapsed. It might be interesting in future research to test whether the amount of relevant information provided in between a target person's behaviors and their judgments changes the rate of discounting.

Second, the passage of time was described to participants rather than experienced by them. Even in Experiments 3 and 4, where participants were sequentially presented with ostensibly real-time allocations of a player, they were finally provided with a table that showed all past 7 allocations of that player. If elapsed time was experienced rather than described to participants, what would discounting of information look like? Would immoral behaviors still be discounted less than moral behaviors, and competent behaviors be discounted less than incompetent ones? Likewise, if behaviors were experienced rather than described, would immoral and competent behaviors be discounted less than moral and incompetent behaviors? One difficulty in studying discounting of experienced behaviors over experienced time is that different behaviors may be more or less likely to be observed, encoded, and later on retrieved from memory. Given that negative and extreme stimuli are more attention-grabbing and memorable (Fiske, 1980; Pratto & John, 1991), the observed effects of discounting of information may be exacerbated, making immoral behaviors even more influential over time than moral behaviors. Disentangling the effects of perceiving, encoding, retrieving, and weighting information over time seems to be an ambitious but certainly fruitful avenue for future research.

Another potentially fruitful avenue for future research maybe to test whether discounting of moral and incompetent behaviors applies equally to cooperative scenarios. We tested discounting in competitive (dictator games, experiments 3 and 4) and in non-interactive scenarios (experiments 1 and 2). In cooperative contexts such as teamwork, when one has to choose a team member whose past observed performance or morality directly affects the benefits for the entire team, will differential discounting occur across domains? When choosing a team member for a project, would low ability be discounted more than high ability? Or would generosity be discounted more than greediness? Would it matter whether the unintelligent or moral behavior was performed in the nearer or the farther past? Cue-diagnosticity theory would suggest that the results will depend on the specific project at hand: if it is an ability-based task, such as a general knowledge quiz, one would expect ability to be more diagnostic than morality. Thus, one would expect lesser discounting of the former. Vice versa, if the task requires high moral standards, one would expect lesser discounting of morality. But what about a task that requires both ability and morality, such as a group activity in school, which tests knowledge and requires academic integrity? Given that a person's morality is more socially relevant than a person's competence (Cacioppo et al., 1997; Fiske et al., 2007; Fiske, Cuddy, Glick, & Xu, 2002; Peeters, 2002; Reeder, Kumar, Hesson-McInnis, & Trafimow, 2002; Skowronski, 2002), discounting of moral behaviors may be more pronounced than discounting of incompetent behaviors.

Besides contributing to the theory of impression formation, our work also sheds light on the social psychology of trust, and the reasons behind its inherent quality of being a fragile construct, which is built slowly over time but can be destroyed by one single mistake (Rothbark & Park, 1986; Slovic, 1993). Trustworthy behaviors are discounted more than untrustworthy ones because they are perceived as less diagnostic of one's character. Consequently, a past history of trustworthiness could be quickly discounted in light of one single mishap. Similarly, a reputation built with prolonged commitment to truth and morality over time may be abruptly destroyed by one slip into unethical behaviors.

These findings are especially relevant in our current, digital times, as many impressions are formed based on information easily retrieved online. Interactions with technology are bound to affect impression formation mechanisms in ways that, we uncovered, may lead to potential biases. Consider the implications of our findings for the job market, for instance. Many firms, before considering or hiring applicants, check their profiles on social networks (Acquisti & Fong, 2013; Brown & Vaughn, 2011; Wilson, Gosling, & Graham, 2012). By doing so, employers may err on the pessimistic side and overestimate the diagnosticity of information about past immoral/greedy behaviors when judging a person's character. Indeed, in response to the establishment of European citizens' "right to be forgotten,"[1] Google has established the possibility for individuals to request the removal of inaccurate or obsolete information. Similarly to social media background checks, many

[1] Find the ruling at http://ec.europa.eu/justice/data-protection/files/factsheets/factsheet_data_protection_en.pdf.

employers run criminal background checks on job applicants. Lower depreciation of negative as compared to positive information may result in unjustified discrimination against ex-convicts, even when the chances that they will recidivate are effectively no higher than a person with no history of conviction (Blumstein & Nakamura, 2009).

Whether it is to gather clues about a job candidate, or about a potential romantic partner, an acquaintance, a business partner, or a colleague, new digital technologies allow for effortless and immediate retrieval of digital information about the present, the past, and the distant past. As the history of somebody's life unfolds in front of our eyes in a series of snapshots, a Facebook timeline, or a set of Twitter feeds, after a search that lasted a fraction of a second, the actual passage of time may be turning less relevant than the timestamp of the information itself in determining how past positive and negative events affect impressions: different events occurred at different points in time, but we see them all crunched together in a snapshot, here and now. Thus, the new developments brought about by technological progress affect the way we form impressions about each other.

Understanding how the unique features of digital information and technological developments affect impression formation is crucial at a time when more and more behaviors and interactions, which used to take place in the physical world, occur in the digital world instead. Consequences from differential discounting of past negative and positive information may reach far beyond impression formation.

References

Acquisti, A. & Fong, C. M. (2013). An experiment in hiring discrimination via online social networks. http://dx.doi.org/10.2139/ssrn.2031979.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, *5*(4), 323-370.

Benzion, U., Rapoport, A., & Yagil, J. (1989). Discount rates inferred from decisions: An experimental study. *Management Science 35*, 270-284.

Blumstein, A., & Nakamura, K. (2009). Redemption in the presence of widespread criminal background checks. *Criminology*, *47*(2), 327-359.

Brickman, P., Coates, D., & Janoff-Bulman, R. (1978). Lottery winners and accident victims: Is happiness relative? *Journal of Personality and Social Psychology, 36*(8), 917-927.

Brown, V. R., & Vaughn, E. D. (2011). The writing on the (Facebook) wall: The use of social networking sites in hiring decisions. *Journal of Business and Psychology*, *26*(2), 219-225.

Cacioppo, J. T., & Berntson, G. G. (1994). Relationship between attitudes and evaluative space: A critical review, with emphasis on the separability of positive and negative substrates. *Psychological Bulletin*, *115*(3), 401-423.

Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1997). Beyond bipolar conceptualizations and measures: The case of attitudes and evaluative space. *Personality and Social Psychology Review*, *1*(1), 3-25.

Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1999). The affect system has parallel and integrative processing components: Form follows function. *Journal of Personality and Social Psychology*, *76*(5), 839-855.

Cahill, C., Llewelyn, S. P., & Pearson, C. (1991). Long-term effects of sexual abuse which occurred in childhood: A review. *British Journal of Clinical Psychology, 30*, 117-130.

De Quervain, D.J., Fischbacher, U., Treyer, V., & Schellhammer, M. (2004). The neural basis of altruistic punishment. *Science*, *305*(5688), 1254-1258.

Dweck, C.S., Chiu, C., &Hong, Y. (1995). Implicit theories and their role in judgments and reactions: A world from two perspectives. *Psychological Inquiry, 6*, 267-285.

Ersner-Hershfield, H., Garton, M. T., Ballard, K., Samanez-Larkin, G. R., & Knutson, B. (2009). Don't stop thinking about tomorrow: Individual differences in future self-continuity account for saving. *Judgment and Decision Making*, *4*(4), 280-286.

Ersner-Hershfield, H., Wimmer, G. E., & Knutson, B. (2009). Saving for the future self: Neural measures of future self-continuity predict temporal discounting. *Social Cognitive and Affective Neuroscience*, *4*(1), 85-92.

Fiske, S. T. (1980). Attention and weight in person perception: The impact of negative and extreme behavior. *Journal of Personality and Social Psychology*, *38*(6), 889-906.

Fiske, S. T., Cuddy, A. J., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences, 11*(2), 77-83.

Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, *82*(6), 878-902.

Fleming, J., Mullen, P. E., Sibthorpe, B., & Bammer, G. (1999). The long-term impact of childhood sexual abuse in Australian women. *Child Abuse and Neglect*, *23*, 145-159.

Forsythe, R., Horowitz, J. L., Savin, N. E., & Sefton, M. (1994). Fairness in simple bargaining experiments. *Games and Economic Behavior*, *6*(3), 347-369.

Frederick, S. & Loewenstein, G. (1999). Hedonic adaptation. In D. Kahneman, E. Diener, & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology* (pp. 302-329). Russel Sage, New York, NY.

Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002). Time discounting and time preference: A critical Review. *Journal of Economic Literature, 40*(2), 351-401.

Gardner, W. L., Pickett, C. L., & Brewer, M. B. (2000). Social exclusion and selective memory: How the need to belong influences memory for social events. *Personality and Social Psychology Bulletin*, *26*(4), 486-496.

Gilbert, D. T., Pinel, E. C., Wilson, T. D., Blumberg, S. J., & Wheatley, T. (1998). Immune neglect: A source of durability bias in affective forecasting. *Journal of Personality and Social Psychology, 75,* 617–638.

Gilbert, D. T., & Wilson, T. D. (2007). Prospection: Experiencing the future. *Science, 317,* 1351–1354.

Goodwin, G. P., Piazza, J., & Rozin, P. (2014). Moral character predominates in person perception and evaluation. *Journal of Personality and Social Psychology*, *106*(1), 148-168.

Ito, T. A., Larsen, J. T., Smith, N. K., & Cacioppo, J. T. (1998). Negative information weighs more heavily on the brain: The negativity bias in evaluative categorizations. *Journal of Personality and Social Psychology*, *75*(4), 887-900.

Jones, E. E., & Davis, K. E. (1965). From acts to dispositions: The attribution process in person perception. *Advances in Experimental Social Psychology*, *2*, 219-266.

Jones, E. E., & McGillis, D. (1976). Correspondent inferences and the attribution cube: A comparative reappraisal. *New directions in Attribution Research*, *1*, 389-420.

Krizan, Z., & Windschitl, P. D. (2007). The influence of outcome desirability on optimism. *Psychological Bulletin*, *133*(1), 95-121.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*(3), 480-498.

Levy, S.R., Plaks, J.E., Hong, Y., Chiu, C,. & Dweck, C.S. (2001). Static versus dynamic theories and the perception of groups: Different routes to different destinations. *Personality & Social Psychology Review*, *5*, 156-168.

Loewenstein, G. (1987). Anticipation and the valuation of delayed consumption. *Economic Journal, 97*, 666-684.

McGuire, J. T., & Kable, J. W. (2015). Medial prefrontal cortical activity reflects dynamic re-evaluation during voluntary persistence. *Nature Neuroscience*, *18*(5), 760-766.

Moore, D. A., & Healy, P. J. (2008). The trouble with overconfidence. *Psychological Review, 115*(2), 502-517.

Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making*, *5*(5), 411-419.

Peeters, G. (1971). The positive–negative asymmetry: On cognitive consistency and positivity bias. *European Journal of Social Psychology*, *1*, 455–474.

Peeters, G. (1983). Relational and informational patterns in social cognition. In W. Doise & S. Moscovici (Eds.), *Current Issues in European Social Psychology* (Vol. 1, pp. 201–237). Cambridge: Cambridge University Press.

Peeters, G. (2002). From good and bad to can and must: Subjective necessity of acts associated with positively and negatively valued stimuli. *European Journal of Social Psychology*, *32*(1), 125-136.

Peeters, G., & Czapinski, J. (1990). Positive-negative asymmetry in evaluations: The distinction between affective and informational negativity effects. *European Review of Social Psychology*, *1*(1), 33-60.

Pizarro, D. A., & Tannenbaum, D. (2011). Bringing character back: How the motivation to evaluate character influences judgments of moral blame. In M. Mikulincer & P. R. Shaver (Eds.), *The social psychology of morality: Exploring the causes of good and evil* (pp. 91-108). APA Press, Washington, DC.

Pratto, F., & John, O. P. (1991). Automatic vigilance: The attention-grabbing power of negative social information. *Journal of Personality and Social Psychology*, *61*(3), 380-391.

Reeder, G. D., & Brewer, M. B. (1979). A schematic model of dispositional attribution in interpersonal perception. *Psychological Review*, *86*(1), 61-79.

Reeder, G. D., Kumar, S., Hesson-McInnis, M. S., & Trafimow, D. (2002). Inferences about the morality of an aggressor: The role of perceived motive. *Journal of Personality and Social Psychology*, *83*(4), 789-803.

Rosen, J. (2010). The web means the end of forgetting. *The New York Times*, p. MM30 (July 21).

Rothbart, M., & Park, B. (1986). On the confirmability and disconfirmability of trait concepts. *Journal of Personality and Social Psychology*, *50*(1), 131-142.

Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review*, *5*(4), 296-320.

Shakespeare, W. (1591). Henry VI, Part 3, Act 4, Scene 4.

Silver, R. L., Boon, C., & Stones, M. H. (1983). Searching for meaning in misfortune: Making sense of incest. *Journal of Social Issues, 39,* 81-102.

Skowronski, J. J. (2002). Honesty and intelligence judgments of individuals and groups: The effects of entity-related behavior diagnosticity and implicit theories. *Social Cognition*, *20*(2), 136-169.

Skowronski, J. J., & Carlston, D. E. (1987). Social judgment and social memory: The role of cue diagnosticity in negativity, positivity, and extremity Biases. *Journal of Personality and Social Psychology*, *52*(4), 689-699.

Skowronski, J. J., & Carlston, D. E. (1989). Negativity and extremity biases in impression formation: A review of explanations. *Psychological Bulletin*, *105*(1), 131-142.

Slovic, P. (1993). Perceived risk, trust, and democracy. *Risk Analysis*, *13*(6), 675-682.

Styron, T., & Janoff-Bulman, R. (1997). Childhood attachment and abuse: Long-term effects on adult attachment, depression and conflict resolution. *Child Abuse and Neglect*, *21*, 1015-1023.

Tausch, N., Kenworthy, J. B., & Hewstone, M. (2007). The confirmability and disconfirmability of trait concepts revisited: Does content matter? *Journal of Personality and Social Psychology*, *92*(3), 542-556.

Thaler, R. H. (1981). Some empirical evidence on dynamic inconsistency. *Economic Letters, 8,* 201-207.

Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Inferences of competence from faces predict election outcomes. *Science, 308* (5728), 1623-1626.

Tversky, A., & Kahneman, D. (1974). Decision making under uncertainty: Heuristics and biases. *Science*, *185*, 1124-1130.

Vosgerau, J. (2010), How prevalent is wishful thinking? Misattribution of arousal causes optimism and pessimism in subjective probabilities. *Journal of Experimental Psychology: General*, *139*(1), 32- 48.

Walker, W. R., Skowronski, J. J., & Thompson, C. P. (2003). Life is pleasant—and memory helps to keep it that way! *Review of General Psychology*, *7*(2), 203-210.

Weber, E. (1994). From subjective probabilities to decision weights: The effect of asymmetric loss functions on the evaluation of uncertain outcomes and events. *Psychological Bulletin*, *115*(2), 228-242.

Weiss, E. L., Longhurst, J. G., & Mazure, C. M. (1999). Childhood sexual abuse as a risk factor for depression in women: Psychosocial and neurobiological correlates. *American Journal of Psychiatry*, *156*, 816-828.

Wilson, A. E., & Ross, M. (2001). From chump to champ: People's appraisals of their earlier and present selves. *Journal of Personality and Social Psychology*, *80*(4), 572-584.

Wilson, R. E., Gosling, S. D., & Graham, L. T. (2012). A review of Facebook research in the social sciences. *Perspectives on Psychological Science*, *7*(3), 203-220.

Wojciszke, B. (2005a). Affective concomitants of information on morality and competence. *European Psychologist*, *10*(1), 60-70.

Wojciszke, B. (2005b). Morality and competence in person-and self-perception. *European Review of Social Psychology*, *16*(1), 155-188.

Yechiam, E. & Hochman, G. (2013). Losses as modulators of attention: Review and analysis of the unique effects of losses over gains. *Psychological Bulletin*, *139*(2), 497-518.

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, *35*(2), 151-175.