AGAINST THE ODDS: PRESCHOOLERS, LIKE ADULTS, PREDICT OUTCOMES THAT ARE DESIRABLE BUT UNLIKELY

A Thesis
by
ZACHARY HOLLINGSWORTH MORGAN

Submitted to the Graduate School
at Appalachian State University
in partial fulfillment of the requirements for the degree of
MASTER OF ARTS

May 2016
Department of Psychology

AGAINST THE ODDS: PRESCHOOLERS, LIKE ADULTS, PREDICT OUTCOMES THAT ARE DESIRABLE BUT UNLIKELY

A Thesis
by
ZACHARY HOLLINGSWORTH MORGAN
May 2016

APPROVED BY:

Robyn L. Kondrad, Ph.D.
Chairperson, Thesis Committee

Andrew R. Smith, Ph.D.
Member, Thesis Committee

Twila Wingrove, Ph.D.
Member, Thesis Committee

James C. Denniston, Ph.D.
Chairperson, Department of Psychology

Max C. Poole, Ph.D.
Dean, Cratis D. Williams School of Graduate Studies

Abstract

AGAINST THE ODDS: PRESCHOOLERS, LIKE ADULTS, PREDICT OUTCOMES THAT ARE DESIRABLE BUT UNLIKELY

Zachary Hollingsworth Morgan
B.A., Appalachian State University
M.A., Appalachian State University

Chairperson: Robyn L. Kondrad, Ph.D.

Adults' expectations for future outcomes are often biased by their desires, a phenomenon known as the desirability bias. The current study was the first to investigate the desirability bias in 4- and 5-year-old children and adults using a modified version of the marked card paradigm. In the paradigm, participants predicted whether a critical card would be drawn from decks containing 10 cards. Decks contained one of five different ratios of critical to noncritical cards (1:9, 4:6, 5:5, 6:4, and 9:1). For some decks, drawing a critical card resulted in gaining a reward and in other decks, it resulted in a loss. Results revealed that children, like adults, exhibited a desirability bias. Furthermore, children showed a stronger bias and were more likely than adults to predict a critical card when drawing one resulted in a gain and less likely to do so when it resulted in a loss. Children, but not adults, showed a desirability bias for potential losses. Adults tended to use a strategy that minimized losses by predicting undesirable outcomes to occur when there was a potential loss at stake. Future research should explore the role that the relative value of accuracy incentives has on the desirability bias and investigate the presence of

the bias in infants and older adults in order to gain insight into how and why the strength of the desirability bias changes throughout the lifespan.

Acknowledgments

I would like to offer my deepest thanks to everyone who helped me with this thesis. Robyn, thank you for everything you have done from day one (and thanks for everything you'll do in the future while I'm at it). I couldn't have done any of this, and likely wouldn't have even gotten into developmental research, if it weren't for you. I look forward to continuing to collaborate with you as a researcher and a friend. Also, thanks to my stats guru and committee member, Dr. Andrew Smith, as well Dr. Twila Wingrove. You have both been enormously influential to my academic career. Thanks to Shauna, for helping me out with running participants, to Eva and my grandmother for helping me assemble my materials, and to all the adults and children that participated and allowed their children to participate in this research. Lastly, I would like to thank the Office of Student Research, the family of Wiley F. Smith, and the Graduate Student Association Senate of Appalachian State University for helping fund this research as well as my travels.

Table of Contents

List of Figures

Against the Odds: Preschoolers, like Adults, Predict Outcomes that are Desirable But Unlikely

Zachary H. Morgan

Appalachian State University

Abstract

Adults' expectations for future outcomes are often biased by their desires, a phenomenon known as the desirability bias. The current study was the first to investigate the desirability bias in 4- and 5-year-old children and adults using a modified version of the marked card paradigm. In the paradigm, participants predicted whether a critical card would be drawn from decks containing 10 cards. Decks contained one of five different ratios of critical to noncritical cards (1:9, 4:6, 5:5, 6:4, and 9:1). For some decks, drawing a critical card resulted in gaining a reward and in other decks, it resulted in a loss. Results revealed that children, like adults, exhibited a desirability bias. Furthermore, children showed a stronger bias and were more likely than adults to predict a critical card when drawing one resulted in a gain and less likely to do so when it resulted in a loss. Children, but not adults, showed a desirability bias for potential losses. Adults tended to use a strategy that minimized losses by predicting undesirable outcomes to occur when there was a potential loss at stake. Future research should explore the role that the relative value of accuracy incentives has on the desirability bias and investigate the presence of the bias in infants and older adults in order to gain insight into how and why the strength of the desirability bias changes throughout the lifespan.

Against the Odds: Preschoolers, like Adults, Predict Outcomes that are Desirable but Unlikely

Having a preference makes it difficult to make unbiased predictions about future outcomes. For example, adults tend to bet on their preferred sports team to win even when they know the odds are against them (Babad, 1987), and expect their preferred politician to be elected even if they are behind in the polls (Krizan, Miller, & Johar, 2010). Adults' predictions are biased by their desires even when they are explicitly told to be objective (e.g., Babad & Katz, 1991), incentivized to be accurate (e.g., Irwin & Graae, 1968), when they know the odds are against them (e.g., Krizan et al., 2010), and when they have read scientific evidence pointing to the contrary (e.g., Bastardi, Uhlmann, & Ross, 2011). The idea that people's desires influence their expectations about a future outcome is called the *desirability bias* (DB).

The link between preferences and expectations is a well-documented phenomenon in social psychology (Babad, 1987; Granberg & Brent, 1983; Hayes, 1936; Lee, Frederick, & Ariely, 2006; Windschitl, Smith, Rose, & Krizan, 2010), but despite speculation about why the DB might exist (see Windschitl et al., 2010), there has been little exploration of its ontogenesis. The primary goals of the present study were to examine whether young children, like adults, exhibit a DB, and to compare the strength of the DB in children and adults. The first section of the introduction reviews studies that use the marked card paradigm to measure the DB in adults, the second section outlines relevant developmental literature, and the third describes the modified marked card paradigm that was created for the current study to measure the DB in both children and adults.

The Desirability Bias in Adults

Adults have been found to show a DB in studies that have used certain games of chance (Windschitl et al., 2010). A card-drawing task known as the marked card paradigm is particularly effective at detecting the bias (Budescu & Bruderman, 1995; Crandall, Solomon, & Kellaway, 1955; Irwin, 1953; Marks, 1951, Windschitl et al., 2010). In the marked card paradigm, participants are asked to predict whether a particular type of card, called the critical card, will be randomly drawn from a deck made up of two types of cards. Participants first learn how many cards in a deck are critical, and that drawing one of the critical cards will result in either a gain or loss of money. They are also told that they are able to gain a small amount of money for accurately predicting which card is drawn. After this, participants predict which card, critical or noncritical, will be drawn. Ideally, predictions should be based on probability: the more critical cards there are in a deck, the greater the chance of drawing one. Based on expected value, if more than 50% of the cards are critical, the optimal prediction is always to guess that a critical card will be drawn.

In reality, participants are biased to predict desirable outcomes. If drawing a critical card automatically results in a positive outcome (i.e., gaining a point), participants tend to predict that a critical card will be drawn even when this is the objectively suboptimal choice. For example, in a deck containing four critical and six noncritical cards, the optimal prediction (i.e., the prediction with the greatest expected value) is that a noncritical card will be drawn. However, if drawing a critical card is desirable because it results in a gain of money or points, then adults tend to predict that a critical card will be drawn (Windschitl et al., 2010). In other words, adults' predictions are based more on their desired outcome than on the optimal outcome.

The strength of a DB that is elicited using the marked card paradigm can be influenced by a number of factors. Of interest for this study are the frequency of critical to noncritical cards, the use of accuracy incentives, and the degree of control participants believe they have over the drawn card (i.e., the illusion of control).

Frequency of critical and noncritical cards. The number of critical and noncritical cards in a deck can influence the strength of the DB. The smaller the difference in the frequency of critical versus noncritical cards in a deck, the more likely it is that adults' desires will influence their predictions. For instance, in a deck with three critical and seven noncritical cards the optimal prediction is always that a noncritical card will be drawn. But when the critical card is associated with a reward, adults predict a critical card to be drawn between 15% and 30% of the time (Crandall et al., 1955; Irwin, 1953; Irwin & Metzger, 1966; Windschitl et al., 2010, study1). In contrast, when the difference in frequency is greater, such as with one critical and nine noncritical cards, adults are less likely to base their predictions on the desirability of the critical card. For a deck with this frequency, it is also optimal to always predict a noncritical card to be drawn, and adults do so between 90% and 100% of the time (Crandall et al., 1955; Irwin, 1953; Irwin & Metzger, 1966).

Interestingly, there is no optimal prediction in decks with equal numbers of critical and noncritical cards (5:5); the expected value is the same regardless of which type of card is predicted. Adults' predictions in this circumstance reflect the value of a critical card. When a critical card is associated with a gain, adults are likely to predict that one will be drawn and when it represents a loss, adults are unlikely to predict one to be drawn (Crandall et al., 1955; Irwin & Metzger, 1966; Windschitl et al., 2010). Thus, the 5:5 ratio provides a baseline measure for the degree to which adults' predictions are influenced by desire when probability provides no directional information for making a prediction.

The opposite relationship is examined through the use of a neutral condition, in which neither the critical nor the noncritical cards are given a value. The neutral condition provides a baseline measure of participants' sensitivity to frequency information when there is no desirability information provided. In it, participants are rewarded only for making an accurate prediction. Because the critical card is not made more or less desirable than noncritical cards, predictions should be based on whatever is the optimal outcome given the odds ratio. Having a baseline condition for each ratio also provides a way to determine whether adults are influenced differently by desirable outcomes than by undesirable outcomes.

Comparing the difference between the neutral condition and the reward conditions is an interesting assessment because there is abundant literature that suggests that adults are more sensitive to losses than they are to gains (Harbaugh, Krause, & Vesterlund, 2002; Johnson, Herrmann, & Terry, 2007; Kahneman & Tversky, 1984; Kahneman, Knetsch, & Thaler, 1991; Steelandt, Broihanne, R omain, Thierry, & Dufour, 2013; Tom, Fox, Trepel, & Poldrack, 2007). People who lose money if they fail to reach a weight loss goal, for example, are more likely to actually shed pounds than are those who earn money for achieving the same goal (Driver & Hensrud, 2013). In past research on the DB, however, adults were equally influenced by both consequences – they overpredicted desirable outcomes to about the same extent in the gain and loss conditions compared to a neutral condition. (Irwin, 1953; Irwin & Metzger, 1966; Irwin & Snodgrass, 1966; Windschitl et al., 2010, study 1).

Accuracy incentives. Some studies using the marked card paradigm have provided participants with incentives for making accurate predictions (Irwin & Metzger, 1966; Windschitl et al., 2010), whereas others have not (Irwin, 1953; Morlock & Hertz, 1964). A meta-analysis of 27 individual studies showed that accuracy incentives produced effect sizes that were not significantly different from those of studies that did not use accuracy incentives (Krizan & Windschitl, 2007b). In other words, giving incentives for making accurate predictions did not appear to influence the strength of the DB in the participants from the included studies.

Despite evidence suggesting that accuracy incentives do not have a big impact on eliciting the DB, there are some good arguments to be made for including these incentives. Importantly, the conclusions drawn from the meta-analysis mentioned above are limited because the authors included studies that used procedures other than the marked card paradigm. Furthermore, because of the small sample, Krizan and Windschitl (2007b) considered studies that stressed the importance of accuracy in the paradigm instructions and studies that provided monetary incentives as one category. On the one hand, it is not clear whether a physical incentive like money has a different effect than an instructional incentive. On the other hand, including an incentive of any type does not appear to limit the ability to detect a DB.

The advantage of including an accuracy incentive is that it makes interpreting the responses clearer. Imagine a participant who predicts a desirable card when there is only a 10% chance of it being drawn. If there is no incentive for making an accurate prediction, then the participant is not risking losing out on that point (because it is not likely that a critical card will actually be drawn). In other words, there is no strategic reason to not predict the unlikely but desirable critical card. But when accuracy incentives are included, then the participant has to decide whether to risk not getting the reward for accuracy. Thus, if a DB is detected even with an accuracy incentive, then it is clear that participants are swayed by their desires.

Illusion of control. Another factor that can influence the strength of the DB is when people incorrectly believe that they have some control over an unknown future outcome, a phenomenon called the illusion of control (Langer, 1975). Some scholars suggest that there is a positive relationship between the illusion of control and the strength of the DB: If participants in the marked card paradigm believe they have some control over what card is drawn from a deck, then they are even more likely to predict improbable but desirable outcomes (i.e., they show a stronger DB; Langer & Roth, 1975).

Other researchers argue instead that the illusion of control is distinct from and uncorrelated with the DB. In one study, participants either drew the cards themselves or an experimenter did so. Although participants who “controlled” drawing the cards reported being more confident about their predictions, both groups demonstrated an equally robust DB (Budescu & Bruderman, 1995, study 1). The participants’ illusion of control eventually disappeared as they made more predictions, but they continued to show a DB. In fact, studies have found a DB when the illusion of control was removed altogether by having an experimenter or a computer program randomly draw the cards for each deck (Krizan & Windschitl, 2007b; Windschitl et al., 2010). These findings suggest that, in adults, the DB is not diminished when participants could potentially have the illusion of control over random outcomes.

Advantages of Using the Marked Card Paradigm

For the present study, a child-friendly version of the marked card paradigm was created to measure the DB in children and adults. This procedure was chosen because it has proven to be more reliable than other procedures at eliciting a DB in adults (e.g., Bar-Hillel & Budescu, 1995; Bar-Hillel, Budescu, & Amar, 2008; Pruitt & Hoge, 1965; Windschitl et al., 2010). There are two main reasons why the marked card paradigm consistently detects a DB: First, participants make discrete outcome predictions rather than likelihood judgments and second, participants make predictions about stochastic (chance) rather than non-stochastic outcomes.

Discrete outcome prediction vs. likelihood judgment. In the marked card paradigm, participants are asked to make discrete outcome predictions. That is, they make a prediction about which of two types of cards will be drawn. When participants are asked to provide a probability estimate of how likely a given outcome is to occur by providing a percentage out of 100 instead of a dichotomous prediction, the DB disappears (e.g., Bar-Hillel & Budescu, 1995; Pruitt & Hoge, 1965). For example, in one study participants were shown a grid of black and white cells and were asked to estimate the probability of randomly drawing one of the colors (Bar-Hillel & Budescu, 1995, Study 1). For some trials, drawing the predetermined critical color was associated with a gain, and for other trials it was associated with a loss. A DB would exist if participants gave higher likelihood judgments when a card resulted in a gain than when it resulted in a loss. Unlike with the marked card paradigm, a DB was not detected using this task;

participants' estimates did not differ, regardless of the prospect of a gain or a loss, and were closely aligned with optimal predictions based on probability.

Why might the DB disappear when using likelihood estimates? Some researchers have theorized that likelihood judgments may lead participants to develop an expectation that the experimenter intends to measure their knowledge of probability (Krizan & Windschitl, 2007b). This expectation could inflate participants' reliance on probability information when making their estimates, reducing their reliance on desirability information and their overall DB. Regardless of the process involved, research using likelihood estimates has shown that this type of decision is not optimal for eliciting a DB.

Stochastic vs. nonstochastic outcomes. The second reason that the marked card paradigm elicits the DB more consistently than other paradigms is because it involves making predictions about stochastic outcomes. Stochastic outcomes are ones that occur randomly such as a dice roll or, in this case, a random card draw. Nonstochastic outcomes, in contrast, are not random (e.g., the answer to a multiple choice question) and are unlikely to elicit a DB (Bar-Hillel & Budescu, study 2, study 4; Krizan & Windschitl, 2007b; Windschitl et al., 2010, study 2).

One theory to explain this difference, referred to as the biased-guessing account (Windschitl et al., 2010), suggests that the DB becomes increasingly pronounced when participants rely more on guessing than on objective information (Windschitl et al., 2010). To illustrate this point, in decks with a 5:5 critical to noncritical card ratio, the objective probability information is not helpful because there is an equal chance of either type of card being drawn. In this case, participants must guess about which outcome will occur, and their guesses tend to align with the more desirable outcome. Furthermore, when participants are specifically encouraged to trust their intuition, the DB is enhanced (the opposite is also true – when guessing is discouraged, the DB is reduced; Windschitl et al., 2010, study 5). In other words, when participants are encouraged to guess or are forced to guess because objective evidence is not available, the DB is more pronounced.

In contrast, the bias is less pronounced as the deck frequencies becomes more differentiated and the uncertainty about the outcome is reduced, as in the 1:9 or 3:7 ratios used in the marked card paradigm (Crandall et al., 1955; Irwin, 1953). For these ratios, participants are more likely to base their predictions on the probability associated with drawing either card, and the DB is diminished (Bar-Hillel & Budescu, 1995, Study 1; Pruitt & Hoge, 1965). The DB is reduced for a similar reason when participants make predictions about non-stochastic trivia questions (Krizan & Windschitl, 2007b; Windschitl et al., 2010, study 2). When adults are told that they can gain money for correctly answering a trivia question (e.g., which animal is louder - A bear or a lion?) and that they can also gain money simply for one of the two answers being correct, they are only swayed by the arbitrary desirability of one of the answers if they are clueless as to what the correct answer might be. For a question that is somewhat familiar to participants such as the one above, most participants will be swayed on way or the other by past knowledge or inference. This type of question, therefore, typically leads to participants focusing less on the desirability of the critical card. This explanation is further supported by the fact that the DB returns when the trivia questions are extremely difficult and unfamiliar to the majority of people (Windschitl et al., 2010, study 3). In summary, a DB is most easily detected through the use of discrete outcome predictions about stochastic events.

Developmental Constructs Related to the DB

Though a number of studies have examined the DB in adults, none have explored its ontogenesis. There has only been one study to measure the DB in children (Marks, 1951). This study was the first to use the marked card paradigm. A DB was detected in children in the study, however the youngest participants were only 9-years-old and there was no adult group for comparison.

A direct comparison between the children in Marks' (1951) study and adults in later studies is difficult because more current studies used slightly different versions of the marked card paradigm than did Marks. Specifically, three methodological concerns limit the conclusions that can be drawn about children from Marks' early study. First, Marks did not provide accuracy incentives. Although accuracy incentives have been shown to minimally influence the way adults make predictions, the same may not be true for children. Second, a neutral condition was not included so it was impossible to determine whether the DB was stronger when children faced a potential gain compared to a potential loss. And third, the noncritical cards used in Marks' study were blank instead of containing a distinct pattern. Research suggests that this procedure might bias predictions towards the critical, non-blank card (Windschitl et al., 2010). To avoid this bias, the marked card paradigm is now conducted using two distinct cards (e.g., a red and a blue) for each deck with the critical and noncritical cards randomly selected between the two types. Thus, the Marks (1951) study provides a starting point for exploring the ontogenesis of the DB, but leaves room for more research with younger participants and with a modified paradigm that addresses these potential issues.

The following sections will provide a brief review of evidence from the developmental literature that supports the hypotheses that preschool-age children will show a DB, and will be more influenced by desirability than adults. The first section explains that children are even more optimistic about future events than are adults (Lockhart, Chang, & Story, 2002; Schuster, Ruble, & Weinert, 1998), the second shows that, like adults, preschoolers make little distinction between desired outcomes and expected outcomes (Babad, 1987; Hayes, 1936; Lipko, Dunlosky, & Merriman, 2009, study 1; Schneider, 1998), and the third shows that children weigh positive information that confirms what they want to believe more heavily than they do negative, disconfirming information (Boseovski, 2010; Rholes & Ruble, 1986; Stipek, Roberts, & Sanborn, 1984). Evidence from these three fields of study suggests that children will be even more likely than adults to be swayed by the desirability of outcomes even though they are capable of understanding concepts related to frequency and probability differences (Denison & Xu, 2010b; Kushnir, Xu, & Wellman, 2010; Xu & Garcia, 2008).

Overoptimism. The DB is a form of a broader construct known as overoptimism. Someone might be overly optimistic if, for example, they have special knowledge or experiences that influence their expectations about the future. This is different from the DB, for which someone’s predictions are specifically influenced by the desire for a particular outcome and not by any other subjective influences. Adults, for instance, tend to be overly optimistic about how healthy they will be in the future compared to the average person (Rothman, Klein, & Weinstein, 1996). This would be an example of a DB only if this optimism resulted from the desire to be healthier than others; it would be general overoptimism if they think they will be healthier because of the knowledge that they eat healthier and exercise more than the average person. Similarly, adults who overestimate how long their romantic relationships will last are demonstrating overoptimism if their estimation is based off of the length of a previous relationship and a DB if it is based off of their desire to stay with their current partner (Buehler, Griffin, & Ross, 1995).

Although adults are clearly overly optimistic in a variety of circumstances, they are more realistic than children when making predictions about the outcome of future events (Boseovski, 2012; Droege & Stipek, 1993; Lockhart et al., 2002; Ross & Levy, 1958; Schuster et al., 1998, study 1). For example, in one study 5- and 6-year-olds, 7- to 9-year-olds, and adults were told about a 5- or 10-year-old storybook character who had an undesirable characteristic (i.e., was messy, excessively freckled, missing a finger) and who really "wished" to change this trait in a positive direction (e.g., "wished she could be neater/had no freckles/had five fingers;" Lockhart et al., 2002). When asked if the storybook characters would, by the age of 21, "continue to show the extreme negative trait," "become more like everyone else," or "change in the opposite, more positive direction," 5- to 6-year-olds were more likely to predict that the character would undergo a positive change, whereas 7- to 9-year-olds and adults were more realistic and tended to predict that the characters would continue to show the negative trait or become more like everyone else. Young children were optimistic even when the outcome was an impossible one, like growing a new finger.

This method measured overoptimism more generally, and not necessarily a DB. This is because participants could have based their predictions on factors unrelated to the DB, such as past experience (e.g., knowing a messy friend who became more organized) or special knowledge (e.g., knowing that a lizard can grow back its tail). Additionally, the participants were making predictions about storybook characters whose fate had no implication for themselves. Previous research has suggested that the DB is diminished in circumstances where the loss or gain is not relevant to the person making predictions (Price, 2000; Stipek et al., 1984). Finally, the predictions made were about a long-term future outcome, rather than a short-term one. Research with adults has suggested that predictions about outcomes far into the future tend to be more extreme and exaggerated than predictions made about outcomes in the immediate future (Buehler & McFarland, 2001; Gilbert, Pinel, Wilson, Blumberg, & Wheatley, 1998; Wilson & Gilbert, 2005). It is possible that children's predictions are also more extreme or exaggerated when thinking about long-term, rather than immediate outcomes. Nevertheless, this study suggests that preschoolers are more optimistic than adults, and supports the hypothesis that they may be more influenced by desirable outcomes than adults in a task that measures the DB.

In summary, young children, like adults, are overly optimistic in a variety of circumstances and, at least in some contexts, are even more optimistic about future outcomes than adults. Because overoptimism represents a more general concept compared to the more specific DB, the results of these studies suggest that children, like adults, will exhibit a DB. Just as they are more optimistic than adults, preschoolers may also be more willing to optimistically accept poor odds in favor of betting on a desired outcome in the marked card paradigm.

Desire-expectation differences. The DB occurs when people's expectations are shaped by their preferences. Overly optimistic predictions about future outcomes are less likely when people more accurately discriminate between what they should expect to occur and what they would like to occur. Adults are able to discriminate between these two factors in some situations, as demonstrated by the lack of a DB in the 1:9 ratio in the marked card paradigm. Here, they might desire the critical card to be drawn, but it is clear that this is not what they expect to occur. Under certain circumstances, however, adults are not as accurate as they could be in recognizing the distinction between preferences and expectations. Their desires and expectations are strongly correlated when, for instance, they expect their desired candidate to win in political elections (Hayes, 1936), expect their desired sports team to win a game (Babad, 1987), and expect their financial investments to pay off as desired (Olsen, 1997).

Young children, like adults, do not tend to differentiate their desires from their expectations (Lipko et al., 2009, study 1; Schneider, 1998). In one study, for instance, researchers asked half of a group of 4- and 6-year-old children to place a flag either where they *wished* to land in a long-jumping game, and asked the other half to place it where they *expected* to land (Schneider, 1998). Here, flag placements made by the “wish” and “expectation” groups differed for the 6-year-olds, but not for 4-year-olds. In a separate study, 4- and 5-year-olds also failed to differentiate between their desires and expectations in a memory task using 10 colored pictures (Lipko et al., 2009, study 1). When asked how many pictures they thought they would be able to remember, children expected to remember an average of three more pictures than they actually could. These results, related to children’s inflated expectations, suggest that young children often fail to differentiate what they expect to happen from what they want to happen. This finding supports the presence of a DB in children; if they are unable to separate their desires from expectations, then they will likely disregard probability information when it does not align with the most desirable outcome.

It is important to recognize that the tasks used in these studies did not measure the DB. Instead, children were asked to make predictions about their abilities, which can be influenced by subjective factors like perceived skill and past experience rather than just desire and objective information. Also, neither study mentioned included incentives for accurately predicting performance, which might have influenced children's expectations. Thus, although the correlation between preschoolers' desires and expectations supports the hypothesis that children will show a stronger DB than adults, a task specifically tailored to measuring the DB is necessary to support this claim.

Bias in social judgments. Further support that children might exhibit a DB comes from research on social judgment and decision-making. Preschoolers have a positivity bias when making personality judgments about others (see Boseovski, 2010 for a review). That is, they have a "people are good" mindset (see Kalish, 2002) that influences how they process behavioral information when making personality judgments.

Because they want to believe the best about other people, they require less evidence that someone is nice to make a positive trait attribution about them than evidence that someone is mean to make a negative attribution (Boseovski & Lee, 2006, 2008; Rholes & Ruble, 1986). Preschoolers have been shown to downplay extensive knowledge about an individual's past behavior in order to make a positive trait attribution, judging someone positively from a single instance of positive behavior, even if there were three prior instances of negative behavior (Rholes & Ruble, 1986). In fact, it takes at least five instances of negative behavior before children are willing to believe that the individual is a "bad person" (Boseovski & Lee, 2006). In other words, children selectively attend to information that aligns with what they expect of other people. In the framework of the marked card paradigm, children could selectively attend to information that aligns with their positive expectations. And since children are not good at differentiating their expectations from their desires, this selective attention to the desirable information would likely increase the strength of the DB.

Children's optimistic mentality also influences their predictions about how other people will behave in the future. They predict that someone who wants to make positive behavioral change will be able to do so, but will not be able to make negative changes (e.g., Heyman & Giles, 2004; Lockhart et al., 2002). For instance, 7-year-olds learned about a character that was nice or mean, but wanted to behave in the opposite way. The children claimed that mean characters could learn to be nice, but that nice characters could not learn to be mean even if that is what they wanted. Thus, children do not tend to make predictions that conflict with the belief that people are good.

Another way of thinking about the positivity bias is that young children will do just about anything to make facts conform to their "people are good" mentality, even if the objective evidence suggests otherwise. This mentality in children relates to the DB because it suggests that children may be willing to disregard probability information in favor of their desires when making decisions about the future. Furthermore, their decreased sensitivity to negative information may lead to children focusing more on desirable outcomes than undesirable ones, which would lead to a DB in the marked card paradigm.

Numeracy. Although a DB exists even when probability information is ambiguous, as in the 5:5 critical to noncritical card ratio, it also can be elicited in instances when the probability information clearly points to an optimal prediction that is different from the desired one, as in the 4:6 ratio. Results from the marked card paradigm are interesting because they show that adults base their predictions on desire even when they stand to gain more from basing them on probability. In the modified version of the marked card paradigm created for the present study, children are asked to compare desirability and probability information to make predictions. The ability to work with and understand numbers and probability, a concept known as numeracy, is necessary when making such a comparison.

Basic probability understanding has been documented in children as young as eight months of age (Xu & Garcia, 2008), and has been consistently reported in older infants and toddlers (Denison & Xu, 2010a, 2010b; Kushnir et al., 2010; Ma & Xu, 2011). For example, 14-month-old children used probability to determine where to search for their favorite type of lollipop. Infants saw two jars of lollipops, one with a 1:3 ratio of desirable to undesirable lollipops and one with a 3:1 ratio. An experimenter took one lollipop from each jar and, without revealing the type of lollipop, placed each one into two opaque cups, and invited infants to choose one of the two cups. The children nearly always searched in the cup where the experimenter had placed the lollipop she drew from the jar that contained more desirable lollipops (Denison & Xu, 2010b). Thus, even before they can talk, children can use basic probability information to guide their decision-making.

Preschoolers also use probability to make inferences about their environments (Kushnir & Gopnik, 2005). In one study, 4-year-olds were shown two blocks that could either activate or fail to activate a detector (a toy that lit up and played music) when placed on top of it (Kushnir & Gopnik, 2005, study 1). For each trial, both blocks were placed upon the detector three times and the number of times each block activated the detector was manipulated. Thus, children could see the general likelihood that each block would activate the detector. After this, participants were asked to "pick the best one to make it go." When children saw block A activate the detector three out of three times, they were much more likely to pick it over block B when block B only activated the detector one out of three times. This trend of making decisions based on probability was found for all trials that did not present ambiguous or conflicting information, showing that 4-year-olds are able to use probability to make decisions about future outcomes.

The evidence related to young children's numeracy suggests that they are sensitive to basic probabilities, and can use such information to make predictions and inform their behavioral decisions. Preschool-age children, therefore, should not have difficulty understanding the differences in critical to noncritical card ratios between decks. They should be able to understand that there are more of some cards than others, and that the more frequent cards are more likely to be drawn. The question of interest, however, is whether their predictions will be based more on this frequency information or on their desires.

Modified Marked Card Paradigm

The DB has never been investigated in preschool-age children. This study aims to measure the DB in children and to compare the strength of the DB in children and adults through the use of a child-friendly version of the marked card paradigm. There are three aspects of the traditional paradigm that the modified version created for the present study addresses in order to make the task appropriate for children: Cognitive load, attention span, and type of incentive.

Recall that, in the adult version of the marked card paradigm, participants are asked to predict whether a critical card will be drawn from a deck. Participants are first offered an incentive to be accurate. They are then told how many cards in a deck are critical cards and whether drawing one will result in either a gain or loss of money (or, in the neutral condition, that there is no value associated with the critical card). Finally, participants predict whether the critical card will be drawn. There are 10 trials for each of three conditions (a gain, loss, and neutral condition) randomly presented across a total of 30 trials.

Cognitive load. The traditional marked card paradigm is associated with a high cognitive load. That is, participants are asked to remember several different rules (e.g., critical card frequency, critical card value, and accuracy incentive) simultaneously. The task is further complicated because the rules change depending on the type of trial (e.g., sometimes participants gain points/stickers and sometimes they lose points/stickers for drawing a critical card). As it is, the task requires sophisticated working memory and executive functioning skills, which are much less developed in 4- to 5-year-old children than in adults (see Carlson, Moses, & Breton, 2002; Case, Kurland, & Goldberg, 1982; Marcovitch, Boseovski, & Knapp, 2007). Therefore, a primary goal of the modifications was to reduce the cognitive demand of the marked card paradigm without compromising its ability to detect a DB.

In the version of the paradigm used for the current study, the differentiation of critical and noncritical cards was simplified. The two cards differed by a single dimension (color: green and yellow) rather than by two dimensions (e.g., color and shape: green circles and yellow squares). There were also several changes made to simplify the procedure. The first was that test trials were block randomized by condition rather than fully randomized across all 30 trials. Participants completed 10 trials for a single condition before moving on to the next set of 10 trials for the next condition. Previous experiments with adults randomized the desirability condition (gain, neutral, or loss) associated with each deck on a per-trial basis. Thus, the desirability of the critical card could change from one deck to the next. Having the desirability condition remain the same for 10 trials in a row made it easier for children to remember and use this information when making their predictions.

A second procedural change was that the colors used to differentiate the critical and noncritical cards were changed after the first five trials for each condition. Previous studies used different types of cards for each deck. Fewer changes made the task less cognitively taxing for young children.

Lastly, participants went through three practice trials before completing the test trials for each condition. These trials allowed them to learn about the accuracy incentive and desirability information separately (first two practice trials), before combining both rules in the third practice trial. The experimenter also provided participants with a visual reminder (a reference sheet) during these trials to help explain each rule. Both rules were repeated to participants before every test trial. All together, these modifications helped to ensure that children were able to keep all of the pertinent information in mind when making predictions.

Attention span. One concern was that the marked card paradigm would be challenging to administer to young children due to their limited attention spans. Preschoolers have been found to have small capacities for sustained attention with a great deal of variability between participants (see Klenberg, Korkman, & Lahti-Nuuttila, 2001). The following modifications were made to help keep children engaged and attentive.

The marked card paradigm is typically computerized for adults who watch as information is displayed on a computer monitor, and then make predictions by clicking one of two options. In the child-friendly paradigm, participants interacted with an experimenter who used real decks of cards. Research has shown that children learn better from live, in-person interactions with real objects than they do from nonphysical interactions, like videos (e.g., the "video deficit" effect; Anderson & Pempek, 2005). Thus, the modified paradigm was presented using actual cards in order to help children focus.

Along with this change, two other procedural modifications were used to help keep children's attention. First, the opportunity to take a break was given after every fifth trial (when the experimenter switches to the next set of cards) as well every tenth trial (when the experimenter explains the rules for the next condition). In addition, the experimenter allowed children to do a short physical activity (e.g., "Let's stand up and do some jumping jacks") if they were especially distracted or fatigued.

Type of incentive. The monetary or point-based incentives typically used in the traditional version of the marked card paradigm were not appropriate for young children, who have a limited knowledge about the concept of money (see Berti & Bombi, 1981 for a review) and may not have any use for money themselves. Therefore, stickers were used as an incentive for children, which have been used successfully in previous studies (e.g., Jaswal, Croft, Setia, & Cole, 2010; Liu, Vanderbilt, & Heyman, 2013; Palmquist, Burns & Jaswal, 2012). Points that accumulated and could be exchanged for candy were used as an incentive for adults. Candy has been used as a successful motivator in previous studies for adults (Heyman & Ariely, 2004).

The Present Study

Adults consistently show a DB when making predictions about uncertain future outcomes (Crandall et al., 1955; Irwin, 1953; Irwin & Metzger, 1966; Windschitl et al., 2010). To date, one study has explored whether children also exhibit this bias (Marks, 1951), but no research has attempted to measure the bias in preschoolers. One reason for this gap is because there is currently no paradigm to measure the DB that is appropriate for this age group. Therefore, the goals of the present study were to measure the DB in 4- and 5-year-old children using a child-friendly version of the marked card paradigm, and to compare the strength of the DB in children versus adults. Because children are overly optimistic in a variety of contexts, fail to differentiate between their desires and expectations, and are biased to neglect probability information in favor of their positive desires, it was hypothesized that 4-and 5-year-olds would exhibit a stronger DB than adults.

Method

Participants

A total of 120 participants was used for this study: 30 4-year-old children (*M* = 55.1 months), 30 5-year-old children (*M* = 64 months) and 60 adult college students (*M* = 20.0 years). Equal numbers of males and females were recruited for both age groups. A statistical power analysis using G*Power (Faul, Erdfelder, Buchner, & Lang, 2009) determined that a total sample of at least 110 would be required to detect the largest interaction effect of interest (f(V) = .25, α = .05, β = .95). Children were recruited from local preschools that agreed to participate and from a local community database made up of community members who agreed to be contacted about participating in research. Adults were recruited from the psychology participant recruit system. They participated to fulfill a course requirement and received candy for participating. Children received stickers and a small thank-you gift (e.g., a ball, t-shirt). Both the adult and child samples were primarily white and middle-class. Two 4-year-old children were excluded because they did not complete at least half of the test trials.

Design

This study used a 3 (Desirability: gain, loss, neutral) x 5 (Frequency: 1:9, 4:6, 5:5, 6:4, 9:1) x 2 (Age group: children, adults) within-subjects design with age group as a between-subjects factor. The dependent variable was whether participants predicted a critical card to be drawn.

All participants completed a total of 30 test trials, with 10 trials in each of three conditions (gain, loss, and neutral). Three practice trials took place before the gain and loss condition, and two practice trials before the neutral condition. The order in which participants completed each condition was counterbalanced such that one-third of the participants completed the gain condition first, one-third the loss condition first, and one-third the neutral condition. The critical card color for each set of five decks (e.g., red or yellow) was also counterbalanced between participants. Each condition consisted of two trials for each of the five ratios (1:9, 4:6, 5:5, 6:4, 9:1). These trials were semi-randomly ordered such that each ratio appeared once in the first five trials, and once in the second five trials per condition.

Materials

Thirty card decks, each with 10 cards (6.3 x 8.8 cm), were used, one for each test trial. There were two colors of cards in each deck (one representing critical cards and the other color representing noncritical cards) in one of the following combinations: red cards paired with yellow cards, pink and purple, orange and brown, black and white, blue and green, stripes and black polka-dots (both in black and white). Reference sheets (an 8.5 x 11 inch piece of paper) were used in the practice trials for each condition to visually display the value of the critical card. A Lenovo Thinkpad laptop was used to keep track of participants' responses.

Procedure

Children and adults completed a single, one-on-one session, lasting approximately 30-minutes for children and 20-minutes for adults. After obtaining written consent from adults and verbal consent from children, the experimenter sat opposite the participant at a small table and explained that they would be "playing a fun game" with decks of cards. The experimenter then led the participant through the practice trials and test trials for each of the three conditions. Stickers were used as incentives for children and candy was used for adults.

Practice trials. Participants completed three practice trials before the gain and loss condition, and two before the neutral condition. Participants first learned that a deck is made up of cards that are one of two colors ("look this deck has orange and blue cards"). The practice trials always used a deck with a ratio of 4:6 because this ratio represents a minimal disparity in frequency and ensured that participants understood the rules related to both desirability and accuracy. The experimenter asked participants to help count how many critical and noncritical cards there were before the first practice trial, and reminded them of these numbers of each for subsequent practice trials. Reference sheets used for practice trials contained two squares representing the two colors used in the deck.

In the first practice trial, the participant learned about the accuracy incentive. The experimenter explained that he/she was "going to shuffle the deck and pick a card without looking", and that the participant got to say if he or she thinks "it was a blue card or an orange card". The experimenter then pointed to the appropriate color on the reference sheet as he explained the rules. For this practice trial, a sticker was placed next to both colors on the reference sheet. If the prediction was correct, the participants learned that they earned a sticker (or candy, for adults). After the experimenter selected a card, reminded the participant of the ratio, and the participant made his or her guess, the identity of the card was revealed and the experimenter explained the result. If the prediction was correct, the experimenter explained that, "you get a sticker because you said it was going to be an orange one, and it was!" If the prediction was incorrect, the experimenter explained that, “you don’t get a sticker this time, because you said it was going to be a blue one, but it wasn’t.” In either case, the experimenter reminded the participant of the opposite outcome (e.g., “If you had said it was a green one, you would have gotten a sticker because it is a green one.”).

The next practice trial followed a similar procedure. In the second practice trial, participants learned that they get (for the gain condition) or lose (for the loss condition) a sticker if the critical card is drawn (“If it’s blue, you get/lose a sticker. If it’s green, you don’t get/lose one”). The reference was used as a visual reminder: The two colors of the deck being used appeared on the left, and the experimenter placed a sticker next to the color of the critical card in the gain and loss conditions. After explaining the value of the critical card, the experimenter reminded children how many of each color card there was, then shuffled the deck, placed the top card on the table, revealed the identity of the card and explained the outcome and its opposite (e.g., “Look, it’s a blue one! That means that you get/lose a sticker because it was a blue card and blue means you win/lose a sticker. If it was a green one, nothing would happen.”) Note that because the critical card was not associated with either a gain or a loss in the neutral condition, participants did not complete this practice trial in that condition.

The final practice trial provided an opportunity for participants to combine the rules for making accurate predictions and for drawing a critical card. The experimenter, pointing to the reference sheet, first reminded participants that they could earn a reward for an accurate prediction (e.g., “So if you say it’s going to be a green one, and it is, you get a sticker. But if you say it will be green and it isn’t, then you don’t get a sticker.”). Then the experimenter reminded them of the critical card value (e.g., in the gain condition, “Remember, you can also get a sticker if a blue card is drawn. So you can earn a sticker if you guess right, and you can earn another sticker if it’s a blue one.”). Note that in the neutral condition, participants learned that they would not get or lose anything if the critical card was drawn during this combination practice trial, and that the only way to earn a reward was with an accurate prediction.

After the experimenter reminded the participants about the rules and the card ratios, he shuffled the deck, placed the top card on the table, face-down, and asked, “Do you think it is going to be a [color] card or a [other color] card?” The card was then revealed, and the experimenter asked the participant to explain the accuracy and critical card rules. For example, in the gain condition: “Look, it’s blue. You said it would be a green one but it’s not, so do you get a sticker for being right? No, you don’t get a sticker this time, because you guessed wrong. But remember that you can get a sticker another way too. Do you remember what color card you can get an extra sticker for? That’s right, you get an extra sticker if it’s a blue one, and it is, so you get a sticker for that! If it had been green, you would have gotten a sticker for being right, but you wouldn’t have gotten another sticker since it’s not blue.”

If participants failed to repeat the rules back to the experimenter correctly, the experimenter corrected them and repeated the same practice trial. If the participant continued to misunderstand the rules, the experimenter thanked them for participating and offered them a thank-you gift.

Test trials. Next, participants completed a total of 10 test trials, blocked into two sets of five trials (one trial for each of the five ratios: 1:9, 4:6, 5:5, 6:4, 9:1). The first set of five trials used one color combination (e.g., blue and green cards) and the second set of trials used a second color combination (e.g., red and yellow cards). The color of the critical card was counterbalanced across participants, and the five trials within each of the two blocks were randomly ordered.

The test trials proceeded in much the same way as the third practice trial. To begin, the experimenter explained that this time, he was going to “change how many blue and green cards we have.” As in the practice trials, participants helped count the number of each type of card. Before each trial, the experimenter repeated the rules and the ratio of cards: “So there are nine blue ones and one green one this time. Remember, you can get a sticker if you are right about which card it is, and you can also get/lose a sticker if it’s a blue one (or – “and guessing correctly is the only way to get a sticker” – in the neutral condition).” Reference sheets were not provided for test trials because their use would have made critical card and accuracy information unequally salient.

Participants predicted which card would be drawn and the experimenter entered their predictions into the computer program. In the traditional marked card paradigm, adults were not given feedback because of its potential to influence participants’ subsequent predictions. In order to modify the paradigm to cater to children’s limited attention span, feedback on the total number of stickers participants either gained or lost was provided to all participants after every five trials.

After 10 test trials, children were given the chance to take a 5-minute break. This break was offered to ensure that they were not getting bored or fatigued. Only two children took a break. Both did so after the second set of 10 test trials and this time was spent doing a fun but cognitively undemanding task (e.g., jumping jacks, bouncing a ball). After the break, participants continued with the practice trials for the next condition.

Quantitative comparison. After participants completed all practice and test trials, the experimenter asked them to indicate which color card was most frequent in three different decks. This measure was used to ensure that participants (particularly children) were sensitive to information about card ratios. Recall that understanding and considering probability information is important because it allows for a clearer interpretation of participants' predictions.

The experimenter used three red/yellow card decks with 1:9, 4:6, and 5:5 respective ratios. These ratios were selected because they represent each ratio used in the paradigm without the reversal of the 1:9 and 4:6 ratios. Participants helped the experimenter count how many of each color cards there were in the deck and then determined if there were more of one type of card than the other type or if there were the same number of each card.

Results

To analyze the results, predicting a critical card was scored as a "1" and predicting a noncritical card predictions was scored as a "0". There was a prediction made in block one and in block two for each of the five deck frequencies. The block one and two responses were averaged to create a composite score. Thus, each participant had a total of fifteen composite scores, five for each frequency in each of the three conditions. For example, a participant who predicted a critical card for the 1:9 ratio deck in the first block of the gain condition and a noncritical card for 1:9 deck in the second block of the gain condition would have a composite score for .5 for the 1:9 ratio deck in the gain condition.

Preliminary analyses showed no effects involving gender, $F(2, 117) = 0.68$, $p = .51$, $\eta_p^2 = .01$, child age (4- or 5-years-old), $F(2, 57) = 0.12$, $p = .89$, $\eta_p^2 = .01$, or the order in which participants completed each condition (gain first, loss first, or neutral first), $F(2, 117) = 2.09$, $p = .13$, $\eta_p^2 = .03$. These factors were not considered in further analyses.

There was missing data from 11 (six 4-year-olds and five 5-year-olds) of the 60 children in the second block of test trials for one or more conditions. All stopped due to boredom, a lack of focus, or a combination of both. A 3(Desirability: gain, neutral, loss) x 5(Frequency: 1:9, 4:6, 5:5, 6:4, 9:1) x 2(Block: block 1, block 2) repeated measures ANOVA showed that, for children who completed both blocks, there was no main effect of block, $F(1, 48) = 1.52$, $p = .22$, $\eta_p^2 = .03$, or an interaction between block and desirability, $F(2, 47) = 0.61$, $p = .55$, $\eta_p^2 = .03$. Therefore, for the analyses reported below, block one predictions made by children who did not complete both blocks were used in place of the composite score. To be conservative, all analyses were run with and without the children who were missing block two data. Results did not differ between the two samples so only results from analyses on the the full dataset are reported.

Hypothesis 1: Children's Desirability Bias

Of primary interest was to determine if children exhibited a DB. Figure 1 shows the percentage of times that children predicted the critical card as a function of its frequency for each of the three conditions. Collapsing across frequencies, children predicted the critical card 67.5% (*SD* = 0.03) of the time when it was desirable (gain a sticker), 48.3% (*SD* = 0.03) of the time when it was neutral, and 26.3% (*SD* = 0.03) of the time when it was undesirable (lose a sticker).

A 3(Desirability: gain, loss, neutral) x 5(Frequency: 1:9, 4:6, 5:5, 6:4, 9:1) repeated measures ANOVA using children's composite scores confirmed hypothesis one: There was a significant main effect of desirability, $F(2, 58) = 41.09$, $p < .001$, $\eta_p^2 = .59$. Children were more likely to predict the critical card in the gain condition than the neutral ($p < .001$) and loss ($p < .001$) conditions. They were also more likely to predict the critical card in the neutral condition than in the loss condition ($p < .001$). In other words, children's predictions were inflated by the desirable outcomes and diminished by undesirable outcomes.

A paired-samples t-test was conducted to compare the difference between children's predictions made in the gain and neutral conditions and the difference between predictions made in the loss and neutral conditions. The difference between children's predictions made in the gain ($M = 0.19$, $SD = 0.30$) and loss ($M = 0.22$, $SD = 0.29$) condition compared to the neutral condition was not significant, $t(59) = -0.47$, $p = .64$, $d = 0.10$. In other words, children were just as biased to predict the critical card when it was associated with a gain as they were to not predict it when it was associated with a loss.

Importantly, Figure 2 shows that children were sensitive to the frequency of critical and noncritical cards across all five deck ratios, $F(4, 56) = 5.78$, $p = .001$, $\eta_p^2 = .29$. Collapsing across desirability condition, children were more likely to predict a critical card to be drawn as the number of critical cards in a deck increased. This pattern suggests that the children took the frequency of critical and noncritical cards into account before making their predictions.

The interaction between the desirability of the critical card and its frequency was not significant, $F(8, 52) = 0.70$, $p = .69$, $\eta_p^2 = .10$ for children. As demonstrated in Figure 1, the strength of the DB, indicated by the difference between the gain and loss conditions, did not differ by critical card frequency.

Individual patterns of responding were consistent with the reported results. Fifty of the 60 children predicted more critical cards when they stood to gain a sticker than when they stood to lose a sticker. In contrast, only 5 of the 60 children did the opposite and 5 showed no difference.

Hypothesis 2: Comparison of the Desirability Bias in Children and Adults

A second goal of this research was to determine if there was a difference in the strength of the DB between children and adults. Adults were first analyzed separately, and then compared to children. Figure 1 shows the percentage of time that adults predicted the critical card as a function of its frequency for each of the three conditions. Collapsing across critical card frequencies, adults predicted the critical card 59.8% (*SD* = .02) of the time when it was desirable (gain a point), 47.8% (*SD* = .01) of the time when it was neutral, and 47.8% (*SD* = .02) of the time when it was undesirable (lose a point).

A 3(Desirability: gain, loss, neutral) x 5(Frequency: 1:9, 4:6, 5:5, 6:4, 9:1) ANOVA using composite scores confirmed that adults showed a main effect of desirability, $F(2, 58) = 16.36$, $p < .001$, $\eta_p^2 = .55$ and frequency, $F(4, 56) = 1{,}052.97$, $p < .001$, $\eta_p^2 = .99$. As in previous research with adults using the traditional marked card paradigm, there was a significant interaction between desirability and frequency, $F(4, 56)$

= 4.58, p < .001, η_p^2 = .55, revealing that the DB was stronger for the 5:5 frequency than any of the other frequencies. These results show that the modified marked card paradigm used in the present study replicated the overall desirability effect found in prior research that used the traditional marked card paradigm. This step was crucial because it suggested that the version of the marked card paradigm used for this study could validly compare children's responses to those of adults.

A paired-samples t-test was conducted to compare the difference between adult's predictions made in the gain and neutral conditions and the difference between predictions made in the loss and neutral conditions. Overall, the difference between adults' predictions made in the gain (M = 0.12, SD = 0.16) and loss (M = 0.00, SD = 0.20) condition compared to the neutral condition was significant, $t(59) = 3.39$, p = .001, d = 0.66. In other words, the strength of adults' bias differed by condition: They were more swayed by their desired outcome in the gain condition than in the loss condition.

A closer look at the desirability main effect, using pairwise comparisons, showed that adults, like children, predicted the critical card in the gain condition more often than they did in the loss condition (p = .004) or neutral condition (p < .001). However, unlike children, adults predicted the critical card in the loss condition as often as they did in the neutral condition (p = 1.00). This result is interesting because in the traditional marked card paradigm, adults typically overpredict the critical card in the gain condition to the same extent that they underpredict the critical card in the loss condition (e.g., Irwin, 1953; Irwin & Metzger, 1966; Irwin & Snodgrass, 1966; Windschitl et al., 2010, study 1). Thus, although the modified marked card paradigm used in the present study replicated the overall DB, caution must be used in interpreting a comparison across ages, specifically for the loss condition.

In order to directly compare children's responses to those of adults, a 3(Desirability: gain, loss, neutral) x 5(Frequency: 1:9, 4:6, 5:5, 6:4, 9:1) x 2 (Age group: children, adults) ANOVA was conducted with age group as a between-subjects variable. There was a significant interaction between desirability and age group, $F(2, 117) = 15.75$, $p < .001$, $\eta_p^2 = .21$. Specifically, children were more influenced by the desirability of the critical cards than adults were. Pairwise comparisons revealed that children predicted the critical card more often than did adults in the gain condition, ($p = .03$). Similarly, children predicted the critical card less often than did adults in the loss condition, ($p < .001$). These results suggest that, as hypothesized, children have a stronger bias than adults.

Children's reliance on frequency information was different than adults'. As demonstrated in Figure 2, a significant interaction between frequency and age group, $F(4, 115) = 72.8$, $p < .001$, $\eta_p^2 = .72$, emerged. Pairwise comparisons showed that adults were more likely than children to predict a critical card in decks containing more critical than noncritical cards (6 critical: 4 noncritical, $p < .001$ and 9 critical: 1 noncritical, $p < .001$). Similarly, adults were less likely to predict a critical card in decks containing fewer critical than noncritical cards (4:6, $p < .001$ and 1:9, $p < .001$). In other words, adults made more use of the frequency differences than children when making predictions. Predictions did not differ between the two age groups for decks with 5:5 ($p = .50$) ratio, suggesting that both children and adults understood that there were no probability differences for these decks.

Additional Findings

Participants were asked to identify which color card was more frequent in each of three different decks at the end of the procedure in order to provide a second measure of (mostly children's) sensitivity to frequency. None of the children answered zero questions correctly, three answered one question correctly, 22 answered two questions correctly, and 35 answered all three questions correctly. This distribution is significantly greater than what would be expected by chance (15 children in each category), X^2 (3, N = 60) = 56.4, $p < .001$, Yate's correction for continuity applied. These results show that most children could determine which color was more frequent. All of the adults answered these questions correctly.

A 3(Desirability: gain, loss, neutral) x 5(Frequency: 1:9, 4:6, 5:5, 6:4, 9:1) x 2(Quantitative Comparison: three correct, two or fewer correct) ANOVA was conducted determine if there was a relationship between the number of frequency questions answered correctly and the strength of the DB in children. For this analysis, children who correctly answered all three questions were included in one group and children who missed one or more question were included in the other. The analysis revealed a significant interaction between numeracy and desirability condition, $F(2, 57) = 5.06$, $p = .009$, $\eta_p^2 = .15$. Specifically, children who correctly answered all three frequency questions exhibited a weaker DB than children who missed at least one question. However, when only the 35 children who answered all three questions correctly were compared to adults, these children still showed a stronger DB, $F(2, 93) = 5.46$, $p = .001$, $\eta_p^2 = .11$. In other words, even the children who were the most number conscious were more biased than adults.

Finally, a 3(Desirability: gain, neutral, loss) x 5(Frequency: 1:9, 4:6, 5:5, 6:4, 9:1) x 2(Position: left first, right first) repeated measures ANOVA was conducted separately for children and adults to ensure that the critical card position (and color) did not impact participants' predictions. Results showed no significant interaction between critical card position and desirability for adults, $F(2, 57) = 0.20$, $p = .82$, $\eta_p^2 = .01$, indicating that the position and color of the critical card did not influence their predictions. Children's predictions were significantly different depending on the critical card position, $F(2, 57) = 3.70$, $p = .03$, $\eta_p^2 = .12$. Children were more likely to predict a critical card when it was located on the left for the first block of test trials (orange cards) and when it was located on the right for the second block (polka-dotted cards), however this was only the case in the neutral condition ($p < .001$). For unknown reasons, children were more likely to predict orange cards to be drawn over brown ones and polka-dotted cards to be drawn over striped ones. Nevertheless, children's predictions in the neutral condition reflected a sensitivity to critical card frequency.

Discussion

The DB has been found repeatedly with adults using the classic marked card paradigm (Budescu & Bruderman, 1995; Windschitl et al., 2010). The present study is the first to show the same bias in children as young as four years of age using a modified, child-friendly version of the paradigm. As hypothesized, 4- and 5-year-old children exhibited a stronger desirability bias than adults, especially when faced with a loss. Interestingly, adults showed a bias only when they stood to gain a reward but not when they stood to lose one. Several artifactual and theoretical explanations for why children might show a bias in the first place, and why their bias is stronger than adults' are considered.

Recall that, in order to make optimal predictions in the marked card paradigm, children should have predicted whichever type of card was most frequent and ignored the desirability associated with drawing either type of card. But the 4- and 5-year olds in this study had difficulty making optimal predictions; instead, the objective probability information was overshadowed by the (un)desirability of the critical card. Children tended to overpredict the critical card when it was associated with the gain of a sticker and underpredict it when it was associated with a loss, and they did so to the same extent when compared to a neutral condition.

Although a direct comparison between the children in the present study and the 9- and 11-year-olds used in Marks' (1951) original study is not possible due to important procedural differences, it still informative to visually compare the data from the critical card frequencies shared between the two studies (1:9, 5:5, & 9:1). For the 1:9 critical to noncritical card ratio, the older children from Marks' study predicted a critical card, on average, 47% of the time when it resulted in a gain and 0% of the time when it resulted in a loss, whereas the children in the present study predicted a critical card 53% of the time when it resulted in a gain and 16.7% of the time when it resulted in a loss. Similarly, for the 5:5 ratio, older children predicted a critical card 90% of the time when it resulted in a gain and 14% of the time when it resulted in a loss, where younger children predicted one 73% of the time when it resulted in a gain and 28% of the time when it resulted in a loss. Finally, for the 9:1 ratio, older children predicted a critical card to be drawn 100% of the time when it resulted in a gain and 56% of the time when it resulted in a loss, whereas younger children predicted one 78% of the time when it resulted in a gain and 29% of the time when it resulted in a loss. These numbers suggest that the magnitude of the desirability bias, quantified by the difference between critical card predictions in the gain and loss conditions, was fairly similar for younger and older children. Interestingly, the older children in Marks' original study were more sensitive to the frequency of critical cards in a deck than children in the present study. Compared to the adults in the present study, however, children in both ages were less sensitive to frequency. This difference in the extent to which children consider frequency information may help to explain the difference in the strength of the DB between children and adults.

One artifactual explanation for children's bias is that they did not understand the frequency information presented to them. There are two reasons to discount this explanation. First, results showed that children were increasingly likely to predict a critical card as the number of critical cards in a deck increased. Second, at the end of the study the majority of children correctly indicated which of two colors was more frequent in three different decks and only 5 children got two or more incorrect. Thus, there is clear evidence that children paid attention to the probability information during the study and used it when considering how to respond.

Given past research on young children's numeracy, the fact that children were capable of using probability information was not particularly surprising (e.g., Kushnir & Gopnik, 2005; Ma & Xu, 2011; Xu & Garcia, 2008;). Even infants as young as 6-months old are capable of detecting when there is more or less of something and of using that information to guide their predictions. In one study, for example, infants saw a box of ping-pong balls, 80% of which were red and 20% of which were white (Xu & Garcia, 2008). When an experimenter closed her eyes and randomly drew five balls out of the box, infants were surprised if she drew an unlikely sample, such as four white balls and one red ball, as indicated by the length of their gaze. They were not surprised when she drew a more likely one, such as four red balls and one white ball.

In addition, two modifications were purposefully made to the modified marked card paradigm used in the present study in order to make the number of critical and noncritical cards particularly salient for children. First, children counted the number of both types of cards in each deck with the experimenter and second, the experimenter reminded the participant of these frequencies right before they made each prediction. These same modifications may have influenced adults' responses in the loss condition, which is discussed below. What was impressive in the present study was that, despite the intentional emphasis on probability and children's apparent sensitivity to this information, children were still strongly influenced by the (un)desirability of the critical card.

Adults also showed a desirability bias, but, unlike the children in this study and adults from previous studies, the bias was only present in the gain condition. Adults in previous research using the traditional marked card paradigm appear to be influenced by desirability to the same extent in both the loss and gain conditions (Irwin, 1953; Irwin & Metzger, 1966; Irwin & Snodgrass, 1966; Windschitl et al., 2010, study 1). One reason for the unexpected finding in the loss condition is that adults developed a new hedging-their-bets strategy for minimizing loss. The reason they might have been able to do so in the present study but not in previous studies is twofold. First, the cognitive load was reduced which could have enabled them to formulate and use such a strategy, and second, the point they could earn for being accurate could offset the loss of a point from drawing a critical card.

In the loss condition, participants could earn one point for making an accurate prediction but could also lose one point if the critical card were drawn. For example, participants who correctly predicted a critical card in a deck of four critical and six noncritical cards would gain a point for accuracy and lose a point for drawing a critical card. In other words, if they made an accurate prediction, the point they could earn for being right would offset the point they would lose for drawing a critical card. Hedging their bets ensured that they would not lose anything, while also making it impossible to gain anything. The same strategy is not relevant in the gain condition because participants could only earn points, they could not lose anything. As the data show, adults were still biased for predictions involving potential gains. In past studies, the accuracy incentive was worth only one-quarter of what participants stood to lose, so a hedging-your-bet strategy would not have been relevant because the accuracy incentive could not fully offset the value of the critical card.

Although inadvertent, this finding is interesting because previous research has shown that stressing accuracy to the participants or providing them with physical accuracy incentives did not seem to alter the desirability bias (e.g., Budescu & Bruderman, 1995; Crandall et al., 1955). Indeed, a meta-analysis has shown that effect sizes did not differ between studies that offered incentives for or stressed the importance of being accurate and studies that did neither (Krizan & Windschitl, 2007b). Because it was believed that incentives mattered little in terms of eliciting a bias, the accuracy incentive in the present study was worth just as much as the prediction incentive. This helped increase its salience for young children and was simply more convenient (stickers and candy are not as easy to fractionalize as monetary incentives).

However, the present study suggests that the strength of the desirability bias, at least in the loss condition, can be shifted by how valuable the accuracy incentive is relative to the desirability of the critical card (or perhaps how salient the accuracy incentive is; see Vosgerau, 2010 for a discussion). A closer look at research on the effect of accuracy incentives shows that it has not been widely studied – only 27 individual studies were included in the meta-analysis that examined the effects of accuracy incentives (Krizan & Windschitl, 2007b), and only 19 used a version of the marked card paradigm. It is possible that these incentives matter more than previously thought. One study suggests that small incentives used with the marked card paradigm do not influence subjective probabilities, or people's estimates of the general likelihood of an outcome occurring (subjective probabilities are typically reported as a percentage chance). Instead, the researchers argued that incentives might influence prediction thresholds – the evidence required to predict one dichotomous outcome over another (Price & Marquez, 2005, as cited in Krizan & Windschitl, 2007b). Thus, accuracy incentives may have more of an impact on dichotomous predictions, like those made in this study, than on subjective probability estimates.

In addition to increasing the relative value of accuracy, the cognitive load was dramatically reduced in order to make the paradigm accessible to children. Unlike in the traditional marked card task, adults in this study had the rules broken down into parts accompanied by several practice trials, had all the trials for a given condition presented in succession rather than randomly, and were reminded of the rules before each prediction.

Reducing the cognitive load was necessary to create a task that would allow for a comparison by age, but it may have provided an opportunity for adults to reason differently about their predictions. Given the change in accuracy incentive, many adults likely ended up developing a new strategy for minimizing losses.

In order to explore whether the hedging hypothesis had legs, a subset of 19 of the adult participants were asked one to two follow up questions (unsystematically) at the end of the study. They were first asked if that they had used a different strategy when the critical card resulted in a loss than when it resulted in a gain. If they answered yes, they were then asked to explain the difference in strategy in an open-ended response. Twelve of the 19 adults reported using the hedging strategy.

These exploratory data are the basis for a new study that is currently investigating whether the increased accuracy incentives were the reason behind adults' use of a loss-aversive strategy for potential losses. The procedure is identical to the present study except that the value of the critical card is increased to four times the accuracy incentive (four points are lost for drawing the critical card, and one point is won for accuracy), as it has been in previous studies with adults. In this case, a hedging strategy would not work because being accurate would not offset the potential loss of drawing a critical card. Response patterns in the loss condition would likely be more similar to previous studies in this case, but children would still be expected show a stronger bias in the loss condition (as they do in the present study).

It is worth speculating about theoretical reasons for why adults did not show a desirability bias in the loss condition. One speculation is that adults are more loss-averse than are children. In the marked card paradigm, adults were willing to forsake the optimal amount of points they could have earned in order to ensure that they did not lose anything. In contrast, children were more willing to risk losing something because they were more optimistic that the undesirable outcome would not occur.

There is abundant literature that supports loss aversion in adults. In a variety of circumstances, adults show more sensitivity to losses than they do to gains (Harbaugh et al., 2002; Johnson et al., 2007; Kahneman & Tversky, 1984; Kahneman et al., 1991; Steelandt et al., 2013; Tom et al., 2007). People who lose money if they failed to reach a weight loss goal, for example, have been found to be more likely to lose weight than are those who earn money for achieving the same goal (Driver & Hensrud, 2013). Similarly, adults who possess an item, and therefore have the possibility to lose it, place more value on the item than they do if they don't own it, a phenomenon known as the endowment effect (Thaler, 1980). Given their sensitivity to loss, it is actually a bit surprising that adults who have completed the traditional, computerized version of the marked card paradigm do not show a reduced desirability bias in the loss condition.

The fact that children exhibited a desirability bias in the loss condition coincides with previous work indicating that children are less sensitive to negative information than positive information when making judgments about others (Boseovski & Lee, 2006, 2008, Rholes & Ruble, 1986). In one study for instance, preschoolers required five times more evidence of another child being mean to consider them a bad person than evidence of another child being nice to consider them a good person (Boseovski & Lee, 2006). In the marked card paradigm, children's predictions in the loss condition reflected this decreased sensitivity to negative information. Children, unlike adults, were willing to take risks that opened them up to a potential loss in order to gain a reward. Future research should explore how the value of the accuracy incentive relative to the value of the critical card influences the desirability bias.

Another question for speculation is whether the processes that have been hypothesized to underlie adults' desirability bias are the same processes that underlie the bias in children. Signal detection theory and the biased-guessing account are two explanations for why adults,' and potentially children's, desires influence their judgments, even when they are incentivized to be accurate. Windschitl et al., (2010) provide a review of several other accounts that seem less likely mechanisms for the bias for adults and are not discussed as possibilities to explain children's bias here.

Signal detection is the ability to correctly indicate when a signal is present (a "hit") or not (a "correct rejection"). Saying a signal is present when it is not is considered a "false alarm" while saying it is absent when it is present is a "miss." In the marked card paradigm, predicting a critical card when one ends up being drawn would be considered a "hit." Participants might place more importance on getting "hits" when it is associated with a gain, and therefore de-emphasize probability information. In other words, the threshold for predicting a desirable outcome based on probability is lowered from baseline. Predicting a critical card when it is not actually drawn would be considered a "false alarm." According to this account, in the marked card paradigm participants might place more emphasis on avoiding "false alarms" in the loss condition (Tanner & Swets, 1954). That is, the threshold for predicting when an undesired outcome based on probability is increased from baseline. Adjusting the threshold up (for gains) or down (for losses) would lead to a greater number of predictions that go against the odds.

A related explanation, called the biased-guessing account, suggests that the desirability bias in adults results from a weighted guessing process. This account posits that individuals assess probability information but make a prediction that goes against this information because of the possibility of a less probable but more desirable outcome occurring (Windschitl et al., 2010). Furthermore, it specifies that adults exhibit a desirability bias due to their belief that part of their prediction is arbitrary because it takes the form of a guess. When this arbitrary guessing element is removed by making likelihood estimates (i.e., what is the likelihood of an outcome occurring) instead of dichotomous predictions, the desirability bias vanishes (Windschitl et al., 2010). This account explains why adults tend to show a desirability bias when the number of critical and noncritical cards are similar (e.g., 4:6) and do not show a bias when they are very different (e.g., 1:9). For decks with similar frequencies of cards, a desirable but unlikely outcome is only slightly less likely to occur than an undesirable outcome. This leads adults to be more likely to predict a desirable outcome for this type of deck than for a deck in which a desirable outcome is much less likely than an undesirable one.

Like signal detection theory, the biased-guessing account can potentially explain why adults, as well as children, predict desirable outcomes despite understanding that they are unlikely. The primary difference between the two can be seen in a hypothetical example involving a deck with four red cards and six yellow cards where the red (critical) card automatically results in a gain. For this deck, signal detection theory would assert that an individual would use a criterion for whether or not they would be willing to predict the desirable but unlikely outcome (a red card), and would thus make this prediction every time when facing this specific deck. The biased-guessing account, however, would involve the individual realizing that a red card would be drawn 40% of the time, and therefore result in him or her predicting the desirable but unlikely outcome some of the time, but not always.

In addition to the two mechanisms discussed, there are a number of other theories, such as the biased-threshold and biased-evaluation accounts (see Windschitl et al., 2010), about why adults make biased predictions. The present study cannot draw strong conclusions about which process is leading children to make biased predictions, but the biased-guessing approach appears more likely because, while children frequently predicted desirable but unlikely outcomes, they did not appear to use a consistent threshold when making these predictions.

A deeper question than what mechanism causes biased predictions would be why children or adults would want to make predictions that go against probability in the first place, when they stand to gain the most and lose the least when they base predictions on the odds alone? Whether the bias has roots in biased-guessing, signal detection theory, or something else, why might the desirability bias or other forms of optimistic reasoning exist, and why do these biases appear to be enhanced in children?

Given that children seem to have a stronger bias than adults, one suggestion is that being overly optimistic is an adaptive and perhaps even an inborn tendency. It might be adaptive to focus on what you want rather than what is likely to occur at first. The bias is then perhaps reduced as a result of socialization processes (e.g., being overconfident about your own performance may be seen as immodest) and experience (e.g., not getting what you want) or interfered with by other competing information (e.g. moral judgments about “losing”).

Imagine an infant first learning to sit up. The infant tries, and fails repeatedly, perhaps hundreds of times, before having any success. If the infant were to predict whether or not, on the next try, they would succeed or fail, the odds would point to failure. But if infants believed they would fail, the motivation to continue trying might decline and they would perhaps never achieve this important milestone. Many kinds of developmental achievement, indeed perhaps even survival in the beginning, may depend in part on an infant's willingness to ignore the odds in favor of optimism. For adults, however, it might pay to be somewhat pessimistic, or at least realistic, at times.

If biased reasoning, specifically the desirability bias, is an innate or adaptive tendency, then it should be present very early in life, at least to the extent that perspective taking or mentalizing (understanding others' mental lives) is not involved. One way to test for this bias in infants would be to borrow a paradigm from the infant numeracy literature. In such a procedure, infants would first be shown a sample of blue and red marbles. In a habituation period, drawing a blue marble could be made desirable for infants by pairing the display of a blue marble with the appearance of a cheerful, positive stimulus. In contrast, drawing a blue marble could be made undesirable by pairing it with the removal of the positive stimulus. The ratio of blue to red marbles could be manipulated and time spent looking at a randomly drawn marble would determine the extent to which infant's desire (or lack thereof) for a blue marble influenced their expectations. This procedure would necessarily involve giving feedback, something that was not done in the marked card paradigm. This could be accounted for, however, by using a single-trial design and increasing sample size. Like the marked card paradigm, this procedure mimics the type of games in which the desirability bias seems to be most robust.

Optimistic thinking persists under a variety of circumstances, at least into middle childhood. Preschoolers believe that they will jump substantially further than they ever have in any previous practice trials, even though they understand that their previous jumping distance is a good predictor of their future jumping distance (Schneider, 1998). 4-year-olds optimistically predict that they will remember picture pairs better than they actually do (Lipko et al., 2009, study 1) and 4-year-olds predict that they will be able to improve their performance for a simple skill-based game even when the task had previously been designed to ensure that participants would fail (Stipek et al., 1984). In all of these circumstances one could argue that children's optimism is adaptive: Believing they can perform a task better than they ever have before or that what they desire is what they can expect may help children persevere even after experiencing failure or sub-optimal experiences time and time again.

Strategic optimism is important for adults as well. Being optimistic about landing that desirable graduate school position or dream job after failing to do so the last several times, and despite knowing how many applicants typically apply for the limited number of opportunities, may help keep the applicant from landing instead in a pit of depression. Abundant research has shown that optimistic thinking confers a wide range of benefits including better physical and mental health, more successful relationships, higher self-esteem, better coping and stress management skills, and more energy compared to people who tend to have pessimistic styles of thinking (see Carver, Scheier & Segerstrom, 2010, for a review).

But with experience, adults learn that in some circumstances, like when gambling in a casino (e.g., Gibson & Sanbonmatsu, 2004), being pessimistic might be more ideal than being optimistic. Being strategically pessimistic can protect against intense disappointment by diminishing the value of a particular outcome. Krizan and Windschitl (2007a; 2007b) suggest that strategic pessimism in adults seems to occur most frequently when outcomes are not within one's own control (e.g., Shepperd, Findley-Klein, Kwavnick, Walker, & Perez, 2000; Taylor & Shepperd, 1998). Since adults did not have any control of the random card drawn in the marked card paradigm, strategic pessimism could help to explain why adults developed a loss-prevention strategy for facing potentially undesirable outcomes. According to one study, children's degree of optimism tends to look more like adults' (who are more realistic) during middle childhood (Lockhart et al., 2002). It would be interesting to test children ages 10- to 13-years in the marked card paradigm to explore whether this shift towards realism occurs at about the same time as when children's desirability bias looks more like adults'.

Another intriguing possibility for why children show a more pronounced bias than adults in the present study has to do with children's tendency to engage in magical thinking. Magical thinking refers to attributing a causal relationship between two events that could not reasonably be causally related. For example, one study showed preschool-aged children a box with crumpled postage stamp inside (Subbotsky, 2004). An experimenter closed the box, said some magic words, and then opened the box again to reveal a new, non-crumpled stamp (a hidden trap door was used to switch the stamps). Out of 4-, 6-, and 9-year-olds, only 4- and 6-year-olds accepted the experimenter's explanation that this was the same stamp they had earlier seen crumpled up, but had transformed due to magic. 4- and 6-year-olds were also more likely than 9-year-olds to offer a magical explanation when given a chance to explain the phenomenon on their own.

In another study, 6-year-olds indicated that if someone had a missing finger, they could grow it back if they wanted it to badly enough (Lockhart et al., 2002). Despite knowing that fingers do not normally grow back, children's belief that magical powers can make wishes come true led them to make optimistic predictions about a biologically impossible event. In other words, young children are willing offer impossible explanations for seemingly impossible events, especially when the outcome is something desirable. For children completing the marked card paradigm, predicting an unlikely outcome might make perfect sense simply because, to them, the outcome of the random card draw is susceptible to what can be thought of as "magical intervention." Adults, who have lost the "magic" of childhood, would not likely consider magic as having anything to do with future outcomes.

One final explanation for children's more pronounced bias is due to a cognitive issue called feature binding. Feature binding is when individual pieces of information are integrated and remembered together. This process requires attention and working memory. Because of the cognitive requirement, young adults are better at it than both children and older adults (e.g., Cowen, Naveh-Benjamin, Kilb & Saults, 2006). Remembering a blue circle from independently coded features (blue object, circle) is one example of feature binding, but the concept applies to integrating the separate features of the marked card paradigm as well (see Hommell, 2004, for a review).

In the marked card paradigm, the independent features are the probability information and the value of the critical card. In the present study, children were more sensitive to probability in the neural, compared to the gain or loss condition. Children also based their predictions the value of the critical card for the 5:5 critical card ratios when probability information is ambiguous (they show a desirability bias even though there is an equal chance of drawing either type of card). Although children can remember both pieces of information individually, they may have more difficulty compared to adults at combining the two features because of the cognitive requirements. Similarly, one reason why adults may not have shown a bias in the loss condition is because they were better able to integrate both pieces of information due to the reduced cognitive demands of the modified marked card paradigm. One study provides some support for this hypothesis. Five-and 6-year-olds and adults had to predict which of two kids broke a window based on the testimony of two witnesses. Both children and adults had no difficulty basing their predictions on how accurate the witnesses were in one study or on how confident they were in a second study. But when required to use both the accuracy and confidence of the witness, children ignored accuracy whereas adults considered both features (Tenney, Small, Kondrad, Jaswal, & Spellman, 2011). When the adults were placed under additional cognitive demands, they, like children, downplayed one feature in favor of another.

It would be interesting to explore further whether cognitive demands can push the strength of the desirability bias up or down. One way to do so while maintaining the ability to compare across age groups is to use the modified marked card paradigm but ask adults to complete a secondary task simultaneously. Another interesting possibility is to test older adults. Older adults' desirability bias should look more like children's if this feature binding account is valid. Feature binding abilities take a U-shaped function by age, with older adults and children performing worse than younger adults (e.g., Chalfonte & Johnson, 1996). In addition, older adults, like children, tend to have a positivity bias (Charles, Mather, & Carstensen, 2003; Mather & Carstensen, 2005).

The present study was the first to look for and successfully find a DB in preschool-age children. In order to do so, a child-friendly version of the marked card paradigm was created to test both children and adults. Results suggest that children have a stronger DB than adults in both the loss and gain conditions. Adults did not show a bias in the loss condition, possibly because of the reduced cognitive demands of the task and the relative value of the accuracy incentive. These results can provide a framework for understanding the development of other biases such as overoptimism and overconfidence. This research is the first to broaden this specific area of developmental literature by demonstrating what many parents would claim to have known from experience: Children expect to get what they want even when they understand that it is clearly unlikely.

References

Anderson, D. R., & Pempek, T. A. (2005). Television and very young children. *American Behavioral Scientist*, *48*, 505-522.

Babad, E. (1987). Wishful thinking and objectivity among sports fans. *Social Behaviour, 3,* 231–240.

Babad, E., & Katz, Y. (1991). Wishful thinking—against all odds. *Journal of Applied Social Psychology, 21*, 1921-1938.

Bar-Hillel, M., & Budescu, D. (1995). The elusive wishful thinking effect. *Thinking & Reasoning*, *1*, 71-103.

Bar-Hillel, M., Budescu, D. V., & Amar, M. (2008). Predicting world cup results: Do goals seem more likely when they pay off? *Psychonomic Bulletin & Review, 15,* 278-283.

Bastardi, A., Uhlmann, E. L., & Ross, L. (2011). Wishful thinking belief, desire, and the motivated evaluation of scientific evidence. *Psychological Science*, *22*, 731-732.

Berti, A. E., & Bombi, A. S. (1981). The development of the concept of money and its value: A longitudinal study. *Child Development*, *52*, 1179-1182.

Boseovski, J. J. (2010). Evidence for “rose-colored glasses”: An examination of the positivity bias in young children’s personality judgments. *Child Development Perspectives*, *4*, 212-218.

Boseovski, J. J. (2012). Trust in testimony about strangers: Young children prefer reliable informants who make positive attributions. *Journal of Experimental Child Psychology*, *111*, 543-551.

Boseovski, J. J., & Lee, K. (2006). Children's use of frequency information for trait categorization and behavioral prediction. *Developmental Psychology*, *42*, 500-513.

Boseovski, J. J., & Lee, K. (2008). Seeing the world through rose-colored glasses? Neglect of consensus information in young children's personality judgments. *Social Development*, *17*, 399-416.

Budescu, D. V., & Bruderman, M. (1995). The relationship between the illusion of control and the desirability bias. *Journal of Behavioral Decision Making*, *8*, 109–125.

Buehler, R., Griffin, D., & Ross, M. (1995). It's about time: Optimistic predictions in work and love. *European Review of Social Psychology*, *6*, 1–32.

Buehler, R., & McFarland, C. (2001). Intensity bias in affective forecasting: The role of temporal focus. *Personality and Social Psychology Bulletin*, *27*, 1480-1493.

Carlson, S. M., Moses, L. J., & Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory *Infant and Child Development*, *11*, 73– 92.

Carver, C. S., Scheier, M. F., & Segerstrom, S. C. (2010). Optimism. *Clinical Psychology Review*, *30*, 879-889.

Case, R., Kurland, M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology*, *33*, 386–404.

Chalfonte, B. L., & Johnson, M. K. (1996). Feature memory and binding in young and older adults. *Memory & Cognition*, *24*, 403– 416.

Charles, S. T., Mather, M., & Carstensen, L. L. (2003). Aging and emotional memory: the forgettable nature of negative images for older adults. *Journal of Experimental Psychology: General, 132,* 310-324.

Cowan, N., Naveh-Benjamin, M., Kilb, A., & Saults, J. S. (2006). Life-span development of visual working memory: When is feature binding difficult? *Developmental psychology*, *42*, 1089-1102.

Crandall, V. J., Solomon, D., & Kellaway, R. (1955). Expectancy statements and decision times as functions of objective probabilities and reinforcement values. *Journal of Personality*, *24*, 192–203.

Denison, S., & Xu, F. (2010a). Integrating physical constraints in statistical inference by 11-month-old infants. *Cognitive Science*, *34*, 885-908.

Denison, S., & Xu, F. (2010b). Twelve- to 14-month-old infants can predict single-event probability with large set sizes. *Developmental Science*, *13,* 798–803.

Driver, S. L., & Hensrud, D. (2013). Financial incentives for weight loss: A one-year randomized controlled clinical trial. *Journal of the American College of Cardiology*, *61*(10_S).

Droege, K. L., & Stipek, D. J. (1993). Children's use of dispositions to predict classmates' behavior. *Developmental Psychology*, *29*, 646-654.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149-1160.

Gibson, B., & Sanbonmatsu, D. M. (2004). Optimism, pessimism, and gambling: The downside of optimism. *Personality and Social Psychology Bulletin, 30*, 149-160.

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

Granberg, D., & Brent, E. (1983). When prophecy bends: The preference–expectation link in US presidential elections, 1952–1980. *Journal of Personality and Social Psychology*, *45*, 477-491.

Harbaugh, W. T., Krause, K., & Vesterlund, L. (2002). Risk attitudes of children and adults: Choices over small and large probability gains and losses. *Experimental Economics*, *5*, 53-84.

Hayes, S. P., Jr. (1936). The predictive ability of voters. *Journal of Social Psychology*, *7,* 183–191.

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

Heyman, G. D., & Giles, J. W. (2004). Valence effects in reasoning about evaluative traits. *Merrill-Palmer Quarterly (Wayne State University. Press)*, *50*, 86-109.

Hommel, B. (2004). Event files: Feature binding in and across perception and action. *Trends in Cognitive Sciences, 8*, 494-500.

Irwin, F. W. (1953). Stated expectations as functions of probability and desirability of outcomes. *Journal of Personality, 21*, 329-335.

Irwin, F. W., & Graae, C. N. (1968). Tests of the discontinuity hypothesis of the effects of independent outcome values upon bets. *Journal of Experimental Psychology*, *76*, 444-449.

Irwin, F. W., & Metzger, M. J. (1966). Effects of probabilistic independent outcomes upon predictions. *Psychonomic Science, 5,* 79–80.

Irwin, F. W., & Snodgrass, J. G. (1966). Effects of independent and dependent outcome values upon bets. *Journal of Experimental Psychology*, *71*, 282-285.

Jaswal, V. K., Croft, A. C., Setia, A. R., & Cole, C. A. (2010). Young children have a specific, highly robust bias to trust testimony. *Psychological Science*, *21*, 1541-1547.

Johnson, E., Herrmann, A., & Terry, K. (2007). *Individual-Level Loss Aversion in Riskless and Risky Choices* (No. 02). CeDEx Discussion Paper.

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1991). Anomalies: The endowment effect, loss aversion, and status quo bias. *The Journal of Economic Perspectives*, *5*, 193-206.

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist*, *39*, 341-350.

Kalish, C. W. (2002). Children's predictions of consistency in people's actions. *Cognition*, *84*, 237-265.

Klenberg, L., Korkman, M., & Lahti-Nuuttila, P. (2001). Differential development of attention and executive functions in 3-to 12-year-old Finnish children. *Developmental Neuropsychology*, *20*, 407-428.

Krizan, Z., Miller, J. C., & Johar, O. (2010). Wishful thinking in the 2008 US presidential election. *Psychological Science, 21*, 140-146.

Krizan, Z., & Windschitl, P. D. (2007a). Team allegiance can lead to both optimistic and pessimistic predictions. *Journal of Experimental Social Psychology*, *43,* 327-333.

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

Krizan, Z., & Windschitl, P. D. (2009). Wishful thinking about the future: Does desire impact optimism? *Social and Personality Psychology Compass, 3*, 227-243.

Kushnir, T., & Gopnik, A. (2005). Young children infer causal strength from probabilities and interventions. *Psychological Science*, *16*, 678-683.

Kushnir, T., Xu, F., & Wellman, H. M. (2010). Young children use statistical sampling to infer the preferences of other people. *Psychological Science*, *21*, 1134-1140.

Langer, E. J. (1975). The illusion of control. *Journal of Personality and Social Psychology*, *32,* 311-328.

Langer, E. J., & Roth, J. (1975). Heads I win, tails it's chance: The illusion of control as a function of the sequence of outcomes in a purely chance task. *Journal of Personality & Social Psychology*, *32,* 951-955.

Lee, L., Frederick, S., & Ariely, D. (2006). Try it, you'll like it: The influence of expectation, consumption, and revelation on preferences for beer. *Psychological Science*, *17*, 1054-1058.

Lipko, A. R., Dunlosky, J., & Merriman, W. E. (2009). Persistent overconfidence despite practice: The role of task experience in preschoolers' recall predictions. *Journal of Experimental Child Psychology*, *103*, 152-166.

Liu, D., Vanderbilt, K. E., & Heyman, G. D. (2013). Selective trust: Children's use of intention and outcome of past testimony. *Developmental Psychology*, *49,* 439-445.

Lockhart, K. L., Chang, B., & Story, T. (2002). Young children's beliefs about the stability of traits: Protective optimism? *Child Development*, *73*, 1408-1430.

Mather, M., & Carstensen, L. L. (2005). Aging and motivated cognition: The positivity effect in attention and memory. *Trends in Cognitive Sciences, 9,* 496-502.

Marcovitch, S., Boseovski, J. J., & Knapp, R. J. (2007). Use it or lose it: Examining preschoolers' difficulty in maintaining and executing a goal. *Developmental Science, 10,* 559–564.

Marks, R. W. (1951). The effect of probability, desirability, and "privilege" on the stated expectations of children. *Journal of Personality, 19,* 332-351.

Morlock, H. C., & Hertz, K. J. (1964). Effect of the desirability of outcomes on decision making. *Psychological Reports, 14,* 11–17.

Olsen, R. A. (1997). Desirability bias among professional investment managers: Some evidence from experts. *Journal of Behavioral Decision Making, 10,* 65–72.

Palmquist, C. M., Burns, H. E., & Jaswal, V. K. (2012). Pointing disrupts preschoolers' ability to discriminate between knowledgeable and ignorant informants. *Cognitive Development*, *27*, 54-63.

Price, P. C. (2000). Wishful thinking in the prediction of competitive outcomes. *Thinking & Reasoning*, *6*, 161-172.

Price, P. C., & Marquez, C. A. (2005). Wishful thinking in the predictions of a simple repeatable event: Effects on deterministic versus probabilistic predictions. Unpublished manuscript.

Pruitt, D. G., & Hoge, R. D. (1965). Strength of the relationship between the value of an event and its subjective probability as a function of method of measurement. *Journal of Experimental Psychology*, *69*, 483-489.

Rholes, W. S., & Ruble, D. N. (1986). Children's impressions of other persons: The effects of temporal separation of behavioral information. *Child Development*, *57*, 872-878.

Ross, B. M., & Levy, N. (1958). Patterned predictions of chance events by children and adults. *Psychological Reports*, *4*, 87-124.

Rothman, A. J., Klein, W. M., & Weinstein, N. D. (1996). Absolute and relative biases in estimations of personal risk. *Journal of Applied Social Psychology*, *26*, 1213-1236.

Schneider, W. (1998). Performance prediction in young children: Effects of skill, metacognition and wishful thinking. *Developmental Science*, *1*, 291-297.

Schuster, B., Ruble, D. N., & Weinert, F. E. (1998). Causal inferences and the positivity bias in children: The role of the covariation principle. *Child Development*, *69*, 1577-1596.

Shepperd, J. A., Findley-Klein, C., Kwavnick, K. D., Walker, D., & Perez, S. (2000). Bracing for loss. *Journal of Personality and Social Psychology*, *78*, 620–634.

Steelandt, S., Broihanne, M. H., Romain, A., Thierry, B., & Dufour, V. (2013). Decision-making under risk of loss in children. *PloS One*, *8*, e52316.

Stipek, D. J., Roberts, T. A., & Sanborn, M. E. (1984). Preschool-age children's performance expectations for themselves and another child as a function of the incentive value of success and the salience of past performance. *Child Development*, *55*, 1983-1989.

Tanner Jr, W. P., & Swets, J. A. (1954). A decision-making theory of visual detection. *Psychological Review, 61*, 401-409.

Taylor, K. M., & Shepperd, J. A. (1998). Bracing for the worst: Severity, testing and feedback timing as moderators of the optimistic bias. *Personality and Social Psychology Bulletin*, *24,* 915–926.

Tenney, E. R., Small, J. E., Kondrad, R. L., Jaswal, V. K., & Spellman, B. A. (2011). Accuracy, confidence, and calibration: How young children and adults assess credibility. *Developmental Psychology, 47,* 1065-1077.

Thaler, R. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization*, *1,* 39-60.

Tom, S. M., Fox, C. R., Trepel, C., & Poldrack, R. A. (2007). The neural basis of loss aversion in decision-making under risk. *Science*, *315*, 515-518.

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

Wilson, T. D., & Gilbert, D. T. (2005). Affective forecasting knowing what to want. *Current Directions in Psychological Science*, *14*, 131-134.

Windschitl, P. D., Smith, A. R., Rose, J. P., & Krizan, Z. (2010). The desirability bias in predictions: Going optimistic without leaving realism. *Organizational Behavior and Human Decision Processes, 111,* 33-47.

Xu, F., & Garcia, V. (2008). Intuitive statistics by 8-month-old infants. *Proceedings of the National Academy of Sciences*, *105*, 5012-5015.

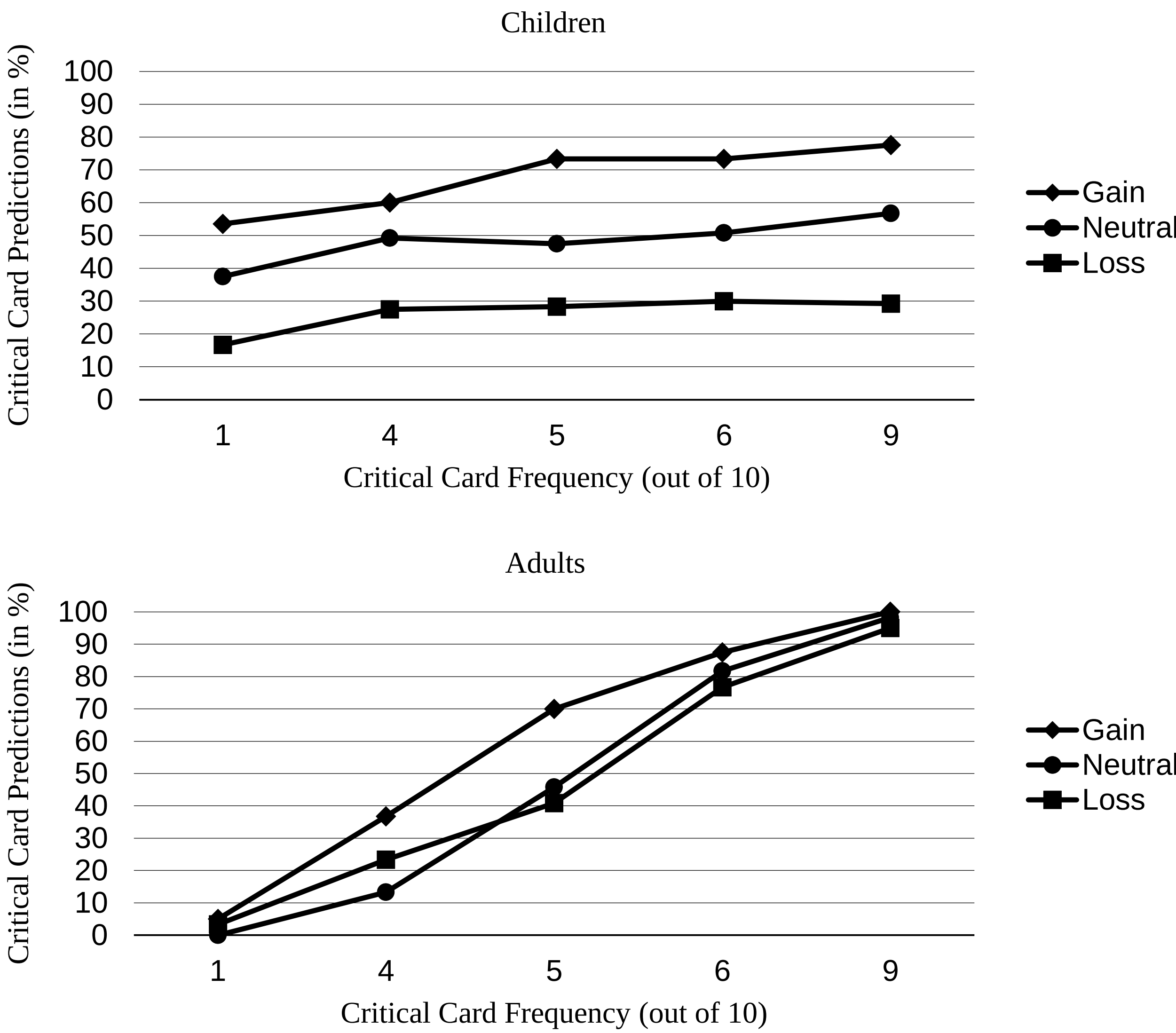

Figure 1. Critical card predictions by frequency of critical card for children and adults. Children were more likely to predict a critical card in the gain condition compared to the loss and neutral conditions and were less likely to predict a critical card in the loss condition compared to the neutral condition. Adults were more likely to predict a critical card in the gain condition compared to the loss and neutral conditions but were no less likely to predict a critical card in the loss condition compared to the neutral condition.

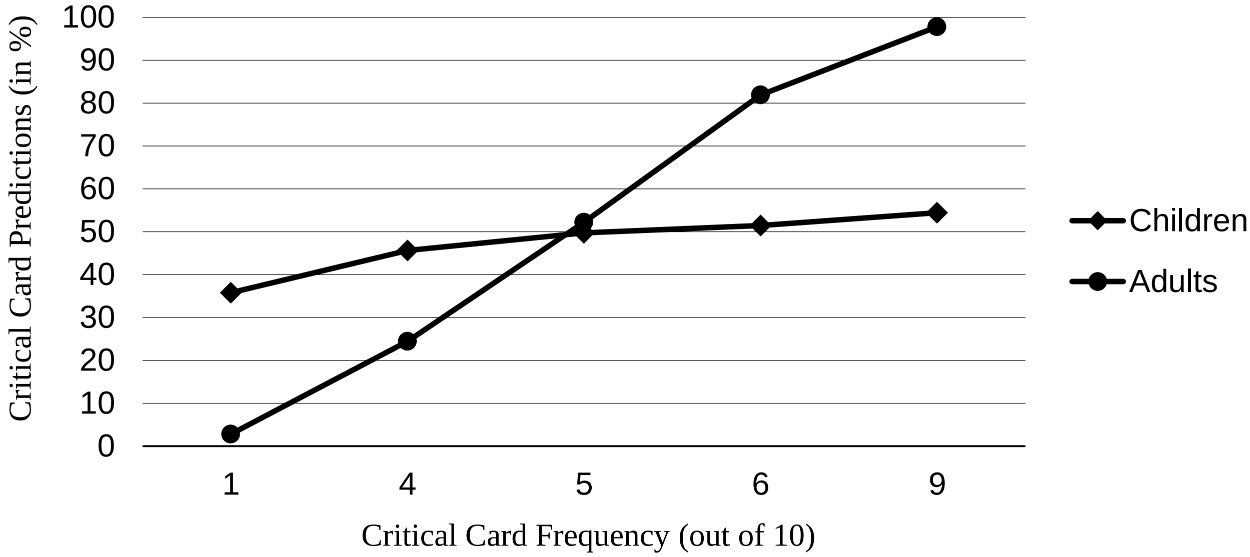

Figure 2. Critical card predictions by frequency of critical card for adults and children. When collapsed across condition, critical card predictions increased as the number of critical cards increased for both adults and children. Compared to children, adults were less likely to predict a critical card for frequencies of one and four and more likely to predict a critical for frequencies of six and nine.

Appendix

IRB Approval Form

IRB <irb@appstate.edu> 10/24/14

to morganzh, kondradrl, smithar3

To: Zachary Morgan

EMAIL

From: Dr. Lisa Grizzard, Institutional Review Board Chairperson
Date: 10/24/2014
RE: Notice of IRB Approval by Expedited Review (under 45 CFR 46.110)
Study #: 15-0076

Study Title: Wishful Thinking
Submission Type: Initial
Expedited Category: (6) Collection of Data from Recordings made for Research Purposes,(7) Research on Group Characteristics or Behavior, or Surveys, Interviews, etc.
Approval Date: 10/24/2014
Expiration Date of Approval: 10/23/2015

The Institutional Review Board (IRB) approved this study for the period indicated above. The IRB found that the research procedures meet the expedited category cited above. IRB approval is limited to the activities described in the IRB approved materials, and extends to the performance of the described activities in the sites identified in the IRB application. In accordance with this approval, IRB findings and approval conditions for the conduct of this research are listed below.

Regulatory and other findings:

The IRB has determined that the research presents minimal risks to participants, adequate provisions are made for soliciting assent of minors, and obtaining the consent of one parent or guardian (45 CFR 46.408).

Approval Conditions:

Appalachian State University Policies: All individuals engaged in research with human participants are responsible for compliance with the University policies and procedures, and IRB determinations.

Principal Investigator Responsibilities: The PI should review the IRB's list of PI responsibilities. The Principal Investigator (PI), or Faculty Advisor if the PI is a student, is ultimately responsible for ensuring the protection of research participants; conducting sound ethical research that complies with federal regulations, University policy and procedures; and maintaining study records.

Modifications and Addendums: IRB approval must be sought and obtained for any proposed modification or addendum (e.g., a change in procedure, personnel, study location, study instruments) to the IRB approved protocol, and informed consent form before changes may be implemented, unless changes are necessary to eliminate apparent immediate hazards to participants. Changes to eliminate apparent immediate hazards must be reported promptly to the IRB.

Approval Expiration and Continuing Review: The PI is responsible for requesting continuing review in a timely manner and receiving continuing approval for the duration of the research with human participants. Lapses in approval should be avoided to protect the welfare of enrolled participants. If approval expires, all research activities with human participants must cease.

Prompt Reporting of Events: Unanticipated Problems involving risks to participants or others; serious or continuing noncompliance with IRB requirements and determinations; and suspension or termination of IRB approval by an external entity, must be promptly reported to the IRB.

Vita

Zachary Hollingsworth Morgan was born in Charlotte, North Carolina to John "Mett" and Kimberly Morgan. He graduated from Myers Park High School in May 2010. Later that year, he entered Appalachian State University (ASU) to study Psychology, Statistics, and Spanish, and was awarded the Bachelor of Arts degree in the spring of 2014. In the fall of 2014, he accepted a research assistantship in Developmental Psychology at ASU and began study towards a Master of Arts degree. The M.A. was awarded in May 2016. In the fall of 2016, Zach continued his work at ASU and served as an adjunct professor of Social Psychology.

Throughout his time at Appalachian State, Mr. Morgan presided as the president of the disc golf team, leading ASU to a berth in the National Collegiate Disc Golf Championships three years in a row.