Predicting the Attitudes, Interests, and Opinions of the Average American Consumer:

Has Anything Changed in the Last Quarter Century?

SHELLE SANTANTA

VICKI G. MORWITZ

DARIA DZYABURA*
* All authors contributed equally and order of authorship is reverse alphabetical. Shelle Santana is a Doctoral candidate in the Marketing department, Vicki G. Morwitz is the Harvey Golub Professor of Business Leadership and Professor of Marketing (Phone: 212.998.0518, Fax: 212.995.4006, email: vmorwitz@stern.nyu.edu) and Daria Dzyabura is Assistant Professor (Phone: 212.992.6843, Fax: 212.995.4006, email: ddzyabur@stern.nyu.edu), all at the Stern School of Business, New York University. Address: 40 West Fourth Street, New York, NY 10012. Please direct all correspondence to Shelle Santana (Phone: 212.998.0729, Fax: 212.995.4006, email: ssantana@stern.nyu.edu). The authors thank Rajan Sambandam at TRC for providing access to and fielding our survey with their panel. The authors also thank Steve Hoch for his advice and suggestions, and Xinyi Duan for her invaluable assistance in data analysis.
Contribution Statement

Over 25 years ago, Hoch (1988) examined how accurately different groups of people could predict the attitudes, interests, and opinions (AIO) of average Americans. He found that: 1) marketing experts were no better in their predictions than novices, and 2) in contrast to false consensus research, people did not project their own views enough when making predictions about others. He concluded that in the absence of valid sources of outside information, people are well served by projecting their own views more than they do. This effects paper takes a look at whether consumers’ ability to accurately predict what others are like has changed in any substantive way over this long time period. This is an important question to ask given the proliferation of data now available to consumers about others as well as firms’ increased reliance on consumer-generated marketing techniques and their decreased reliance on traditional marketing research data. We also extend the original work by Hoch to examine whether: 1) consumers can more accurately predict how average Americans answer general interest questions for which more valid outside information should be available than for AIO questions, and 2) subjects commonly used in consumer behavior research - undergraduate students and mTurk panel members - differ from a representative sample in how accurately they can predict the attitudes, interests, opinions, demographics, and behaviors of average Americans.
Abstract

Over 25 years ago Hoch (1988) found that every day consumers performed poorly at predicting the attitudes, interests, and opinions (AIO) of average Americans. We examined whether every day consumers’ predictive accuracy has changed in any substantive way since then, given the massive increase in data available about others through specialized and targeted media sources, social media, and user generated content. We compared the predictive accuracy of a representative sample of consumers to two groups commonly used by academic researchers: undergraduate students and workers on Mechanical Turk, and examined whether accuracy was higher for general interest questions than for AIO questions. We found that every day consumers now have less valid outside information and make less accurate predictions for AIO questions than in the past. Accuracy was higher for general interest questions. Undergraduate students and Turkers were as accurate in their predictions as a representative sample consumers.
INTRODUCTION

Over a quarter century ago, Hoch (1988) examined how accurately marketing experts and every day consumers could predict the attitudes, interests, and opinions (AIO) of average Americans. He found that neither group was particularly skilled at intuitively predicting the AIO of average Americans. Even marketing experts, whose living depended on it and who were often exposed to data about consumers in their daily work, did not have good intuition about average Americans. In fact, Hoch showed that experts were no more accurate in their predictions than were novices.

Prior research on the false consensus effect (Ross, Greene, and House 1977) suggests that when people make predictions about others, they will overweight their own opinions, and falsely assume that others are similar to them. Based on these findings, it is plausible that Hoch’s (1988) experts and every day consumers were inaccurate because they falsely assumed that other consumers shared their own attitudes, interests, and opinions. However, this was not the case. Hoch (1987) found that even when people over-predicted consensus, they still would have been more accurate if they had relied even more on their own views when making predictions about others. In subsequent research, Hoch (1988) found that both experts and every day consumers would have been more accurate in their predictions of average Americans if they had relied more heavily on their own views and projected these more in making their predictions. Hoch (1987, 1988) concluded that in the absence of valid sources of outside information, people are well served by projecting their own views more than they do.

The objective of our research is to examine the extent to which consumers’ ability to accurately predict what others are like has changed over the past quarter century. This is an
important question for several reasons. First, information technology has led to dramatic changes in the depth, breadth, and speed at which consumers can acquire information about others. Therefore it is an appropriate question to ask whether access to this information has changed consumers’ predictive ability. If consumers now have access to better information than they did in the 1980’s, they may no longer need to project from their own AIO to accurately make predictions about others. Second, firms are increasingly relying on input from a small self-selected set of every day consumers when making their marketing decisions (e.g., consumer generated advertisements, promotions, and product design) and from insights gleaned from consumer generated input, social media, and crowd sourcing, when assessing the attitudes, opinions, and preferences of their target customers. Firms, for example, often rely on user generated content to assess consumers’ brand perceptions, the importance they place on different product attributes, and more generally their product-related needs. Therefore the extent to which these consumers can accurately articulate the views of other consumers is an important question.

To that end we examine how well a representative sample of American consumers can predict the answers of average Americans to most of the same AIO questions used by Hoch (1987, 1988). We also extend the original work by Hoch to examine whether subjects commonly used in consumer behavior research - undergraduate students and mTurk panel members - differ from a representative sample in how accurately they can predict these AIO responses. Finally, we compare how accurately members of these groups can predict how average Americans respond to general interest questions (e.g., “I am living below the poverty line,” “I drink coffee on a daily basis,” “I own a gun”) versus the more private AIO questions (e.g., “I would rather spend a quiet evening at home than go out to a party”), since information about the former types of questions may be more available through outside sources of information than for the AIO
questions used in Hoch (1987, 1988).

We find that every day consumers now have less valid outside information and make less accurate predictions about the AIO of average Americans than Hoch (1988) found in the late 1980’s. Like Hoch, we found that people under-project and can increase their accuracy by relying more on their own position when making predictions about others. We also find the validity of outside information about AIOs of average Americans has decreased since the time of Hoch, despite the proliferation information available through the internet, media, and social platforms. We extend the findings of Hoch and find that accuracy is higher when people make predictions for general interest questions, for which they have valid outside information. Finally we find that undergraduate business students and workers on Mechanical Turk are as accurate in their predictions as are a representative sample of every day consumers. In the next section we discuss the importance of the problem we investigate, and delineate and motivate our research goals.

**BACKGROUND**

Research Motivation

Consumers have always been concerned about what others think and about what products, services, and experiences they value (Bearden and Etze 1982), but they are now increasingly using technology to obtain information about other consumers’ views about products and services and to share their opinions, thoughts, and preferences about products and services with others. With these technological advances consumers are now participating in
virtual communities to gather and share knowledge from others (Chiu, Hsu, and Wang 2006) and are relying on the opinions of many others in making their own purchase decisions (Dhar and Chang 2009; Ye et al. 2011). The value of these opinions may well depend on the degree to which one consumer’s preferences are relevant to those of another. For example, a consumer who is reviewing a hotel experience on an online platform may try to anticipate the preferences of others in deciding what information to share and how to best summarize his or her own experience. A consumer who is reading a hotel review written by someone else may try to determine how closely his or her own preferences align with those of the writer in trying to determine how much to weight the review in his or her own decision making. Thus an important issue is how accurately consumers are able to predict the preferences and opinions of others.

The question of how accurately people can predict the attitudes, preferences, interests, and opinions of others has been of interest to consumer researchers for some time (Alba and Hutchinson 2000), given its importance in a wide variety of important business and consumer-related contexts (Brenner and Bilgin 2011; Lerouge, Davy and Warlop 2006). For example, marketers try to predict which products, services, prices, promotions, and advertisements will best serve their target customers and often rely on their intuition or on a few key informants in making these decisions. In addition, marketers are increasingly shunning reliance on traditional marketing research and the use high-paid experts in favor of reliance on everyday consumers to help develop new products, promotions, and advertising campaigns. For example, Frito-Lay North America conducts a “Do Us A Flavor” competition each year where consumers submit their ideas for a new flavor of potato chip, and the winner receives $1 million (Market Watch 2014). So the ability of consumers to accurately predict the attitudes, likes, and dislikes of other consumers is becoming more critical.
In an organizational setting, when managers make decisions on behalf of their employees, they need to be able to predict how their employees will react to these decisions. When one bargains or engages in negotiations, one’s predictions of what is important to the other party will affect the success of the negotiations (Bottom and Paese 1997). Politicians often try to take into consideration how their constituents will react to different policies and their success in politics may well depend on how well they can anticipate the preferences of voters. News reporters and newscasters try to anticipate which stories will interest their audience. Various agents help consumers decide what clothes to buy, what wine to select with dinner, what houses to buy, and what to purchase for family parties and major events, and in all these cases need to make judgments about the preferences of their clients (Solomon 1986). Whether or not these clients are satisfied and recommend these agents to other customers will likely depend on how well they were able to find products that matched their clients’ preferences. In addition to professional agents, friends and family members act as agents when they buy presents for others, plan surprise parties, and vacations for others (West 1996), where the success of those efforts depend in large part on their ability to successfully predict the preferences of others. Similarly, when consumers simply make a product or service recommendation to others, the accuracy of their recommendations may depend on how well they can predict what others are like (Alba and Hutchinson 2000). More generally, for joint and group activities, such as which restaurant to go to or which movie to attend, the decision makers for the group often need to predict and take into consideration the preferences of others in the group. Even when consumers make individual purchase decisions, since they value social approval, they also sometimes make predictions about and take into consideration the preferences of others (Bearden and Etze 1982; Childers and Rao 1992; Lerouge and Warlop 2006). For example, when consumers buy a new car that they hope
will impress others, a new suit to impress a potential new boss, or a new outfit to woo a date, they rely on predictions they make about how others will react to their purchases.

While these predictions are often based on intuition, they are not always. In some cases, data are available to better inform decision makers about the preferences of others. For example, politicians can poll their constituents to gather data on their attitudes, opinions, and preferences. Similarly, marketers often gather marketing research data to determine the attitudes, opinions, and preferences of their target customers and use those data to design products that provide value to their customers. However, in many cases politicians, marketers, agents, and friends, lack these types of hard data about others’ preferences. When this occurs, they need to make decisions based on other sources of information such as their own opinions, those of people they know, or things they observe through the media and in the world.

Hoch (1987) took a Brunswikian (Brunswik 1952) perspective. He discussed how when people are making a prediction about a target population, they have multiple cues or sources of information to rely on. First they have their own attitudes, opinions, and behaviors. Second they may have some outside or idiosyncratic knowledge that might be relevant. Third, they have may have knowledge or a belief about the degree to which they themselves are similar or different from the target. Hoch (1987) discussed how the accuracy of people’s predictions about a target then will depend on whether they can identify the cues available to them that are predictive of the target and whether they weight them correctly. Thus inaccuracy in prediction can come from incorrect weighting of their own views, lack of availability of other helpful information, or inappropriate weighting of that information, when it is available.

Thus, two important questions are (1) how accurate are these intuitive predictions, and (2) what factors influence their accuracy. Hoch (1988) compared the ability of three different groups
of people to accurately predict the AIO of average Americans. He examined how the actual similarity between each group and the target group for which they were making predictions, the amount each group projects their own views in making predictions, and the validity of the other information each group used combined to affect their predictive accuracy. His central findings were that overall predictive accuracy was low and that marketing experts were no more accurate than a representative sample of Americans in predicting the AIO’s of average Americans.

Another important finding was that, in contrast to prior research on the false consensus, people tended to underweight their own views when making predictions about others. False consensus research would predict that people would tend to view others as more similar to themselves than they really are. However from Hoch’s work and the work of others, it is now more widely known that people may not always display false consensus effects (Krueger 1998), and that projection can be useful, especially when there is not much other information to rely on when making predictions about others (Alba and Hutchinson 2000; Dawes and Mulford 1996).

Hoch concluded that, although both experts and novices could have improved their predictive accuracy by projecting their own AIOs more, the reasons for their respective under-projection differed. Relative to novices, he argued that experts had access to more reliable information, but were less similar to average Americans. Therefore, the data they held about their own views was less helpful in predicting the views of average Americans. In contrast, every day consumers (i.e., novices), had less access to useful outside information than did experts but were more similar to the target than the experts were. Thus their own views did provide useful information, and their low accuracy came from not projecting their own views as much as they could have.
Primary Research Goal

Our primary goal is to examine, in the quarter century that has passed since Hoch published this work, what, if anything, has changed in how every day consumers make predictions about others. One thing that has changed in this time period is that information technology has led to dramatic changes in the depth, breadth, and speed at which consumers can acquire information about others. Every day consumers now have at their fingertips a vast array of new data sources, that did not exist 25 years ago, to inform them of the views of others including more targeted media, such as Fox news or MSNBC, social media platforms such as Facebook and Twitter, and user generated reviews from platforms such as Yelp and Tripadvisor. Marketers too are relying more and more on insights gleaned from consumer generated input, social media, and crowd sourcing, when assessing the attitudes, opinions, and preferences of their target customers.

This proliferation of information sources raises the question of whether these new inputs offer helpful data for those who need to make intuitive predictions about others. To that end, our goal was to see whether the predictive accuracy of every day consumers has changed in the years since Hoch (1988) was published. Our research is exploratory in nature as we believe that there are three plausible possibilities. First, there is good reason to believe that there has been no change in the last quarter century in how accurately every day consumers make predictions. Projection is a fundamental and a difficult psychological process. It is as difficult now as it was in the 1980’s and therefore it is likely that people still fall prey to the same errors in how they make predictions as they did in the past. That is, though they now may have more and/or better data, they may still fail to use it appropriately, and may still tend to under-project their own
Second, every day consumers may be better at predicting the views of average Americans than they were a quarter century ago. Given the technological advances of the last quarter century, many people have at their fingertips the ability to gather more facts and figures about others than they ever could before. People are connected online to a more diverse group of people through social media and user generated content, than they were before. Research suggests that the time people spend with others online neither increases nor decreases the time they spend face to face or on the phone with others (Wellman et al. 2001), so net people may just have more information about others than they did in the past. This may lead them to be able to better calibrate how similar or dissimilar their own beliefs, attitudes, opinions, and behaviors are than others and may give them access to more valid outside information.

Third, every day consumers could be worse at predicting what consumers are like than they were a quarter century ago. People are watching fewer national and local news broadcasts on television, reading fewer newspapers, and spend less time listening to radio news broadcasts (Chyi 2009) that could potentially inform them about the general population. While one may logically think that people have simply moved from traditional news media to obtaining news from the internet, studies show that the majority of Americans are not obtaining news online either, and that those who do tend to be the same people who also obtain news from traditional media sources (Ahlers 2006; Pew Research Center for the People & the Press 2004). Other studies show that people are obtaining some news from non-traditional media which gives rise to a selection bias. For example, nearly half of all adult users of Facebook (or 30% of the overall population) obtain news from Facebook (Mitchell et al. 2013), which gives them information about news that is of interest to their friends, and not necessarily to the general population.
Further, even among those who do still follow the news, there is a growing availability and tendency to follow partisan news broadcasts like Fox News and CNBC. These sources tend to confirm, rather than challenge, the viewers’ existing views (Prior 2005; Sunstein 2001) and thus may lead viewers to be less informed about the views of dissimilar others. Thus, with their increasing ability to select which media to use as information sources, people may have less opportunity to learn about others as reflected in reports by the news media.

In addition, consumers may be spending more time reading comments from people who they are connected to via social media networks, such as Facebook and Twitter, and reading idiosyncratic accounts from individual consumers via user generated content. Since these social networks and the consumers who contribute to online communities are not necessary representative of the larger population of every day American consumers, the views people are exposed to may provide a biased assessment of the attitudes, opinions, beliefs, and behaviors of others. For example, following the shootings in Newtown, Connecticut, two thirds of the conversation on Twitter related to calls for stricter gun control, while a representative sample of Americans was more evenly split between a desire for more gun control and a desire to protect gun rights (Matsa and Mitchell 2014). Thus though consumers may have access to more information than they did in the 1980’s, the quality of the information they have access to may be worse than it was years ago and reliance on it may lead to less accurate predictions.

Additional Research Goals

In addition to our main goal of examining how, if at all, the predictive accuracy of every day consumers has changed in the past quarter century, we also explore two secondary goals.
First, we examine whether every day consumers are better able to predict how average Americans respond to general interest questions (e.g., “I am living below the poverty line,” “I drink coffee on a daily basis,” “I own a gun”) than the more private attitudes, interests, and opinions of others. The AIO questions asked in the Hoch (1988) study may have been especially challenging for people as they represented beliefs that people may hold privately, they are not things that can generally observed, and they are not topics that are often examined in public opinion polls that may be reported in the news. For example, one of Hoch’s original questions was “The dirt you can’t see is worse than the dirt you can see.” We wager to guess that responses to this or similar questions are not frequently reported in news reports of public opinion polls. On the other hand, the news often does report statistics on topics the media believes are of general interest, such as how many people live in poverty, how many regularly attend church or synagogue or some other religious service, and how many own a gun. In addition, these are things that people can generally observe about others. Thus answers to general interest questions might be easier for people to estimate because they have access to more valid outside information. We explore the relative accuracy of every day consumers’ predictions about AIO items versus more general interest items.

Hoch (1988) focused on the predictions of every day consumers and marketing experts (and also of MBA students). These are not the only groups of people who need to make predictions about what others are like. Academics who work in the area of consumer research also increasingly rely on the input of their research study participants – typically undergraduate students (Peterson 2001), but increasingly workers on Mechanical Turk, and more rarely representative samples of consumers – to make inferences about how the larger population of consumers thinks and behaves. Therefore a third goal of our research is to assess how well the
samples we typically use in academic research can predict the views of more representative sets of consumers.

To that end, we use three different samples in our work. The first is a marketing research firm’s proprietary nationally representative online sample of Americans, which is similar to and can be compared to the representative sample in Hoch (1988). The second is a sample of workers on Mechanical Turk (mTurk). This sample can be viewed as a modern day counterpart to the convenience sample of Hoch (Chicago-based focus group participants), but it is also a group that is of interest in and of itself. mTurk is increasingly being used for research by academics and by firms so it is important to know how similar those who participate in mTurk studies are to average Americans, and how accurately they can make predictions about average Americans. The third is a sample of undergraduate business students from a large private northeastern university. Undergraduate business students are also a very common subject base of much academic research in marketing.

The use of mTurk and undergraduate business students is not without controversy in our field so we examine: (1) how similar the views of members of these groups are to those of average Americans and (2) how well they can make predictions about average Americans. Johnson (2001) discusses how there is rarely a theoretic reason to use student subjects for research, and noted that rather, we tend to use them because they are less costly in terms of time and money than other, more representative samples. He and others (Peterson 2001; Sears 1986) raise the issue that student subjects are less diverse than other samples, and, importantly, display less variation on constructs of interest to consumer researchers. As Sears (1986) reported, undergraduate students tend to come from a narrow age range and are highly educated. Compared to adults, undergraduate students often have less well formed attitudes, are more
prone to change their attitudes, have a weaker sense of self, differ in their peer relationships, and have stronger cognitive skills and tendencies to comply to authority figures. Of particular interest for the current research, Sears discusses how undergraduate students’ self-assessments and attitudes are often influenced by comparisons they make to others. In addition, since young people in general are more egocentric, they may be more prone to assume others are like them. Johnson (2001) argues that our reliance on student samples also decreases the credibility of our research to managers and policy maker audiences. Peterson (2001) noted that while there has been much debate in ours and other fields about the reliance on student subjects, there has been less research on the appropriateness of their use, something we hope to address.

Johnson (2001) predicts that with the diffusion of information technology we would see reduced reliance on student subjects, greater use of more diverse panels including online ones, use of other data sets, and larger sample sizes. A perusal of our academic journals confirms that Johnson’s predictions have come true, as we are seeing an increased number of studies conducted using online panels, and in particular with studies using Mechanical Turk (mTurk). mTurk is a crowdsourcing site where requestors can post jobs and workers can select jobs and be paid for doing them (Mason and Suri 2012). Over 100,000 people participate as workers on mTurk (commonly called Turkers), and its use for research has grown in popularity because the sample is large, willing to do research for very small payments, and is more demographically diverse than most student subject pools (Burhmester, Kwang, and Gosling 2011). mTurk offers a low cost and fast way to conduct cross sectional studies, as well as studies that require a panel, since its panel of workers is quite stable over time (Mason and Suri 2012).

A few studies have been conducted to determine whether Turkers respond differently to research than other subject populations. Burhmester et al. (2011) asked participants on mTurk
and a large internet sample to complete six personality scales. They found no differences in the results and noted that there was high test-retest reliability on mTurk. Paolacci, Chandler, and Iperiotis (2010) gave traditional judgment and decision making problems like the Asian disease and the Linda problem to Turkers, subject pool members at a Midwestern U.S. university, and people recruited through online discussion boards. They found very similar results to these problems across the samples.

Despite this evidence that Turkers behave similarly to participants in other research samples and despite the greater diversity in mTurk samples relative to student samples, some concerns about the reliance on mTurk for research remain. One concern is whether samples from mTurk, like other online platforms, adequately represent the true populations of interest (Johnson 2001; Paolacci et al. 2010). For example, Paolacci et al. (2010) report that Turkers are younger and more educated but have lower income than the general population. A second concern is whether the data quality from mTurk is high enough, especially given that Turkers are paid so little for their research participation (Paolacci et al. 2010). We therefore examine the extent to which Turkers are similar to a more representative sample and the degree to which they can accurately predict the views of average Americans in exchange for a small payment.

In summary, the question of how well different groups of individuals can predict the AIOs of other Americans and what information they rely on to make these predictions remains an important one. When Hoch (1988) first studied this question over 25 years ago, the primary interest was whether marketing experts (and MBAs) were any better than novices (defined as convenience and representative samples of adults) at predicting the AIOs of the average American. He found that they were not, and that every day consumers could improve their predictive accuracy by projecting their own views onto the target population since other valid
information was not available to them. We revisit the question of prediction accuracy with different goals in mind. First, rather than focus on the difference between experts and novices, we focus on the difference between three groups of novices—business school undergraduates, a convenience sample of mTurk workers, and a representative sample of consumers drawn from a marketing research firm’s representative online panel. We focus on the quality of novice predictions since the ability of one group of consumers to accurately predict what other consumers are like is an important question both for individuals and for businesses alike. Second, we explore whether the quality of every day consumers’ predictions differ for private AIO questions versus more general interest questions where more outside information should be available. Finally, our work sheds light on the debate over the validity of using mTurk and undergraduate subjects in consumer behavior research.

The remainder of this paper is organized in the following manner. First, we review the model and approach used by Hoch (1987, 1988). We then present our study and discuss the findings for our three populations—a representative sample, Turkers, and undergraduate students—for both Hoch’s (1988) original AIO questions and for a new set of general interest questions. We conclude with a discussion of how our findings compare to those of Hoch, our incremental contribution, and what remains to be explored in future research.

**Hoch’s (1987, 1988) Model of the Interpersonal Prediction Task**

Hoch (1987, 1988) presented a model of how an individual makes an interpersonal prediction ($p$) about the position of some target of interest ($t$), which in Hoch (1988) was the position of average Americans. In his framework (see Figure in Hoch 1988), the individual
making the prediction has two pieces of information: his own position \((o)\) and all other relevant information about the target \((z)\).

Hoch examined the relationship between these factors using correlations. In this framework, \(r(t,p)\) represents predictive accuracy or the correlation between the actual position of the target \((t)\) and the individual’s prediction about the target \((p)\). \(r(t,o)\) represents the actual similarity between the target and the individual making the prediction, \(r(p,o)\) represents how much the individual projects his own position in making a prediction, and \(r(t,z)\) represents the validity of other sources of information.

Using this framework, predictive accuracy is affected both by identification of the relevant cues (own position and other information) and the weight that is placed on those cues. As Hoch (1987, 1988) discussed, one’s own position should be a highly accessible source of information and therefore it is likely to be identified. The important issue affecting accuracy then will be how one’s own position is weighted, or more specifically, whether it is relied on too much or too little. In contrast he discusses for other information how predictive accuracy will be affected both by the identification of other information and how much weight is placed on it. As Hoch (1988) stated, “Because relevant other knowledge is fragmentary, decision makers (DMs) may be distracted by salient but irrelevant details.”

Hoch (1987, 1988) used a one-cue lens model (Hursch, Hammond, and Hursch 1964) to estimate how individuals making predictions using their own position and outside information combine these two sources of information. In this model a respondent’s own position \((o)\) is observed, but outside information \((z)\) is inferred from the model residuals. Hoch (1987) discusses how while in theory outside information, or some aspects of it, could be measured, in practice it is difficult since the information is likely to be highly individualistic and idiosyncratic. His
model of prediction is:

\[ p = b_o + \sqrt{1 - b^2} z \]  \hspace{1cm} (1)

Hoch (1987, 1988) also derived the optimal manner in which an individual making predictions should assign weights to his own position \((o)\) and to other relevant information \((z)\) to maximize predictive accuracy. He showed that these weights depend on the correlations of \(o\) and \(z\) with the target position \((t)\). If \(b'\) represents the optimal weight the individual making the prediction should assign to his own position \((o)\), and \(p'\) represents what the prediction would be when own position and other information are optimally weighted then:

\[ p' = b'o + \sqrt{1 - b'^2} z \]  \hspace{1cm} (2)

To maximize predictive accuracy, \(r(t, p')\), the weighting of one’s own position, \(b'\) should be selected in order to maximize \(r(t, p')\): Since

\[ r(t, p') = b'r(t, o) + \sqrt{1 - b'^2} r(t, z), \]  \hspace{1cm} (3)

then

\[ \frac{b'}{\sqrt{1 - b'^2}} = \frac{r(t, o)}{r(t, z)} \]  \hspace{1cm} (4)

Thus, the optimal weight of the individual’s own position, \(b'\) in comparison to the weight given to the other information, \(\sqrt{1 - b'^2}\) should be proportional to the ratio of \(r(t, o)\), the actual similarity between the target and the individual making the prediction, and \(r(t, z)\), the validity of other sources of information.

We used Hoch’s (1987, 1988) framework, model, and methods in the study that follows.

**STUDY**

This study is very similar to Hoch (1988). Like Hoch, we compare the ability of three
different groups to predict the AIOs of American consumers. In contrast to Hoch, who used every day consumers, marketing professionals, and MBA students, we used every day consumers, Turkers, and undergraduate business students. We used almost all of the same AIO questions as in Hoch (1988) but also augmented his questions with 14 general interest questions. We used these data to estimate the parameters of the same model used by Hoch and described in the previous section for each of the samples and for both sets of questions.

Stimuli used in the current research

We used 20 of the same 21 AIO items as in Hoch (1988). The items that Hoch used were selected from an annual Life Style study that was conducted by DDB Needham in the late 1980’s. The one item from the original Hoch paper that we changed was his number 13, “Communism is the greatest peril in the world today.” We opted to revise this item because we felt it was more relevant to the time period for which it was originally asked than for the current time period. We replaced this item with a new one, “Terrorism is the greatest peril in the world today.”

Participants in all three groups had to estimate what percent of married American men and women would agree with the statement made in each item in 2011, the year the study was conducted, as well as in 1986 the year Hoch’s data were collected. We did not analyze the predictions for 1986, since we did not have those respondents’ raw data, and therefore could not calculate the \( r(t,o) \) between our sample and the 1986 sample. Therefore, we do not discuss the 1986 response data further. Participants also had to indicate their own level of agreement with these statements on a 7-point scale ranging from strongly disagree (1) to strongly agree (7). In
addition we added 14 new general interest items to the survey. For these items, participants were asked to estimate the percentage of American adults, age 18 or older, that would respond positively to a set of general interest questions (regardless of whether they were married or not), and they were also asked to respond to the same statements for themselves. For example, participants were asked to estimate the percentage of Americans that pay their credit card bills in full each month, and they were also asked if they paid their credit card bills in full each month. These two sets of survey items are included in Appendix A and B, respectively.

Respondents

Representative prediction sample. The representative sample groups (the prediction and the truth groups combined) consisted of 172 consumers sampled from a representative online marketing research panel. A link to the questionnaire was sent by email to 422 panel members. Of the 422 panelists who were sent the link, 153 either did not respond, did not provide their consent, or did not complete the survey, which was very long and took between 15-30 minutes to complete. From the 269 panelists who consented and completed the survey, we removed the data from 97 respondents who did not answer or who incorrectly answered an instructional manipulation check. Please note this rate is in line with the percent of participants who have failed instructional manipulation checks in the studies conducted by Oppenheimer, Meyvis, and Davidenko (2009), which developed and introduced the instructional manipulation check. Thus, the analysis below is based on the remaining 172 respondents. We selected this sample size to match what was used in Hoch (1988) where surveys were sent to 275 consumers and complete data were available from 141 of them.

mTurk. We recruited 299 U.S. adults with a HIT approval rate of 98 or better to
participate in this study. Participants were paid $.85 upon completion of the study. We removed the data from 68 participants because they did not finish the questionnaire, and also removed the data from another 20 participants who failed an instructional manipulation check. The analysis was based on the remaining 211 respondents. This sample size was also selected to match the size of the sample used in Hoch (1988) for the representative sample.

Undergraduate Business students. 149 undergraduate business students from a large northeastern university participated in this study as part of their Introduction to Marketing subject pool class requirement. They received partial class credit for their participation. These participants came to the research laboratory at a specified time, and completed the survey on a computer after clicking on a link to the survey. We eliminated 36 of the 149 students due to failed instructional manipulation check, and we removed another 6 for incomplete responses. The analysis for the undergraduate students is therefore based on the remaining 107 participants. This sample size was the largest that the authors could obtain from this subject pool during this time period.

Representative truth sample. The truth or norms for the actual answers were obtained from the same online panel as the representative prediction sample. When computing the accuracy of this sample’s predictions, we used a Leave One Out Cross Validation (LOOCV) procedure to avoid biasing the results. The procedure is described in detail in the Data Analysis section.

Procedure

The surveys were conducted online for all three sample groups. Each respondent first read a consent page, and if they agreed to participate, clicked through to an online survey where
they completed a questionnaire consisting of 20 of the 21 original AIO questions used in Hoch (1988), the one revised AIO question, and the 14 additional general interest questions. The undergraduates only responded to 13 of the general interest questions as they were not asked the question about whether they had a college degree.

For each of the 21 AIO items, respondents made three different judgments. Using a seven-point strongly disagree (1) to strongly agree (7) scale, they indicated their own personal level of agreement with each item. Next, using a zero to 100 percent scale, they estimated separately the percentage of married men and the percentage of married women who they believed agreed with each item. For the 14 general interest items, respondents estimated the percentage of Americans, 18 and older, who either met the criteria or agreed with each item by entering a percentage between 0 – 100 in a space provided. Respondents were reminded that these questions were in reference to all Americans age 18 or older, not just married Americans. Participants were also asked to state their own position for each item. For both the AIO and the general interest questions, and consistent with the original Hoch (1988) paper, we controlled for possible order effects by counterbalancing whether participants responded for themselves first and then the target, the target and then themselves, or themselves and the target simultaneously.

We also included a few additional items where participants were asked to estimate the market performance (i.e., market penetration or market share) of some products for a different research project. These items were not examined for the current research.

Toward the end of the survey we also asked respondents to complete the items for three different individual difference scales—Self Monitoring (Snyder 1974), Social Comparison Orientation (Gibbons and Buunk 1999), and Perspective Taking (Davis 1980). We did not find any effects related to any of these scales and therefore do not discuss them further. We also asked
a few additional attitudinal questions for use in another research project. The survey concluded with an instructional manipulation check and several demographic measures which for the representative and mTurk samples included gender, household income, education level, occupation, marital status, and state of residence, and for the undergraduates included gender, age, the number of years lived in the U.S., and country of origin.

**ANALYSIS**

Similar to Hoch, we compute each participant’s projection, predictive accuracy, and actual similarity, and infer the validity of other information and optimal weighting of own position. We report the means and 95% confidence intervals of the population distributions. We then compare the resulting confidence intervals to the values reported in Hoch (1988). Note that this test is conservative since we do not have the confidence intervals for Hoch’s estimates and therefore can only test whether our confidence intervals overlap with his point estimates.

When analyzing the data from the representative sample, we use a Leave One Out Cross Validation procedure (LOOCV) to ensure that the focus respondent is not part of the truth sample, and to use as much of the data as possible (see, e.g. Bishop 2006). The procedure involves taking a single respondent, computing the truth \( (t) \) value using the population excluding the focus respondent, and then computing the respondent’s predictive accuracy, projection, and actual similarity with respect to this truth value. Then, the same procedure is repeated for each respondent in the sample. This ensures that the focus respondent’s reported values do not enter the truth estimate and thus bias the similarity, projection, and predictive ability estimates.
RESULTS

As was done in Hoch (1988), we calculated the actual similarity, $r(t,o)$, the level of projection, $r(p,o)$, the validity of other information, $r(t,z)$, and the predictive accuracy of each group $r(t,p)$. We also derived the optimal projection weighting of own position that each group should have used to maximize their predictive accuracy ($b'$). We found no significant gender effects, so therefore, like Hoch (1988), report means averaged over male and female target populations in our tables.

Table 1 shows the actual similarity $r(t,o)$, level of projection $r(p,o)$, validity of other information $r(t,z)$, and the predictive accuracy $r(t,p)$ for the original Hoch convenience and representative samples and for our three new samples. We focus first on a comparison of the 21 AIO items and then move on to a discussion of the 14 general interest questions.

--- Insert Table 1 around here ---

Analysis of the 21 AIO Items

*Similarity.* The first column of Table 1, actual similarity ($r(t,o)$), reflects how similar each sample’s own responses to the AIOs were to the actual responses of the target population of average Americans. Like Hoch, our convenience and representative samples differed significantly from each other, with the representative sample being more similar to the target population than the convenience sample. This is not surprising since the representative sample was designed to reflect the demographics of average Americans and the mTurk sample was not. Our representative sample was also significantly more representative of the target than was the
case in Hoch (1988). This may be because although like Hoch’s, our representative sample and the target were comprised of members from the same survey panel population, in our case a bootstrapping method was used with one sample. Our mTurk sample was also significantly more representative of the target population than was the convenience sample used in Hoch (1988). This too is not so surprising since Hoch’s sample were focus group participants for an advertising agency in the Chicago area, whereas our mTurk sample was comprised of a heterogeneous group of adults from across the U.S. In addition, like a representative group of online panelists (our target population), Turkers represent people who are willing to participate in online surveys. Finally, our undergraduate sample was significantly less representative of the target than were either of Hoch’s samples or our representative and convenience samples, which is logical given the difference in demographics between college sophomores and heterogenous groups of adults.

**Projection.** We next turn our attention to each group’s level of projection. As in Hoch (1988) we operationalized the degree of projection as the correlation between respondents’ own positions and their predictions of the target’s position \( r(p,o) \) across all items. The results are shown in the second column of Table 1.

Our undergraduate sample projected at significantly lower levels than both Hoch’s samples and our other two samples. Although it was possible that undergraduates may have projected more than other groups since they tend to be more egocentric (Sears 1986), this is not what we found. They may have projected less because they realized they were less similar to the target than all of the other groups were and that their own views were not relevant or they may have felt they had access to other more valid sources of outside information, or both.
Additionally, although our representative sample was more similar to our target than Hoch’s representative sample was to theirs, our group did not project any more than Hoch’s group did. Our mTurk sample, which was also more similar to the target than was Hoch’s comparable sample, also projected at significantly lower levels than his convenience sample. These samples may have projected less than Hoch’s since given the rise of the internet since his analysis (and the presumed increase in the validity of the information gleaned from the internet), they may have other information they can rely on. Lastly, although our mTurk and representative samples differed in how similar they were to our target group, they did not differ from each other in their level of projection.

Validity of other information. The members of our two online panels--our representative sample and mTurk-- had significantly lower validity of other information than their Hoch (1988) counterparts. These two groups also did not significantly differ from each other. This is an interesting and surprising outcome, given the ease with which individuals today can learn about others relative to when Hoch (1988) collected his data. However, ease of access and accuracy of information are not the same, and this variable may be a reflection of the latter. Therefore, it’s possible that even though today’s individuals spend more time online and have greater access to information, the information they tend to access about the AIO of others is more idiosyncratic and less representative than information that was available to similar others 25 years ago. In contrast, the validity of other information for the undergraduate business students was significantly higher than that for either of the Hoch samples and for our other two samples. The undergraduate students used in this research attend a large urban university, come from all over the country and the world and are well traveled. Perhaps these factors, combined with being in an
academic environment accounts for their access to more valid other information.

**Predictive accuracy.** The combination of the previous three factors culminates in a measure of predictive accuracy. Conceptually speaking, given a certain level of similarity to a target, the appropriate level of projection and access to and incorporation of valid outside information should result in maximum predictive accuracy by an individual. Our three samples did not differ significantly from each other in their predictive accuracy, even though they did differ from each other on the underlying variables. Our representative and mTurk samples had significantly lower predictive accuracy than Hoch’s representative and convenience samples, despite being more similar to the target than were his samples. This suggests that given the low predictive validity of the information available to them, they should have projected their own opinions more, a topic we explore later. Our undergraduate sample, while having no higher predictive validity than our other samples, did not differ significantly from Hoch’s convenience sample. The fact that the least similar sample group from the target was the most accurate in their predictions of AIOs may seem surprising, though they were also the group with the most valid outside information.

In sum, when it comes to predicting the AIO of average Americans, it appears, that if anything, people perform worse than they did a quarter century ago. Given that our mTurk and representative samples were more similar to their target than their Hoch (1988) counterparts, one might expect that they would have higher predictive accuracy, all else being equal. However, this was not the case. In fact, with the exception of our undergraduates, our samples were less accurate in their predictions than were Hoch’s (1988) samples. These results appeared to be
driven by three factors. First, our samples failed to fully recognize their similarity to the target, and may have under-projected their own positions onto the target population when estimating the latter’s AIO, something we explore later by estimating the optimal projection weights. Second, this tendency on its own need not be problematic, provided that the individuals in question had access to and incorporated valid information from outside sources to help inform their prediction. However this was not the case either. Despite having greater access to technology that could potentially provide access to better information about others than was available in the late 1980’s, this was not the case as the validity of other information was lower for these samples than for Hoch’s. Taken together, these factors resulted in lower accuracy predictions for the representative and convenience samples than what Hoch observed a quarter century ago.

Conversely, our undergraduates, who were dissimilar to all of the other groups (Hoch’s and ours) and to the target population, had the highest prediction accuracy of our three groups (though the difference between it and our other groups was not significant) and was the only one of our groups to have comparable accuracy to any of Hoch’s samples. Though the difference in accuracy between the undergraduate students and the other groups is small, it is interesting to speculate why they were the most accurate group and not the least accurate. The undergraduates sample may have recognized their difference from the target population, prompting them to project less and to instead rely on outside information. What’s even more interesting is that the validity of the outside information recruited by this group of individuals was superior to that of the mTurk and representative groups. Whether our undergrads recruited different information than our online groups or simply assimilated the same information more fully is unknown, but it is ironic that this population would outperform two groups that are far more similar to the target,
Analysis of the 14 General Interest Items

Thus far, our results have been largely consistent with Hoch’s (1988) at a macro level, with important differences at a more granular level. That is, like in Hoch’s (1988) work, our participants showed poor predictive ability for the AIOs of average Americans, which may be caused, at least in part, by under-projection of individuals’ own AIOs. Our participants were more similar to their target than Hoch’s were, but they projected at similar or lower levels. And while one may have expected that in this internet era people could access and use more valid information to make accurate predictions, the validity of information was lower for our representative and mTurk samples than for Hoch’s counterparts. Combined, our groups were even less accurate than Hoch’s with the surprising exception of our undergraduate student group.

Perhaps one reason that we observed low predictive accuracy in the late 1980’s in Hoch’s (1988) work and here was due at least in part to the nature of the task itself. That is, it is simply difficult to accurately predict attitudes and/or other people’s unobservable belief systems and valid outside information for these may simply not be available even in an internet era. For this reason, we also asked our participants a series of general interest questions where we a priori expected that more valid outside information should be available. Since we expected more valid information to be accessible and used in these predictions, we expected that predictive accuracy for these items would be higher than that for the AIOs across all three sample groups, even if the similarity between our groups and the target group was no higher than that for the AIOs.

Similarity. The representative and mTurk samples did not differ significantly in their similarity to the target for these new items compared to their similarity for the AIO items. The
undergraduate sample surprisingly was significantly more similar to the target for these items than for they were for the AIO items.

*Projection.* Although the representative and mTurk samples had not differed in similarity to the target for these items, they did significantly project their position on these items more than they did for the AIO items. We also note that these levels of projection are similar to the levels observed in Hoch (1988) for the AIO items. The undergraduate group, perhaps because they were more similar for these items versus the AIO items, also projected significantly more.

*Validity of other information.* Perhaps the most noteworthy difference between the findings for these items and the AIO ones lies in the validity of other information. For all three groups the validity of outside information was significantly higher than what we observed for these same groups for the AIO items. The representative and mTurk samples did not differ significantly from each other on this measure, but as we saw for the AIO items, the undergraduate students had significantly higher values on this measure than the two other groups.

*Predictive accuracy.* All three groups had significantly higher predictive accuracy for these items than for the AIO items. The three groups did not differ significantly from each other, though we note that at least directionally, again, the undergraduate sample was the most accurate. This large increase likely occurred because even though the actual similarity between the mTurk and representative groups and the target group was similar to that of the AIOs, the validity of the outside information used by these groups to make their predictions was
significantly higher than that for the AIO predictions. Interestingly, though the undergraduate group was more similar to the target for these items than for the AIO ones, projected more, and had more valid information, in the end they were no more accurate than the other groups.

In summary, these data suggest that predicting attitudes and belief systems that are perhaps more private in nature is more difficult for people than predicting general interest questions, regardless of how similar the respondents might be to the target population. Though we found that all three samples projected their own position more for these items compared to the AIO items, the increase in predictive accuracy seems to be primarily due to the validity of outside information. Thus one reason why our samples and Hoch’s (1988) performed so poorly on the AIO items was simply because no outside information was available for these items. Hoch (1988) discussed that when valid outside information is lacking, increasing projection is the best way to increase predictive accuracy. We examine this in the next section, where we compare the optimal weights for projection across the different samples (ours and Hoch’s) and for the two different groups of items.

Optimal Weighting Analyses

As a final step in our analysis, we estimated the optimal weighting that each group should have applied to their own position in making predictions for both sets of items. Table 2 reports the observed weight that members of each group placed on their own position when making predictions ($b$) and the optimal weight ($b'$) that we computed that members of each group should have placed on their own position to maximize their predictive accuracy. We also report the
difference between the optimal and observed weights, and the percent of people who projected too much and who could have increased their predictive accuracy by projecting less. Finally the last two columns of Table 2 report the maximum accuracy if the optimal weights for own position had been used and the amount of improvement if optimal weighting had been used.

--- Insert Table 2 around here ---

*AIO items.* This analysis shows that one main reason why all three of our groups showed low predictive accuracy for the AIO questions was that they under-projected their own positions. Specifically for all three groups the optimal weight for their own position was significantly higher than the actual empirical weight. The vast majority of people projected too little, and in all three groups only a small percent projected too much.

As in Hoch (1988), all three of our groups could have improved their predictive accuracy if they projected their own AIOs more. This seems to be particularly true for the mTurk and representative samples, which were similar to the target population, and did not have access to or failed to assimilate valid outside information. It is interesting to note that for these two groups, since the validity of outside information was lower than for the comparable groups in Hoch (1988), their optimal weight for their own position were higher than in Hoch. Had these two groups projected at an optimal level, they could have more than doubled their predictive accuracy. Surprisingly, even our undergraduate sample, which had lower levels of similarity to the target group, could have benefited by projecting more. Our data show that their predictive accuracy could have improved by more than 60% had they projected at the optimal level.
General Interest Items. The optimal weight for own positions for these items was significantly lower than for the AIO items, probably because valid outside information was more available for these items. Since participants projected more for these items, the difference between the optimal and the observed weights for own position were smaller for all three groups than for the AIO items. Interestingly for these items, for all three groups, there was also a larger percent of people who over-projected. The percent of people who over-projected was greater for undergraduate students than for the other two samples. Finally, the maximum possible accuracy for these items with optimal weighting was significantly higher than for the AIO items. This confirms our conjecture that one reason why people performed poorly in predicting AIO items in Hoch’s (1988) samples and ours was that these were difficult tasks.

DISCUSSION

Summary of Findings and Implications

Over 25 years ago Hoch (1988) examined how accurately different groups of people could predict the attitudes, interests, and opinions (AIO) of average Americans. He found that people did not project their own views enough when making predictions about others and concluded that when valid sources of outside information are lacking, the only helpful information people do have is their own position on the issue and they should project their own views more than they do.

The main goal of this paper was to examine whether consumers’ ability to accurately predict what others are like has changed in any substantive way over this long time period. Many
things have changed in the world in the last quarter century including a massive change in the data available to consumers and firms about others through the internet, specialized and targeted media sources such as Fox News and CNBC, social media like Facebook and Twitter, crowd sourcing, and user generated content. We explored whether things have changed in an open minded fashion since we thought three different patterns were all plausible. First we thought that since projection is a basic psychological process, we might observe very similar patterns to Hoch’s. Second, we thought that given the massive increase in data that are available to consumers now about others, they would have more valid outside information to rely on in making intuitive predictions and that could lead their predictive accuracy to increase. Third, we thought that although there is more information available now than in the late 1980’s, it is not necessarily more valid information. If true, reliance on this information could lead to less accurate forecasts.

We found support for the third pattern. Both our representative sample and the mTurk sample were less accurate than Hoch’s representative and convenience samples, respectively. Interestingly our undergraduate sample, while not significantly different from our other two samples, was less accurate than Hoch’s representative sample, but did not differ from his convenience sample.

Why were our representative and mTurk samples less accurate than Hoch’s samples? It was not because of similarity; as they were more similar to the prediction target than were Hoch’s samples. It was also not because of large differences in how much they projected their own position. Our representative sample did not differ from Hoch’s in the degree of projection. The mTurk sample did project significantly less than Hoch’s convenience sample, but not to a large degree. However, all three of our samples did differ from Hoch’s in the validity of other
information. For our representative sample and the mTurk sample, their validity of other information was lower than Hoch’s and this seems to be the main factor leading to their lower predictive accuracy. In contrast, the undergraduate sample had higher validity of other information than any of the other samples (ours and Hochs). This likely led to their comparable predictive accuracy to our other two groups, despite their lower similarity to the target.

Our analysis of the optimal weights showed that, as Hoch (1988) had found, all three of our groups could have increased their predictive accuracy if they had projected their own position to a greater degree. Thus like he found, false consensus does not seem to be a problem for this task since very few people in any of the groups projected too much and rather the vast majority under-projected. If anything, people seem to be worse at predicting the AIO of average Americans than they were in the late 1980’s, and this seems to be because they are underweighting their own position to a greater extent than they did previously. It also seems to be due to the lower validity of information available (at least for the representative and mTurk samples). We cannot confirm from the data available to us, but it could be that consumers today feel that they know more about other consumers because of the information available, they do not realize that information is anecdotal and not representative of the average American, and because of that rely less on their own position than they did in the past.

In addition to examining whether Hoch’s findings still held over 25 years later, our research expanded on Hoch’s work in two ways. First, we explored whether the low predictive validity for the AIO items may have been due not just to people under-projecting, but also because the AIO questions used both by Hoch and by us largely asked about private thoughts people hold that are typically not covered in the media. For these types of questions it would therefore be reasonable to assume that individuals making prediction do not have access to valid
outside information on which to make a prediction. We therefore asked the participants in our three groups to also predict how they think average Americans would answer general interest questions, where more outside information should be available to those making predictions, to examine both how this improves predictive accuracy and how it changes the degree to which people project their own position.

We did indeed find higher validity of outside information for all three groups for the general interest questions compared to the AIO questions. This led to a dramatic increase in predictive validity for all three groups. We also found that for these questions and for all three groups that people were projecting at closer to the optimal level and that while the majority of people were still under-projecting, there was now a sizeable group of people in each sample who were over-projecting their own position.

The second way in which we expanded on Hoch’s (1988) work was by including two new samples that are often relied on by consumer behavior researchers – undergraduate students and research participants on Mechanical Turk. An important issue for some research questions is how generalizable the views of participants in these groups are to the larger population of consumers. For the AIO questions we found that the mTurk sample had lower similarity than the representative sample, but not by much, which is good news for researchers who conduct mTurk studies. In contrast undergraduate students had much lower similarity. The news was better when we looked at the general interest questions. There all three groups had higher levels of similarity than for the AIO questions, and for these too there was no difference in similarity for the mTurk and the representative samples. The student sample still had lower similarity.

A second question is whether researchers and firms can rely on these samples to predict what a larger more representative group would say. In other words, rather than going to a large
representative sample, could firms or researchers use mTurk or student samples as almost a convenience sample and ask them to predict how others would respond to their research questions? For this, the news is excellent. There is no significant loss to predictive accuracy if one relies on an mTurk or a student sample compared to a representative sample, either for AIO questions where respondents attempt to predict responses of a more private nature, or for general interest questions where respondents also have valid outside information to help in developing their intuitive predictions.

These results suggest that if survey participants complete predictive tasks, such as Frito-Lay’s “Do Us a Flavor” competition, a convenience sample is as generalizable as a representative sample. However, because the two convenience samples have lower similarity to the representative sample, we conclude that having a representative sample is important if it is going to be used for extrapolating the sample’s measured AIO to the general population.

Limitations and Opportunities for Future Research

One limitation of our approach is that we do not have the data to determine why everyday consumers are less accurate at predicting the AIO of average Americans than they were over 25 years ago. It could be, as we speculated, that the information people have about others through the media is biased. It could also be that there is now too much media information which overwhelms people and makes it more difficult for them to use it in constructing intuitive predictions. Finally it could be that the diversity of the average American target has increased in the last quarter century and because of that people have a more difficult time bringing to mind a consistent target stereotype or deciding what outside information is appropriate to rely on.
Consistent with this viewpoint, Hoch (1987) found higher levels of the validity of outside information when people made predictions about a known group, peers or their spouses, than when they made predictions about other consumers. Different methods than what Hoch (1987, 1988) and we employed would be required, but future research should explore these possibilities.

Our method also does not allow us to disentangle different aspects of the validity of outside information. When we observe low validity of outside information, we cannot differentiate between whether those making predictions simply lack access to outside information, whether they lack access to high quality, valid information, or whether they have access to valid information but do not employ it when making intuitive predictions. Future research using other methods should explore these possibilities and more fully examine the role of various forms of outside information (e.g., news reports, information obtained from social media, information obtained from user generated content) in intuitive predictions.

In the current research we compared the mTurk sample and the undergraduate student sample to a representative sample of American consumers. Hoch’s representative sample members were people in Market Facts’ consumer panel who received the questionnaire by mail. Our representative sample was its modern day counterpart; people in a marketing research firms’ proprietary online panel. The online panel that we used is designed to be representative of all American consumers. Though some research questions the validity of internet panels (Couper 2000), others find them to offer more accurate results than other traditional samples such as probability telephone samples (Chang and Krosnick 2009). Though the internet sample we employed matches the demographics of the greater population of Americans, still we cannot be certain that it represented the attitudes and opinions of Americans.

Related to this, Hoch’s “truth” sample came from the Market Fact panel as did his
representative sample. In our case we also relied on the same propriety market research internet panel for both our representative and our truth groups. In contrast to Hoch, rather than use two independent samples, we used a bootstrapping method to assess the accuracy of predictions in order to make best use of our data. However, since the predictions and truth came from the same group of participants, this may have artificially increased the level of similarity for the representative sample. We note though, that even though the similarity was higher for our representative sample than was Hoch’s, their predictive accuracy was lower. Still is it possible that predictive accuracy would have been even lower had we employed a completely independent representative sample.

Hoch (1988) compared the predictive accuracy of experts to those of novices – every day consumers. We focused just on the latter group, every day consumers, but future research should examine whether the intuitive predictions of marketing experts has changed in the last quarter century. Like every day consumers, they have at their fingertips access to much more data about their customers, but to the extent to which those data are not valid, it is possible that the predictive accuracy of marketing experts, like for every day consumers, has declined in the last quarter century.

We compared predictive accuracy for two different sets of questions; the AIO questions used in Hoch (1988) and a set of new general interest questions. We believe the latter set of questions differ from the former in that more outside information is available about them. However these two sets of questions differ in many different ways. Future research could develop a typology of different types of questions that vary in how easy or hard they are for people to accurately predict others’ answers to study factors affecting predictive accuracy. Predictive accuracy may for example depend on what is observable about consumers, what is
reported in the news and in surveys in the popular press, and what types of topics people tend to talk about on social media.

Finally future research should examine how well people can make other predictions relative to consumers and markets. For example, future research could see to what extent everyday consumers and marketing experts can predict facts about the market place (e.g., the market share of different brands) or theories about markets (e.g., does advertising change behavior?).

Over 20 years ago Armstrong (1991) asked academics, practitioners, and high school students to predict whether hypotheses from published JCR articles were supported or not. He found that no one group was significantly better than others and that none preformed above chance rates. He also found that experts were no better than novices for these tasks. Given the increased reporting of social science findings in the popular press, an interesting question to explore would be whether predictive accuracy for hypotheses or findings in consumer or market-relevant academic journals has changed over time.
REFERENCES


Mason, Winter, and Siddharty Suri (2012), “Conducting behavioral research on Amazon’s


577-592.


Table 1

Actual Similarity, Level of Projection, Validity of Other Information, and Predictive Accuracy for Each Population

<table>
<thead>
<tr>
<th>Respondent population</th>
<th>Actual similarity $r(t,o)$</th>
<th>Level of projection $r(p,o)$</th>
<th>Validity of other information $r(t,z)$</th>
<th>Predictive accuracy $r(t,p)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hoch Samples – Hoch AIO items</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Convenience</td>
<td>.222</td>
<td>.362</td>
<td>.098</td>
<td>.184</td>
</tr>
<tr>
<td>Representative</td>
<td>.311&lt;sup&gt;b&lt;/sup&gt;</td>
<td>.302</td>
<td>.107</td>
<td>.202</td>
</tr>
<tr>
<td>New samples – Hoch Items</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Undergrads</td>
<td>.172 (.137,.214)</td>
<td>.054 (.009,.100)</td>
<td>.151 (.114,.189)</td>
<td>.163 (.129,.198)</td>
</tr>
<tr>
<td>Mturk (Convenience)</td>
<td>.322 (.296,.351)</td>
<td>.258 (.239,.298)</td>
<td>.035 (.007,.063)</td>
<td>.115 (.090,.141)</td>
</tr>
<tr>
<td>Representative</td>
<td>.413 (.380,.443)</td>
<td>.281 (.241,.320)</td>
<td>.032 (.002,.062)</td>
<td>.153 (.123,.181)</td>
</tr>
<tr>
<td>New samples – General Interest Items</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Undergrads</td>
<td>0.282 (.241,.322)</td>
<td>0.385 (.336,.435)</td>
<td>.569 (.533,.605)</td>
<td>0.567 (.542,.591)</td>
</tr>
<tr>
<td>Mturk (Convenience)</td>
<td>0.362 (.327,.395)</td>
<td>0.351 (.318,.384)</td>
<td>.484 (.455,.512)</td>
<td>0.530 (.508,.551)</td>
</tr>
<tr>
<td>Representative</td>
<td>0.405 (.367,.437)</td>
<td>0.348 (.307,.383)</td>
<td>.479 (.447,.511)</td>
<td>0.530 (.505,.560)</td>
</tr>
</tbody>
</table>

Note: The combined means represent unweighted means. The 95% confidence interval for each variable is shown in parentheses under each of the means.
Table 2
Optimal Weighting Compared to Empirical Weighting of Own Position Averaged Across
the Separate Male and Female Consumer Targets

<table>
<thead>
<tr>
<th>Respondent population</th>
<th>Empirical Weight $b = r(p,o)$</th>
<th>Optimal Weight $b'$</th>
<th>Difference ($b' - b$)</th>
<th>Percent $b' &lt; b$ projecting too much</th>
<th>Maximum accuracy $r(t,p')$</th>
<th>Improvement due to optimal weighting</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hoch Samples – Hoch AIO items</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Convenience</td>
<td>.362</td>
<td>.525</td>
<td>+.163</td>
<td>26%</td>
<td>.353</td>
<td>.169</td>
</tr>
<tr>
<td>Representative</td>
<td>.302</td>
<td>.685</td>
<td>+.383</td>
<td>16%</td>
<td>.426</td>
<td>.224</td>
</tr>
<tr>
<td><strong>Our samples – Hoch Questions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Undergrads</td>
<td>.054 (.009,.100)</td>
<td>.606 (.548,.665)</td>
<td>+.552</td>
<td>7%</td>
<td>.262 (.221,.304)</td>
<td>.099</td>
</tr>
<tr>
<td>Mturk (Convenience)</td>
<td>.258 (.239,.298)</td>
<td>.810 (.778,.842)</td>
<td>+.552</td>
<td>10%</td>
<td>.318 (.287,.318)</td>
<td>.203</td>
</tr>
<tr>
<td>Representative</td>
<td>.281 (.241,.320)</td>
<td>.860 (.830,.890)</td>
<td>+.580</td>
<td>5%</td>
<td>.403 (.367,.439)</td>
<td>.253</td>
</tr>
<tr>
<td><strong>Our samples – General Interest Questions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Undergrads</td>
<td>0.385 (.336,.435)</td>
<td>.453 (.406,.500)</td>
<td>+.073</td>
<td>41%</td>
<td>.685 (.637,.704)</td>
<td>.120</td>
</tr>
<tr>
<td>Mturk (Convenience)</td>
<td>0.351 (.318,.384)</td>
<td>.554 (.511,.597)</td>
<td>+.203</td>
<td>24%</td>
<td>.661 (.637,.686)</td>
<td>.131</td>
</tr>
<tr>
<td>Representative</td>
<td>0.348 (.307,.383)</td>
<td>.596 (.554,.638)</td>
<td>+.248</td>
<td>22%</td>
<td>.671 (.643,.698)</td>
<td>.141</td>
</tr>
</tbody>
</table>
Appendix A

Attitude, Interest and Opinion Questions

1. I would like to spend a year in London or Paris.
2. If I had my life to live over, I would sure do things differently.
3. I am an impulse buyer.
4. I am a homebody.
5. A nationally advertised brand is usually a better buy than a generic brand.
6. I would rather spend a quiet evening at home than go out to a party.
7. Information from advertising helps me make better buying decisions
8. I like to pay cash for everything I buy.
9. I am more concerned about nutrition than most of my friends are.
10. Television is my primary form of entertainment.
11. Our family is too heavily in debt today.
12. Police should use whatever force is necessary to maintain law and order.
13. Terrorism is the greatest peril in the world today.
14. A woman’s place is in the home
15. I am concerned about how much sugar I eat.
16. I would rather live in or near a big city than in or near a small town.
17. Children cannot get a good education in schools today.
18. The government should exercise more control over what is shown on television.
19. The dirt you can't see is worse than the dirt you can see.
20. I have somewhat old-fashioned tastes and habits.
21. There is too much talk these days about what is good and bad for you when it comes to food
Appendix B

General Interest Questions

1. What percent of Americans have a college degree?
2. What percent of Americans attended church or synagogue or some other religious service in the last seven days?
3. What percent of American households owns a gun?
4. What percent of Americans are vegetarians?
5. What percent of Americans are gay, lesbian, or transgender?
6. What percent of Americans have a passport?
7. What percent of Americans eat 2 or more servings of fruit per day?
8. What percent of Americans played golf at least once in the past year?
9. What percent of Americans drink coffee on a daily basis?
10. What percent of Americans own a Smartphone?
11. What percent of Americans do not have any health insurance?
12. What percent of Americans live below the poverty line?
13. What percent of Americans earn at least $1 million per year?
14. What percent of Americans have net assets of $100,000 or more?