

**Language of Administration as a Source of Measurement Error:
Implications for Surveys of Immigrants and Cross-cultural Survey
Research**

by

Emilia A. Peytcheva

**A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Survey Methodology)
in The University of Michigan
2008**

Doctoral Committee:

**Professor Norbert W. Schwarz, Chair
Professor Robert M. Groves
Professor Trivellore E. Raghunathan
Research Associate Professor Frederick G. Conrad**

Table of Contents

List of Figures	iv
List of Tables	v
Chapter I. Introduction	1
Chapter II. Effect of Language of Survey Administration on the Response Formation Processes.....	7
1. Introduction.....	7
2. Comprehension	10
2.1 <i>Cultural Frame Switching</i>	12
2.2 <i>Spatial Frames of Reference</i>	14
2.3 <i>Codability</i>	16
3. Retrieval and Judgment.....	18
3.1 <i>Behavioral Reports</i>	18
3.2 <i>Attitude Questions</i>	25
4. Response Formatting	31
4.1 <i>Codability</i>	31
4.2 <i>Cultural Frame Switching</i>	32
5. Response Editing	34
5.1 <i>Cultural Frame Switching</i>	34
6. Summary	35
Chapter III. Effect of Language of Survey Administration on Responses to Survey Questions	39
1. Introduction.....	39
1.1 <i>Language, Cognition and Culture</i>	39
1.2 <i>Language in Surveys of Bilingual Respondents</i>	43
2. Data and Methods	49
2.1 <i>Dataset</i>	49
2.2 <i>Causal Inference in the Absence of Random Assignment</i>	51
3. Results	54
3.1 <i>Propensity Score Model</i>	54
3.2 <i>Hypotheses Tests</i>	59
4. Discussion and Conclusions.....	66
5. Appendix	70
Chapter IV. Does Language Matter: An Investigation of the Effect of Language Assignment on Answering Behavior among Hispanic Bilingual Survey Respondents	74
1. Introduction.....	74
2. Data and Methods	79
3. Results	83

3.1	<i>Hypotheses Tests</i>	83
3.2	<i>As-treated Analyses and Propensity Model</i>	88
3.3	<i>Hypotheses Tests Revisited</i>	93
4.	Discussion and Conclusions.....	96
5.	Appendix	100
	Chapter V. Conclusions.....	108
1.	What have we learned?.....	108
2.	Implications for survey practitioners	114
3.	Next Steps.....	116
	References	119

List of Figures

Figure 1. Effect of Language on Survey Response Formation Processes	10
Figure 2. Overlap of Estimated Propensity Scores for Those Interviewed in English and Those Interviewed in Spanish.....	58
Figure 3. Outcome Measures of Interest from the NIS Questionnaire	73
Figure 4. Outcome Measures of Interest from the NLAAS Questionnaire	106
Figure 5. Overlap of Estimated Propensity Scores for Those Interviewed in English and Those Interviewed in Spanish.....	107

List of Tables

Table 1. Logistic Regression Coefficients for Likelihood of Selecting English Language for the Survey Interview	57
Table 2. Number of Cases in Propensity Strata by Language of Survey Administration.....	59
Table 3. Relative Risk for the English Language Group versus Spanish across Measures by Expectations of Language Effect	60
Table 4. Distribution of Outcome Variables by Language of Administration by Propensity Strata.....	62
Table 5. Coefficients for Logistic Regression Models for Being Diagnosed with Emotional and Health Problems, Being Depressed and Alcohol Consumption ..	65
Table 6. Distributions of Correlates of Language Selection and Demographic Characteristics by Propensity Strata.....	70
Table 7. Relative Risk for the English Language Assignment Group across All Bilinguals and Immigrants	85
Table 8. Means for the English and Spanish Language Assignment Groups across All Bilinguals and Immigrants	86
Table 9. Logistic Regression Coefficients for Likelihood of Being in the English Language Compliant Group	92
Table 10. Relative Risk for the English Language Compliant Group versus Spanish across Measures with Different Language Effect Expectations Reflecting Stratification by Propensity Strata.....	95
Table 11. Means for the English and Spanish Language Compliant Groups across Measures with Different Language Effect Expectations Reflecting Stratification by Propensity Strata.....	95
Table 12. Relative Risk for the English Language Group versus Spanish of Immigrants who Came to the U.S. at Age 13 or Older across Measures where Language Effects are Expected (n=76)	100
Table 13. Means for Respondents who Came to the U.S. at Age 13 or Older across Measures where Language Effects are Expected (n=76)	100
Table 14. Demographic Characteristics and Correlates of Acculturation by Compliance Groups.....	101
Table 15. Language of Survey Administration by Propensity Strata.....	102
Table 16. Balancing of Propensity Model Predictors across Propensity Strata	103

Chapter I. Introduction

In the age of globalization, modern societies are characterized by an increasing number of people with multiple cultural backgrounds. Often, such individuals speak more than one language and are knowledgeable about more than one culture. U.S. surveys of minorities and immigrants usually allow respondents to answer in the language of their choice (e.g., the New Immigrant Survey; the National Survey of Latinos; the National Latino and Asian American Study). Those who choose to answer in their mother-tongue are likely to differ from those who choose to answer in English in their background characteristics (e.g., level of acculturation, education), substantive answers and response patterns (like item missing data or “Don’t Know” responses, see Collet, 2007). While self-selection certainly plays a role in these differences, it would be premature to consider it the sole source of all observed differences between the English and the foreign language versions of the same survey. One obvious source of measurement error is the necessary use of different languages when intending to measure the same phenomena in multiple ethno-cultural groups.

Indeed, recent psychological and linguistic research suggests that the language of administration may influence responses over and above respondents’ self-selection into language conditions. Most importantly,

different languages are spoken in different cultural contexts and are inevitably associated with the respective cultural meaning system. Hence, the language of the interview can serve as a powerful situational cue that primes the associated cultural meaning system in bicultural bilingual respondents.

One likely result is that the language of survey administration affects the respondent's reference frame, potentially influencing how respondents perceive the intent of the questions asked and their affective characteristics, including question sensitivity and need for a socially acceptable answer. These influences may be further exacerbated when the interviewer is a native speaker of the respective language, who can be assumed to share the cultural norms of the respondent. If so, bicultural bilingual individuals may answer differently, based on the cultural frame that is evoked by the language of administration.

Furthermore, language can influence the way we think. Since Sapir (1929) and Whorf (1940) addressed the issue of linguistic relativity and determinism, many studies have supplied evidence that language can influence cognition through its syntactic properties (e.g., Briere and Lanktree, 1983; Khosroshashi, 1989; Ng, 1990), grammatical structure (e.g., Bickel, 1997; Widlock, 1997; Majid et al., 2004) and semantic categories (e.g., Hunt and Agnoli, 1991). Most importantly, several studies have demonstrated that the language in which mental activity is carried out during information encoding can serve as a retrieval cue during information recall (e.g., Schrauf and Rubin, 1998; Marian and Neisser, 2000; Schrauf and Rubin, 2000).

Accordingly, the language of survey administration may facilitate the recall of information from time periods and life domains in which the same language was spoken, but may impede recall from domains and periods when a different language was spoken.

For survey practitioners this would suggest that depending on the language of survey administration, bilingual bicultural respondents may answer the same questions differently. This can be a result of different cultural frameworks activated by the different languages, the fact that the requested information was initially encoded in a different language, or other language-related mechanisms. However, survey questions vary in their susceptibility to measurement error – for example, response accuracy (and item nonresponse) in sensitive questions may depend on perceived risk of disclosure (Couper, Singer and Kulka, 1998; Singer, Mathiowetz and Couper, 1993) and features of the survey interview associated with privacy (Tourangeau et al., 1997; Turner et al., 1998). Thus, it may be expected that language, as an external stimulus, may not perform similarly across survey questions. The limited available research is consistent with this assumption – for example, a study of bilinguals attending an American school in Greece demonstrated that the lowest correlation between responses provided by the same respondent in the two languages of interview occurred when the questions differed in social desirability in Greece and the United States (Triandis, Davis, Vassiliou and Nassiakou, 1965).

As this discussion suggests, the language of survey administration has the potential to influence every stage of the response formation process (see Sudman, Bradburn and Schwarz, 1996; Tourangeau, Rips and Rasinski, 2000), from the encoding of events that precede the survey, to the interpretation of the question, the recall of information from memory, the norms and expectations that figure prominently in the computation of behavioral frequency estimates, and the social desirability considerations that determine response editing. At present, we know from carefully controlled laboratory studies that language *can* have such influences on tasks that are designed to assess them. However, the psychological research provides little insight into whether these influences are of practical relevance to survey research; conversely, the findings in the survey methodology literature fail to isolate the influence of language *per se*. This dissertation attempts to fill in the existing gap by presenting a theoretical framework for language effects and exploring language influences through secondary analyses of observational and experimental data.

The dissertation consists of three related essays. The first one presents a theoretical framework for the possible mechanisms through which language affects each stage of the response formation process. Research from psycholinguistics and cross-cultural psychology is presented and the most studied mechanisms associated with language effects are linked to stages of the response formation model. The discussion focuses on question comprehension, retrieval and judgment, response formatting and editing and

the possible language effects related to them based on published studies. After each section, speculations on how the presented findings may impact survey data are offered.

The second essay attempts empirical estimation of language effects in a national survey of immigrants. The analyses focus on Spanish-English bilinguals who self-select the language of survey interview. We attempt to eliminate the confounding effects of this self-selection through the use of propensity score methods. Two groups of questions are considered – question in which language effects are expected to be the strongest because of differential social desirability associated with the Hispanic and American cultures, and questions related to highly accessible and well defined facts, where no language influences are anticipated. The first group involves questions on mental, physical health and alcohol use – these topics are associated with different stigma in the two cultures. In contrast, the second group consists of questions related to respondent’s marital status, living arrangements or number of biological children.

Despite its promise, the analytic approach used in the second essay is sensitive to propensity model specification and the omission of unobserved covariates. In order to build confidence in the causal conclusions, replication and alternative approaches are needed. The third essay presents analyses of an experimental assignment of bilingual respondents into the language of survey administration. It compares questions on similar topics – self-evaluation of mental and physical health, age of first alcohol use, and

familism – areas where the cultural values and norms for Hispanics and Americans diverge. Again, as a contrast, questions where response differences between the two language groups are not expected are examined.

The results bear on surveys of immigrant populations as well as on cross-cultural surveys administered in multiple languages. By going beyond the current focus on translation issues, the dissertation illuminates how the language of survey administration affects respondents' answers and highlights the methodological implications for surveys. Significant language effects suggest that the present practice of leaving the choice of language to the bilingual respondent may not be a good one. Further, the dissertation motivates the next stage of research – disentangling the mechanisms that can possibly induce language effects at various stages of the response formation process.

The broader impact of this work derives from the relevance of collecting accurate data from immigrants and minorities. If different languages evoke different cultural frames in bilingual biculturals, and if language affects the accuracy of autobiographical recall, the choice of language may be an important factor in collecting accurate data for policy decisions. At present, little is known about these possibilities.

Chapter II. Effect of Language of Survey Administration on the Response Formation Processes

1. Introduction

The link between language and cognition has long fascinated scientists. Aristotle's view of language as a tool to communicate thoughts evolved to theories suggesting that thinking is silent, subvocal speech (Watson, 1925), and then to theories stating that language determines the way we think (Einstein, 1954; Whorf, 1956 and Allport, 1954). In the 1950s, the idea that we can think only in terms of concepts that are represented in our language received most attention. Its best-known proponent, Benjamin Whorf, made popular the principles of linguistic determinism and linguistic relativity, which constitute what is known as the Sapir-Whorf hypothesis (Whorf, 1956). The principle of linguistic relativity states that cultural differences in cognition are correlated with differences in languages across cultures; the principle of linguistic determinism states that people think differently *because* of differences in their languages. Most of the evidence for this hypothesis was based on cross-cultural comparisons; however, the analyses of language samples were not always methodologically objective (see Longacre, 1956). Nowadays, there is a revived interest in a much weaker version of linguistic

determinism - that language and thought are interdependent, but their relationship is not yet fully understood.

Surveys are often administered in more than one language (e.g., the National Survey of Family Growth, the National Comorbidity Survey, the National Survey of American Life). To the extent to which language influences thought processes, responses provided by bilingual respondents would depend on the language of survey administration. In order to study such language influences, we need to assume that the different language versions are free of translational problems and convey the same concepts and ideas. Thus, any observed differences between responses provided by the same respondent in different languages can be attributed to language priming a particular mind frame and influencing the thought processes.

To examine the potential effects of language on survey responding we focus on the response formation process (Sudman et al., 1996; Tourangeau et al., 2000). The right hand-side of Figure 1 presents the tasks that respondents perform in order to answer a survey question. These tasks are not necessarily sequential or independent, but are presented as such for simplicity. First, a respondent has to interpret the question and the response alternatives. Understanding the literal and pragmatic meaning occurs at the comprehension stage. Next, the relevant information has to be recalled from memory (retrieval stage), then evaluated for relevancy and summarized at the judgment stage. At the formatting stage, the response that is ready to be reported is configured to match the presented response

options. Still, before providing an answer, respondents may edit their response (editing stage) due to considerations related to self-presentation and social desirability. Various factors may affect each task, language of survey administration possibly being one of them. The left hand-side of Figure 1 represents the mechanisms related to language that are most likely to affect each stage of the response formation process. Cultural frame switching, language dependent recall, codability and language inherent frames of reference have been the focus of investigation of many psycholinguistic and cross-cultural psychology studies; however, their importance for survey research remains unknown. Next, we examine each respondent task and the possible language influences by reviewing the existing literature from relevant fields and deriving hypotheses about applications to surveys.

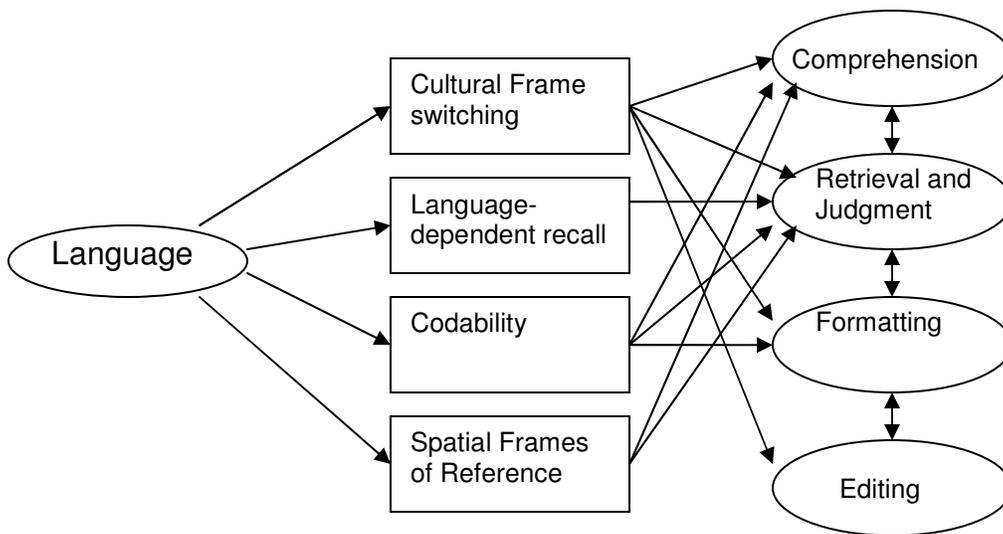


Figure 1. Effect of Language on Survey Response Formation Processes

2. Comprehension

Survey results are meaningful to the extent to which respondents understand the survey questions as they are originally intended. Question comprehension involves processing the syntactic structure and understanding the semantic (literal) and pragmatic (intended) meaning.

While constructing simple sentences that contain clear and familiar words is important (see Converse and Presser, 1986; Fowler, 1995), it is also crucial to keep in mind that respondents infer the intended meaning of a question based on contextual information, such as response options, preceding questions, survey topic, sponsor, etc. (for a review see Schwarz, 1996). In cross-cultural surveys, in addition to the direct impact of

translation, comprehension problems may occur as a result of differences related to cognition. Sentence comprehension is a constructive process that involves general knowledge about the world, not just knowledge specific to language processing (Bransford and Franks 1971; Bransford, Barclay and Franks, 1972). Since language is a tool for information exchange among people of the same culture, it reflects the meaning system of the culture. Thus, word meaning/sentence meaning in language comprehension depends on preexisting background knowledge about not only the grammatical norms associated with the language, but also the cultural norms and practices related to it. For example, the question “Could you close the window?” is not a request for a report on one’s physical capability, but a request for action reflecting a norm of politeness in cultures that use the English language.

Furthermore, lexical ambiguity is inherent to languages and recall of the lexical meaning of words is often context dependent. Languages differ in their context dependency and this is reflected in conversational norms across cultures. For example, many words in Chinese acquire meaning only in the conversational context and can not be translated directly. This is related to the fact that East Asian cultures value ‘reading between the lines’ (for an overview see Nisbett, 2003). Thus, the same question read in Chinese or English by a bilingual bicultural respondent may convey a different meaning depending on how much contextual information is included from previous questions.

2.1 *Cultural Frame Switching*

Psychological research suggests that the use of contextual information differs across cultures – for example, East Asians have been found to be more sensitive to the conversational context than Westerners (for a review see Schwarz, 2003). Such differential context dependency can have consequences for question interpretation when partially redundant information is presented. For example, Haberstroh et al. (2002) asked Chinese and German students to report their academic and general life satisfaction using the academic-life or life-academic question order. As expected, in contrast to German students who used information brought to mind by the academic satisfaction question to answer the general life satisfaction question, Chinese students, being more sensitive to the conversational context and detecting the potentially redundant questions, disregarded the information they already provided about their academic life when reporting general life satisfaction.

More interestingly, however, in bicultural respondents, such context sensitivity is likely to be dependent on which cultural frame is primed by the survey question. Research on acculturation has demonstrated that individuals can possess more than one cultural identity (for example, Berry and Sam, 1996) and move between different cultural meaning systems, depending on situational cues and requirements. This phenomenon is known as “cultural frame switching.” Each cultural meaning system serves as an interpretive frame that affects individual’s cognition, emotion and

behavior (Geertz, 1993; Hong, Chiu and Kung, 1997; Kashima, 2000; Mendoza-Denton, Shoda, Ayduk and Mischel, 1999).

Language can serve as a situational cue for the cultural system associated with it; thus, it may prompt bicultural bilingual respondents to differential question interpretation based on the cultural frame induced by it. For example, Ross, Xun and Wilson (2002) used language to shift the cultural frame adopted by Chinese-born students in Canada in a task of self and mood description. As expected, those asked in Chinese provided collective statements and reported higher endorsement of Chinese views. These respondents also reported equal levels of positive and negative moods, whereas participants who were asked in English, reported a preponderance of positive moods. Similarly, Trafimow, Silverman, Fan and Law (1997) found that bilingual Hong Kong students reported more private traits and less social roles when interviewed in English versus Chinese, consistent with the associated cultural emphasis on individual vs. collective aspects of identity.

Such research suggests that language can be a powerful cue for the interpretive frame bilingual bicultural respondents adopt when answering survey questions. Different conversational dynamics across cultures can change the pragmatics of the survey interview, evoking different Gricean implicatures¹ (Grice, 1989). Thus, depending on the cultural identity primed by the language of survey administration and the socially anchored

¹ Inference drawn from a sentence that preserves the implicit agreement among participants to make meaningful contributions to the conversation (*cooperative principle*, Grice, 1989)

behaviors and values associated with it, respondents may assume that certain features of the survey question and/or the preceding questions are relevant for their response. For example, for Chinese-English bilingual biculturals, language may determine how much contextual information is used when processing a survey question, which in turn would determine its pragmatic meaning. Furthermore, which cultural identity is primed may have consequences for the amount of other background information (e.g., survey sponsor, visual images) that is considered during response formation. Thus, cultural differences in conversational dynamics, such as context sensitivity, should be taken into account in the design of survey questionnaires intended for immigrants and ethnic minorities.

2.2 Spatial Frames of Reference

Languages have inherent frames of reference for describing relationships among objects. Recent research has demonstrated that languages differ in the availability and range of application of these frames of reference (for example, Levinson, 1996). There is a common distinction in psychology, psycholinguistics and brain sciences between relative and absolute languages (also referred as egocentric and allocentric). Relative languages use a viewer-centered perspective, giving rise to descriptions such as “in front of me”, “to the left”. Most Western languages are relative. Absolute languages use external reference frames, such as cardinal directions or up-down axis – for example, speakers of Arrernte (Australia)

will say “the fork is to the north of the spoon”, while speakers of Tzeltal (Mexico) will describe the fork “uphill of the spoon” (Majid et al., 2004).

Such intrinsic language differences may potentially affect comprehension in bilingual speakers of languages with different dominant spatial frames of reference as the latter have been found to determine many aspects of cognition (see Levinson, 2003). Experiments by Pederson et al. (1998) demonstrate that what linguistic frames of reference are dominant in a language reliably correlates with the way its users conceptualize in nonlinguistic domains. For example, speakers of Mopan (Mayan) and Kilivila (Austronesian) can not distinguish between two photographs of a man facing a tree when the position of the man and the tree are left-right reflections of one another as such relationship between the objects is described as “tree at man’s chest” in both photographs.

For survey practitioners, such findings suggest that speakers of languages that use different frames of reference may interpret survey visual images and response scales differently. For example, the orientation of a scale (vertical or horizontal) may influence how similar or distinct response categories are perceived, depending on the language used and its inherent frame of reference. However, such effects are likely to occur only in cases where the dominant frames of reference used in two languages are not functional equivalents of one another (as in the example of Mopan speakers where there were no functional equivalents of “left” and “right” in the

described mirror-image photographs), thus their impact on the survey response processes may be very limited.

However, another aspect of language frames of reference may bear more directly on surveys of bilingual bicultural respondents. The relationship between dominant frames of reference and cultural orientation (individualistic vs. collectivistic²) remains unknown. To the extent to which ego-centered frames of reference are related to individualistic identities across cultures that use such languages and vice versa, the language of administration will be an important factor influencing survey responses. Similar to cultural frame switching, a speaker of languages that use different frames of reference would endorse more individualistic or collectivistic responses depending on the cultural identity evoked by the ego-centric or allocentric frame of reference inherent to the language of survey administration. Such possibility deserves further investigation.

2.3 Codability

Language codability is the ease with which a language can express a concept. Not surprisingly, the most highly codable concepts are presented by the most frequently used words, which are short, easy to write and pronounce (see Whitney, 1998). Codability has been demonstrated to affect mostly cognitive processes such as retrieval (Lucy, 1992, Lucy and Shweder, 1979, Lucy and Wertsch, 1987) and comparative judgment (Kay

² For reviews on documented social and cognitive differences see Oyserman, Coon and Kimmelmeier, 2002 and Oyserman and Lee, 2007

and Kempton, 1984). For example, Kay and Kempton (1984) demonstrated that speakers of a Mexican-Indian language that did not have separate words for blue and green were unable to differentiate among color chips on the blue-green continuum based on color characteristics (see more details in section 3.2.2).

However, codability may also influence question comprehension in surveys. When respondents read a question, they form a representation of the question target. The question target may be very different depending on whether a concrete word exists in the language for a given attitude object or behavior, or only a general word exists. For example, in Chinese, there are separate terms for family members that have only one English equivalent - different words describe whether your “uncle” is your mother’s brother, or father’s brother, and whether he is a younger or older brother. Analogous to the study by Kay and Kempton (1984), it can be hypothesized that when asked in Chinese about two or more related people that can be labeled differently, respondents may think of them differently relative to when asked in English, when a common label is used. This may lead to inclusion errors when respondents are asked in English³ due to failure to draw a lexical distinction across referents. Such interpretational differences across two languages may affect various respondent tasks – for example, household roster construction.

³ An additional complication in this example comes from the fact that “uncle” is often used for male elderly acquaintances who do not necessarily have a family relation to the respondent.

Overall, language may influence the literal and pragmatic meaning of survey questions and hence the mental representation of the question target. Different mechanisms may induce such effects and their magnitudes and importance for surveys in multiple languages await further investigation.

3. Retrieval and Judgment

Once respondents comprehend the question and determine what information is required, they need to recall the relevant information from memory. Rarely, the requested information is readily available (for example, a previously formed opinion; facts such as respondent's gender or birthplace). Most often, respondents need to compute a judgment on the spot. This process is somewhat different for behaviors and attitudes.

3.1 Behavioral Reports

Behavioral questions often ask about past events that took place in respondent's life. When such events have low frequency of occurrence or are of particular importance to the respondent, they may be directly accessible in memory (for reviews see Bradburn, Rips and Shevell, 1987; Schwarz, 1990 and Strube, 1987). However, if the behavior is frequent, respondents may have a ready answer (in the form of a rate estimate) only when the behavior is highly regular (Menon, 1994). Thus, respondents often need to recall relevant information from memory and count instances

of occurrence (enumeration) or compute a judgment (rate-based estimation). The success of retrieval of information and its accuracy depend on time at task (for example, Williams and Hollan, 1981); the elapsed time (Cannell, Miller and Oksenberg, 1981; Means et al., 1989; Loftus, Smith, Klinger and Fiedler, 1992; Smith and Jobe, 1994) and the availability and adequacy of retrieval cues (for a review see Strube, 1987). In addition, Tulving and Thompson (1973) observed that retrieval accuracy further depends on the match between the encoding context and the recall context (principle of encoding specificity). The context may vary from physical context (Godden and Baddeley, 1975; Smith, 1988) to mental and emotional state (Eich, Weingartner, Stillman and Gillin, 1975; Bower, Monteiro and Gilligan, 1978; Bower, 1981). Several studies have demonstrated that the language in which mental activity is carried out during information encoding creates an internal context analogous to a mental state and can serve as a retrieval cue during information recall; similarly, the language spoken aloud during an event creates an external context analogous to a physical context and can serve as a situational cue during event recall (Schrauf and Rubin, 1998; Schrauf and Rubin, 2000; Marian and Neisser, 2000). Thus, a match between language of encoding and language of the recall task in surveys should yield more accurate responses among bilingual respondents.

3.1.1 *Language Dependent Recall*

Language dependent recall is the notion that the language of the question may influence retrospective reports. This phenomenon has been demonstrated in several bilingual groups in terms of number of recalled memories and time in life when the recalled events took place. Bugelski (1977) found that when Spanish-English bilingual immigrants were cued with Spanish words, 43% of their thoughts were related to post-immigration events, but when cued with English words, 70% of their thoughts were related to the post-immigration period. A decade later, Schrauf and Rubin (1998) found that 20% of the memories of elderly Spanish-English bilingual immigrants were recalled to mind in the language that was not spoken at the interview; that is, if an event took place in a Spanish context and had Spanish verbalization, that event could come to memory in Spanish, even when cued by an English prompt. Consistent with the language-dependent recall hypothesis, respondents reported more memories from ages 5-20 in response to Spanish rather than English prompts, but more memories from ages 20-55 in response to English rather than Spanish prompts (the average age at immigration was 28).

In another study of elderly Spanish-English bilingual immigrants Schrauf and Rubin (2000) distinguished between congruent memories (retrieved in the language used during the interview) and cross-over memories (retrieved in an alternate language than the language of interview) when assessing the language of first thought about an event.

Memories retrieved in Spanish came from earlier times in life (29 years of age for congruent and 27 years of age for cross-over) than memories retrieved in English (48 years of age for congruent and 53 years of age for cross-over), supporting the notion that language can facilitate recall in domains where the same language was used during information encoding. The study also found that language of interview influenced, but did not completely determine language of first thought – in the cross-over case, 21% of memories were retrieved in English during Spanish interviews and 30% of memories were retrieved in Spanish during English interviews.

Marian and Neisser (2000) further demonstrated language dependent access to autobiographical memory. Going beyond earlier findings of language-congruity effects, they investigated whether a match between language of encoding and recall facilitated retrieval because the language matched words during the original event or because the language at the time of recall induced a more general mind-set, resembling the processes assumed to underlie state dependent memory. The results showed that the effect of ambient language was significantly stronger than the effect of word-prompt language, further “strengthening the analogy between language-dependent recall and other forms of context dependency” (Marian and Neisser, 2000).

Such results suggest that retrieval in bilinguals, who have learned their first language through socialization in the culture of origin and later on, a second language through socialization in a second culture, is not the

same in both languages. One possible explanation, coming from cognitive psychology, is that remembering is analogous to state-dependent learning (Weingartner, 1978) and recall is more successful when language spoken at the time of retrieval matches the language of initial encoding. Another explanation, coming from cultural psychology, is that the self is presented by multiple identities (Ewing, 1990) that are culturally predetermined (Shweder and Bourne, 1982) and language primed. Thus, when a particular language is spoken, a “language-specific self” is activated, through which memories are encoded and retrieved (Schrauf, 2000).

The implication of such findings for surveys that involve immigrant and ethnic minority populations is that the choice of language of survey administration would affect both quality and quantity of recall. Specifically, first language cues tap into first culture memories, while second language cues activate more recent memories. This suggests that language of survey administration in bilingual bicultural respondents may be switched throughout the survey, depending on life periods for which researchers are interested in collecting data. Additionally, bilingual immigrants or ethnic minorities are likely to use different languages in different life domains – for example, at work and at home. The findings presented so far suggest, that recall of work related information will be facilitated by being asked in the dominant culture language, whereas the recall of home related information will be facilitated by being asked in the mother tongue. Specifically, we can expect that the match between language spoken at home and language of

administration will yield the most accurate information regarding “home” events, the highest number of such reported events, the lowest response latencies for home-related questions and vice versa. Such hypotheses would also suggest a language switch across domains of interest during the survey interview.

3.1.2 *Spatial Frames of Reference*

A different aspect of language dependent recall is demonstrated in studies on spatial cognition - the frames of reference used in a language to describe specific situations are likely to induce the same frame of reference in the nonlinguistic coding of the same situations for memory (Levinson, 2003). Various experiments (Pederson et al., 1998; Wassmann and Dasen, 1998; Levinson, 2003) have shown that when speaker of languages with different dominant frames of reference are given various memory and spatial reasoning tasks, the non-linguistic frames of reference used to carry out these tasks match the dominant frames of reference of the languages. Such tasks involved memory for spatial configuration of objects⁴, memory for motion and path-direction⁵ and spatial reasoning⁶ (see Pederson et al.,

⁴ Participants are seated at a table and shown a card printed with a large and small dot, such that the small dot is towards them, away from them, to the left, or to the right. After a 30 second delay, participants are rotated at 180 degrees and led to another table where they have to identify from a set of four cards the card (or counterpart card) they saw at the first table.

⁵ Participants are watching a toy man moving a particular path on Table 1. After a delay, they are rotated at 180 degrees and led to a maze set up at another table. Participants are asked to replicate the path of the toy man in the maize.

⁶ Participants are shown a cone and a cube arranged on a table. They are then rotated at 180 degrees to another table where they are shown a configuration of a cube and a cylinder. Finally, participants are rotated back to the first table where they find the cone

1998; Wassmann and Dasen, 1998; Levinson, 2003 for detailed description of the experiments). Specifically, speakers of languages that use absolute frames of reference such as Guugu Yimithirr and Arrernte (Australia), Hai//om (Namibia), Tzeltal (Mexico), Longgu (Solomons), Belhare (Nepal) and Balinese (Indonesia) preserved the absolute coordinates of objects when performing these tasks, while speakers of relative languages, such as Dutch, Japanese and Yukatek (Mexico) preserved the relative coordinates of objects (Pederson et al., 1998; Levinson, 2003).

At this point it remains unclear what are the cognitive consequences of being bilingual in languages that use different frames of reference, but the presented research supports the notion that language, through its inherent reference frames, may influence memory and recall. One possible mechanism is differential perceptual tuning due to use of different frames of reference. Studies have shown that language can affect perception such that individuals become more or less attuned to certain features of the environment (Goldstone, 1998; Sloutsky, 2003). For survey practitioners this may mean that what is reported during recall tasks may be related to what language is used during initial information encoding and later during the survey interview. In an extreme example, certain information may not be encoded *because* of the language spoken during an event that predetermines on what speakers focus their attention. On the other hand, similar to language dependent recall, it can be expected that a match

standing alone and their task is to place the cylinder next to the cone, keeping the location consistent with what they have seen.

between language frames during encoding and retrieval would facilitate remembering.

3.1.3 *Codability*

Sometimes, there is no direct correspondence across languages with respect to terms that describe the same phenomenon/object; thus, the use of phrases or multiple words to describe the concept of interest is necessary during translation. A hypothesis about the consequences of lack of direct match of terms for bilingual respondents is not straightforward. Research related to language codability would predict difficulty in recall with difficult-to-code words as easily coded words (therefore, events associated with them) are remembered better (Lucy and Wertsch, 1987; Lucy, 1992). However, multiple words may provide more contextual cues that can ease recall and eventually improve report accuracy. To date, it is unknown how such processes operate for users of two languages with different levels of concreteness for the same concept.

3.2 *Attitude Questions*

Attitude questions often require respondents to form an opinion on the spot in the specific context of the specific survey (Sudman et al., 1996). To do so, they need to form a mental representation of the question target based on the relevant information they are able to retrieve. However, respondents do not search for all relevant information, but rather stop the search process as soon as they have enough information to form a

judgment (Bodenhausen and Wyer, 1987); thus, opinions are often based on the most accessible appropriate information. Various stimuli, such as preceding questions, visual aids, interviewer characteristics, etc., can make certain information more accessible. Language of survey administration can also determine what information is most accessible at any given time in bilingual respondents by activating the cognitive-affective cultural framework associated with it. As discussed earlier, bilingual individuals possess more than one culturally constituted (Shweder and Bourne, 1982) and linguistically conditioned personality (Schrauf, 2000). By using a particular language, a “language-specific self” is activated who acts like a filter through which information is both encoded and retrieved (Schrauf, 2000).

3.2.1 Cultural Frame Switching

Language can affect what information is temporarily accessible by evoking a particular mindset related to the cultural meaning system associated with it. For example, a study of Greek students attending an American school in Greece showed that the correlation between the same attitudinal questions administered in English and in Greek was low for domains in which the Greek and American norms differed in what was considered socially desirable, and high for domains where the cultural values converged (Triandis et al., 1965). Similarly, Marin, Triandis, Betancourt and Kashima (1983) administered a questionnaire to English-Spanish bilinguals in both of their languages. The instrument consisted of

Hispanic emic (culture-specific) items, such as such familism, dignity, respect and obedience, and etic (universal) items from Hofstede's Values Survey (Hofstede, 1980). Higher discrepancies in responses obtained in the two languages were observed when the questions pertained to emic concepts, further supporting the hypothesis that differential social desirability primed by language was the reason for the observed differences between the Spanish and English versions.

Contrary to such findings, in a study of Chinese-English bilinguals, Bond and Yang (1982) reported that culture-confirming answers (bilinguals responding in a more Chinese fashion when asked in English) were present for topics that were considered of particular importance to the participants; however, responses that were appropriate in the culture associated with the foreign language were observed for items that were not considered salient. A possible explanation for the observed differences is ethnic affirmation (Yang and Bond, 1980). In its essence, ethnic affirmation maintains that presenting a questionnaire in a language different from the individual's mother tongue arouses awareness of individual's ethnicity and provokes more ethnic responses as ethnicity becomes a salient part of one's self-perception. Such findings present a promising area for future study of the reasons for cultural frame switching.

Another aspect of cultural frame switching relates to knowledge organization across cultures. Psychological research has demonstrated differences in how Westerners and East Asians organize the world –

Westerners show preference for grouping of objects based on taxonomy or common category membership, while East Asians prefer groupings based on relationships (Chiu, 1972; Ji, Zhang and Nisbett, 2002). For example, when presented with sets of three words, such as panda-monkey-banana and asked to identify the two most closely related, U.S. students used taxonomy (animal kingdom) to group panda and monkey together, while Chinese students used thematic relationship (monkeys eat bananas) to group monkey and banana together (Ji et al., 2002). Furthermore, such grouping preferences can be manipulated by the language used in administration of the cognitive task – Ji, Zhang and Nisbett (2004) found that relationship-based grouping shifted to categorical when Mainland and Taiwan Chinese participants were asked in English. Similar effects however were not reported for Hong Kong and Singapore Chinese⁷. Such findings possibly rule out the hypothesis that perceptual differences are language-only driven in support to the language-priming-culture hypothesis. However, recent studies in psycholinguistics have demonstrated that language can draw attention to what to compare to what (Bowerman and Choi, 2003; Gentner, 2003) and to the extent to which languages may classify according to different criteria, the extracted similarities will also

⁷ A plausible post-hoc explanation offered by the authors is that Mainland and Taiwan Chinese acquired their second language later in life relative to their mother-tongue, thus had separate representations in Chinese and English. In contrast, having English as one of the official languages, Hong Kong and Singapore Chinese were more likely to have acquired it parallel to Chinese, thus shared the same representational system in both languages.

differ (Boroditsky, 2001; Lucy and Gaskins, 2001; Boroditsky, Schmidt and Phillips, 2003).

There are several implications of such findings for surveys of immigrant and ethnic minority bilingual populations. First, what information is accessible to form an opinion will vary depending on language of survey administration. To achieve maximum equivalence of language versions of the same instrument, this would suggest that open ended questions should be avoided. Second, the same questions can be perceived to have different affective characteristics depending on language and the cultural norms it activates; thus, more or less socially desirable opinions will be expressed depending on language. Knowing in advance how cultures differ in terms of a question's affective characteristics may better inform questionnaire design and various techniques may be used to reduce social desirability or sensitivity across language versions. Third, judgments can be language dependent as comparisons are based on culture-approved practices and/or how language systems are organized. Such hypotheses necessitate systematic investigation of language effects and the underlying dynamics across question types.

3.2.2 Codability

Studies in psycholinguistics have demonstrated that codability affects judgement. Kay and Kempton (1984) showed that color-naming practices affect judgments among colors – speakers of Tarahumara (a Mexican-Indian language that does not have separate words for blue and

green) differentiated among color chips on the blue-green color continuum based on their physical characteristics – wavelength of reflected light. In contrast, English speakers differentiated among the same color chips based on labels, such as “shade of green” and “shade of blue” and deemed chips that were very close in wavelength, but labeled as “blue” and “green” more different than chips further apart in wavelength of reflected light, but labeled “green”. Thus, English speakers evaluated colors in terms of categories in which they were easily coded, while Tarahumara speakers, lacking such codability of colors, based their evaluations on physical dimensions.

Hoffman, Lau and Johnson (1986) examined to what extent the codability of personality description (existence of stereotype) in a language influenced the impression about a person. Chinese-English bilingual respondents were asked to read personality descriptions in Chinese and English and write their impressions about the person in the paragraph, as well as make further inferences about this person’s behavior. The study found that terms that were readily available in the language lead to stereotyped impressions and participants were more likely to elaborate on the described person’s characteristics using terms consistent with the stereotype, than when a verbal label was not available.

Such findings may have implications for surveys of bilingual respondents. For example, scales may be judged differently depending on whether scale labels are easily codable in both languages. If label equivalents are not easily codable in one language, then respondents may

be more likely to consider solely the numeric scale when making judgments. Such differential scale use will result in differences across language versions.

In sum, the choice of language of survey administration may affect retrieval and judgment processes and has several implications for surveys of bilingual bicultural respondents: 1) language of interview may determine the quantity and quality of recall through the information that is made accessible; 2) the inherent frame of reference of each language and its codability may affect to what respondents pay attention; thus, comparison and judgment, and 3) cultural norms activated by a language would determine a question's affective characteristics.

4. Response Formatting

Once a judgment is formed, it has to be matched to the provided response alternatives. The latter themselves can affect the response (for a review, see Schwarz and Hippler, 1991). Additionally, the ability to differentiate among response options may be influenced by language codability and the stimuli used to anchor the points of a rating scale may be affected by the cultural meaning system primed by language.

4.1 Codability

Differentiation among objects and categories has been demonstrated to be easier when separate words exist in a language (Kay and Kempton,

1984). The findings reported in section 3.2.2 may have implications not only for retrieval and judgment, but also for response formatting. For example, scales may be used differently by speakers of different languages as a result of different scale label codability; thus, the meaning of the same number on a labeled scale may be affected by what language is used. Taken to an extreme, there are cultures whose languages have terms only for one, two and many (Greenberg, 1978), which further limits the ability of their speakers to make comparisons (Hunt and Agnoli, 1991). At this point it remains unclear how this may affect the cognitive processes in bilinguals whose other language allows them to utilize the whole numeric scale. It can be speculated that the ability to make comparisons may remain language dependent.

4.2 Cultural Frame Switching

Cultural frame switching can further complicate the investigation of language effects at the formatting stage as scale anchoring may be affected by the reference frame primed by language. Such differences in scale anchoring may be reflected in the observed differential response styles across cultures. For example, several studies have reported that East Asians avoid extreme responses (Zax and Takahashi, 1967; Chun, Campbell and Yoo, 1974; Chen, Lee and Stevenson, 1995; Stening and Everett, 1984; Hayashi, 1992). While such differences are often attributed to differential emphasis on conflict avoidance and humbleness, it is unclear whether they are an artifact of self-presentation as a result of language

priming culture, or true differences in perception, independent of language. Thus, based on findings presented so far, it can be speculated that Chinese-English bilinguals would avoid extreme responses when asked in Chinese (a culturally appropriate behavior), but would be likely to use the whole response scale when their Western self is primed by English.

Additionally, the extent to which respondents use the range of a presented frequency scale as a frame of reference when reporting behavioral frequencies is also culture-dependent. A study by Ji, Schwarz and Nisbett (2000) demonstrated that Chinese students were influenced by the range of frequency scales only when asked to report private, unobservable behaviors (e.g., having nightmares, borrowing books from the library). However, no scale effects were found for public behaviors (e.g., being late for class), possibly reflecting the importance of “fitting in” in Asian cultures; thus, monitoring one’s and others’ public behaviors and as a result having better memory representations for such behaviors. In contrast, consistent with previous research on scale effects (for a review, see Schwarz, 1996), American students relied on the presented response scale frequency range to estimate both private and public behaviors. For surveys of bilingual biculturals such findings may suggest that depending on the cultural identity primed by the language of interview, different estimation strategies may be employed.

Overall, how responses options are treated may be influenced by the choice of language of survey administration, resulting in a possibility for

different responses provided to the same questions by the same bilingual bicultural respondent answering in different languages and within different cultural frameworks.

5. Response Editing

Respondents sometimes edit their responses before reporting them, reflecting social desirability and self-presentation concerns (Sudman et al., 1996). Gender, age, socio-economic status, and various survey design characteristics have been found to be correlates of socially desirable responding (for a review, see DeMaio, 1984). Recent work in cross-cultural research suggests that culture influences social desirability through interpretation based on cultural experiences and response editing depends on the necessity to conform with particular social norms (Fu, Lee, Cameron and Xu, 2001; Lee et al., 2001).

5.1 Cultural Frame Switching

Reports on racial differences on social desirability scales support the idea that culture plays an important role in evaluation of target behaviors and opinions (for a review, see Johnson and Van de Vijver, 2003). Thus, the same question may be perceived to have different levels of socially desirable content depending on the respondent's cultural identity. For example, maintaining harmony and face-saving are more socially desirable traits in Asian cultures than in the Western world (Triandis, 1995). Similarly,

mental health is stigmatized in Arab and Hispanic societies (Bazzoui and Al-Issa, 1966; Okasha and Lotalif, 1979; Chaleby, 1987; Silva de Crane and Spielberger, 1981), but not so in Western cultures. For bilingual bicultural respondents this would mean that, depending on the language of the survey interview and the cultural frame primed by it, such questions would be perceived to have different affective characteristics and will provoke socially desirable responding in one case, but will not in the other. That is, respondents would be likely to edit their answers to match the values of the culture associated with the language. Section 3.2.1 discussed studies by Triandis et al. (1965) and Marin et al. (1983) that illustrate this effect. For survey practitioners this would require thorough advance knowledge of where cultural differences related to questions affective characteristics are to be expected in order to determine the language assignment of bilingual bicultural respondents or to employ questionnaire design techniques that reduce the differentially perceived social desirability/sensitivity between language versions.

6. Summary

A substantial body of literature in psycholinguistics and cross-cultural psychology suggests that language of survey interview may potentially impact every stage of the response formation process. Different mechanisms may play a role at the same time at each step of the response

process. As our discussion indicates, depending on language, respondents may answer the same question differently as a result of different question interpretations, different mental representations of the question target, a mismatch between the language of encoding and language of recall, different accessible information at the time of the survey request, differential anchoring of response scales, and differential self-presentation concerns.

Two shortcomings of the presented theoretical framework relate to its application. It is desirable to directly connect the outlined hypotheses to published survey research and possibly reinterpret puzzling results in light of the proposed language influences, but the existing cross-cultural survey data does not offer such opportunity. Thus, the presented framework remains largely speculative. Next, some of the presented mechanisms are demonstrated through research in cultures at a much earlier economic stage of development than cultures that typically conduct surveys. We believe the merit of this work is to present possibilities for language influences and stimulate further discussion and action around these issues.

Systematic research on language effects on the response formation process may have significant impact on the way we conduct surveys of respondents who are fluent in more than one language and competent in more than one culture. For example, differences in context dependency across cultures should be taken into consideration at the questionnaire design stage. Furthermore, it should be acknowledged that the same question may have different affective characteristics across cultures. The

effect of scales and their orientation should also be carefully considered when languages that have different dominant frames of reference are used. The effect of language due to absence of certain words and direct translation in some languages should be examined in order to anticipate the consequences for measurement error. Additionally, the finding that the match between language of encoding and recall facilitates the quality and quantity of retrieved information opens the possibility of conducting surveys with bilingual immigrants in both their languages, depending on the time period or life domains of interest.

Before attempting to disentangle the mechanisms that may produce differences across language versions, it is important to determine whether language influences are indeed of practical relevance to survey research. Chapter III examines the effect of language using observational data from the New Immigrant Survey, while Chapter IV studies the same problem in the experimental setting of the National Latino and Asian American Survey. Both chapters seek insight into what questions are most likely to be affected by language of survey administration and compare sets of questions where differences are to be expected based on divergence of cultural norms and beliefs, and questions where no such differences are expected. Strong language effects may reflect dissimilarities in the response process due to language, true difference between the cultural identities of bilingual bicultural respondents, or a combination of language and culture influences.

Such differences are of particular importance for surveys of immigrants where often the choice of language is left to the respondent.

Chapter III. Effect of Language of Survey Administration on Responses to Survey Questions

1. Introduction

1.1 Language, Cognition and Culture

National and cross-national surveys usually offer multiple languages to accommodate respondents from different ethnic minority groups and immigrants. The survey literature has focused on developing better translation practices to achieve measurement comparability across language versions of the same questionnaire (Harkness, Van de Vijver and Mohler, 2003). However, the language in which a survey interview is conducted is likely to affect the response process in numerous ways that received little attention in survey methodology.

The effect of language on cognition has long puzzled researchers in linguistics and psychology. Aristotle's view that language is a tool to communicate thoughts, thus the different categories in language reflect the categories used in thinking, has evolved into theories that view the language system as unique and separate from the rest of the cognitive system (for example, Chomsky, 1965), or ones that assume that language determines thought (Whorf, 1956). The latter hypothesis has regained

interest in psycholinguistics, and various studies have demonstrated support for its much weaker version, namely, that language and thought are interdependent (for example, Kay and Kempton, 1984; Maass, Salvi, Arcuri and Semin, 1989; Lucy, 1992; McConnell and Fazio, 1996).

Multiple studies have demonstrated that language can be used as a cue to prime culture (Botha, 1968; Feldman, 1975; Bond and Yang, 1982; Trafimow et al., 1997; Watkins and Gerong, 1999; Ross et al., 2002; Watkins and Regmi, 2002; Kimmelmeier and Cheng, 2004; Marian and Kaushanskaya, 2004). The implication of such findings for survey methodologists is that language has the potential to influence each stage of the response formation process. For example, language can affect comprehension by inducing the cultural frame associated with it, thus potentially changing the literal and pragmatic meaning of survey questions (e.g., Trafimow et al., 1997). Similarly, language can affect retrieval and judgment by evoking a particular mindset related to a given cultural meaning system, thus determining what information is temporally accessible and/or serving as a retrieval cue during information recall (e.g., Marin et al., 1983; Marian and Neisser, 2000). Furthermore, the reference frame primed by a language may affect scale anchoring at the formatting stage and/or the necessity to consider particular social norms at the editing stage (e.g., Fu et al., 2001; Lee et al., 2001; for an overview of language effects at each stage of the response formation process, see Chapter II).

Such effects become particularly important when respondents are fluent in more than one language, as the choice of survey language has the potential to influence what responses are given. Bilingual individuals are not only knowledgeable about another language, but often have experienced the culture associated with it, thus possessing dual cultural identities (Berry and Sam, 1996). Such bicultural bilinguals are usually individuals who have lived in another country for an extended period of time (for example, immigrants, international students or foreign workers) or who were born in ethnic minority families. The acculturation literature reveals that some biculturals succeed in developing compatible cultural identities (LaFromboise, Coleman and Gerton, 1993; Rotheram-Borus, 1993; Padilla, 1994; Phinney and Devich-Navarro, 1997), but many perceive the mainstream and ethnic cultures as highly distinct, even oppositional (Gil, Vega and Dimas, 1994; Phinney and Devich-Navarro, 1997; Chuang, 1999). This may be of particular interest to survey methodologists, as cues from different cultures (e.g., language of interview, interviewer characteristics, survey sponsor) may affect survey responses.

Each cultural meaning system serves as an interpretive frame that affects individual's cognition, emotion and behavior (Geertz, 1993; Hong et al., 1997; Kashima, 2000; Mendoza-Denton et al., 1999). Individuals can shift the operative cultural meaning system, depending on situational cues and requirements of the surrounding environment; a phenomenon known as "cultural frame switching." Which cultural identity is dominant can be

affected by generational status (Tsai, Ying and Lee, 2000), language assimilation (Laroche, Kim, Hui and Tomiuk, 1998), sociopolitical climate (Fordman and Ogbu, 1986; Berry, 1990; Gurin, Hurtado and Peng, 1994) or situational cues (Hong et al, 2000, 2001). Of most interest to the present research is language's property to cue culture. For example, Earl (1969) found that Hong Kong bilinguals provided more "Western" responses to a dogmatism scale when they answered in English versus Chinese. Similarly, Trafimow et al. (1997) found that bilingual Hong Kong students reported more private traits and fewer social roles when answering in English versus Chinese, consistent with the associated cultural emphasis on individual vs. collective aspects of identity. Ross et al. (2002) also used language to shift the cultural frame adopted by Chinese-born students in Canada in a task of self and mood description. As expected, those asked in Chinese provided collective statements and reported higher endorsement of Chinese views. Furthermore, a study of Greek students attending an American school in Greece showed that the correlation between the same attitudinal questions administered in English and in Greek was low for domains in which the Greek and American norms differed in what was considered socially desirable (Triandis et al., 1965). In contrast, the correlations between questions administered in the two languages to the same respondents were high for domains in which the cultural values converged. Such studies suggest that questions can be answered within the cultural frame primed by the language of the questionnaire.

Alternatively, the use of a foreign language may cue respondents to more strongly endorse their ethnicity. Yang and Bond (1980) and Bond and Yang (1982) found that Chinese bilinguals from Hong-Kong reported overall more traditional Chinese views when the questionnaire was in English. However, when individual items were analyzed, both ethnic affirmation and cross-cultural accommodation (matching to the norms and values of the culture associated with the language) were found. Furthermore, the nature of the selected items and the post-colonial time period when the experiments were conducted may explain the unusual results – the authors used traditional Chinese beliefs that were well known to be perceived as superstitious by the Westerners.

Overall, the presented results suggest three important consequences for survey methodologists: a) language of survey administration may affect responses, b) different items in the same questionnaire may exhibit differential language effects, and c) the strongest language influences should be expected on topics where two cultures take diverging standpoints.

1.2 Language in Surveys of Bilingual Respondents

A common practice of U.S. surveys that offer more than one language is to leave the choice of language to the respondent (e.g., the Current Population Survey, the New Immigrant Survey, the National Survey of Latinos). Various factors may determine a respondent's self-selection into a language – they can be respondent-related (e.g., level of

acculturation and comfort with the mainstream language), interviewer-related (e.g., ethnicity, accent, and visual cultural cues), and interview-situation-related (e.g., perceived sponsorship and possible privacy concerns when other family members are present). At present, little is known about whether controlling for such factors would make language effects negligible. Further, if that were not the case, we do not know the consequences for measurement error of providing different responses depending on language versions. It can be hypothesized that when the cultural identities of the respondent are in opposition to each other (e.g., premarital sex is acceptable in Western cultures, but not in Arabic), language may be a strong cue to what types of responses are socially acceptable. As a result, different answers to the same question may be provided depending on the cultural norms evoked by the language. However, not all questions should be equally affected by language effects. For example, questions that ask for factual information such as respondent's gender, country of birth, etc., are not likely to receive different answers when the language of interview changes. In contrast, questions that differ in the perceived levels of sensitivity and social desirability across cultures or those directly related to cultural values and norms should exhibit stronger language effects (e.g., Triandis et al., 1965; Marin et al., 1983).

The theoretical foundation presented so far demonstrates that language effects are not culture-specific i.e., we expect to find them in any combination of cultures as long as they are associated with different

languages and the cultures take diverging positions on an issue. Such cultural differences are often related to religion, societal structure, customs, and habits. For example, the Latino culture differs from the American in many aspects. The existing cross-cultural and psychological research has demonstrated that Hispanics and Americans differ in the endorsement of individualist and collectivist dimensions of identity, with Hispanics having a stronger collectivist orientation (Marin and Triandis, 1985). Two values that are central to the Hispanic but not the American culture are *familism* (importance and strong connection to family) and *simpatia* (need for being nice and polite with others; agreeableness with others) (Lisansky, 1981; Triandis, Marin, Lisansky and Betancourt, 1984; Sabogal, Marin and Otero-Sabogal, 1987; Marin and Marin, 1991; Vasquez, 1994; Levine, Norenzayan and K., 2001; De Las Fuentes, Baron and Vasquez, 2003). The latter has also been associated with higher tendency among Hispanics to acquiesce (Marin et al., 1983; Hui and Triandis, 1989; Marin, Gamba and Marin, 1992) or to provide socially desirable responses (for example, Ross and Mirowsky, 1984; Warnecke et al., 1997). Such tendencies have been observed in various domains and may explain some puzzling results. For example, consider the notorious “immigration paradox” in health and mental health sciences. Recent immigrants, despite poverty and more difficult access to health care, report better physical and mental health than their more acculturated or U.S.-born counterparts (Vega et al., 1998; Ortega, Rosenheck, Alegria and Desai, 2000; Grant et al., 2004). This may be

related to social desirability and may reflect the evoked cultural values during the interview (note that the language of survey administration for acculturated immigrants is usually English; for example, Breslau et al., 2007).

In addition, extensive health research reveals that culture often determines whether and what symptoms are reported, whether health care is sought, what meaning is imparted to the illness and how much stigma is attached to it (U.S. Department of Health and Human Services 2001). The value of familism is reflected in different attitudes among Hispanics and European Americans towards family issues, including the belief that the family is responsible for one's health problems (Sabogal et al., 1987) and that mental illness is best treated within the family (Edgerton and Karno, 1971). Furthermore, there is some evidence that psychiatric disorders may have greater stigma attached to them among Hispanics than among European Americans (Silva de Crane and Spielberger, 1981). Thus, we would expect that survey administration in Spanish (versus English) would prompt bilingual bicultural respondents to report fewer psychiatric and general health problems.

Similarly, the 1992 National Longitudinal Alcohol Epidemiologic Study suggests that the self-reported percentage of current and former drinkers is significantly lower for Hispanics relative to non-Hispanics (Dawson, Grant, Chou and Pickering, 1995). Estimates from the 2000 National Household Survey on Drug Abuse (2001) reveal that Hispanic

respondents are less likely to report current use of alcohol (39.8% vs. 50.7% for Whites). Like health reports, reports of alcohol use can be influenced by social and cultural factors, such as group norms and attitudes towards alcohol. Hispanics have been reported to have more conservative alcohol norms and attitudes than Whites (Caetano and Clark, 1999). Weaker intentions to use alcohol, tobacco and marijuana over the course of adolescence have also been reported for Hispanics relative to other ethnic groups (Maddahian, Newcomb and Bentler, 1988). The existing literature suggest that even though alcohol consumption is more socially undesirable, it is not less prevalent in the Hispanic culture (Clark and Hill, 1991; Greenfield and Kaskutas, 1998). Additionally, consistent underreporting of drinking and driving has been found in multiple studies, as compared to official DMV records (Chang, Lapham and Barton, 1996; Jason and Cherpitel, 2004). Thus, subgroup differences in the rates of alcohol consumption are likely to be a result of underreports of alcohol consumption among Hispanics, provoked by social desirability, rather than true differences. Given this, we would expect to see underreporting of alcohol consumption when bilingual bicultural respondents are reminded of their Hispanic identity by being interviewed in Spanish as compared to English.

Survey questions vary in their susceptibility to measurement error and in the mechanisms that induce measurement error. For example, the accuracy of responses to autobiographical questions depends on passage of time (Cannell et al., 1981; Loftus et al., 1992; Means et al., 1989; Smith

and Jobe, 1994), length of reference period (see Sudman and Bradburn, 1973 for a meta-analysis), event salience characteristics (e.g., Thompson, Skowronski, Larsen and Betz, 1996; Wagenaar, 1986), and question aids used to improve recall (e.g., Brewer, 1988; Wagenaar, 1986). Thus, it may be expected that language, as an external stimulus to response formation, may not perform similarly across different types of survey questions. Language is likely to be most powerful when the questions pertain to topics on which the cultures hold different expectations (e.g., Triandis et al., 1965). On the other hand, questions related to well-defined and accessible facts, such as one's marital status, number of biological children, or whether respondent lives alone, are not likely to change depending on the language of interview. Thus, we would examine the hypothesis that language of survey administration will affect responses among Hispanic bicultural-bilingual respondents in the U.S. if and only if the two cultures differ in their social desirability norms relevant to the questions being answered. Specifically, we would predict that Hispanic bicultural-bilingual respondents interviewed in Spanish will report lower rates of general and mental health problems and alcohol consumption relative to those interviewed in English, as a result of language cueing Hispanic cultural values and inducing socially desirable responses for this culture. In contrast, we would expect that language of survey administration will have no effect on responses provided by Hispanic bicultural-bilingual respondents to questions related to

well-defined and stable facts, such as number of biological children, marital status and living situation.

2. Data and Methods

An investigation of the effects of language on survey measurement error would ideally be based on a large scale survey using a random assignment of bicultural bilinguals to one of the two relevant languages. However, the general practice of U.S. surveys that employ multiple languages has been to leave the choice of language to the respondent. As language effects are likely to be item-specific, it is important to be able to study them even in the absence of random assignment to a language. Statistical methods, such as propensity score modeling, albeit dependent on the specification of the propensity model, allow us to explore the influence of language when respondents self-select themselves in a language of survey administration.

2.1 Dataset

The New Immigrant Survey (NIS) is a longitudinal study of legal immigrants to the United States and their children, assessing migration behavior and its impact on economy. The design involves drawing representative samples of new immigrant cohorts every four years and following each cohort over time. Each cohort is selected from the electronic administrative records compiled by the United States Citizenship and Immigration Services and the Office of Immigration Statistics. Eligible

sample persons include adult immigrants and their children legally admitted for permanent residence. Immigrants are interviewed in the language of their choice; however, not all languages are treated equally. NIS classifies languages into tiers based on expected origin-country distribution, native language distribution and preferred language by country. Thus, there are 5 tiers – tier 0 is English and based on the pilot results, it was expected to be the preferred language (Jasso, Massey, Rosenzweig and Smith, 2005); tier 1 is Spanish; tier 2 includes Chinese, Korean, Polish, Russian, Tagalog, and Vietnamese; tier 3 includes Arabic, Croatian, Farsi, French, Gujarati, Hindi, Serbian, Ukrainian and Urdu, and tier 4 includes all other languages. Only Spanish (tier 1) was treated equally to English and received a CAPI implementation. The survey instrument was fully translated only into tier 1 and tier 2 languages. Only key concepts (such as citizen, alien, lottery visa, child support, food stamps, etc.) were translated in the other language tiers and presented to the respondents along with the English original. The survey questionnaire was translated into Spanish by a professional translation firm, and an assessment of the translation was carried out by NORC team of bilingual translation experts.

The questionnaire covers a variety of topics such as health, schooling, English language skills, income and assets, employment, use of government services, social networks, travel and religion. Some of these topics differ in social desirability in the United States and other countries (for example, health, alcohol consumption, etc.).

Data from the first cohort of new immigrants (2003) have been collected and released. The overall response rate is 69%.

For the purpose of this investigation, the analyses are limited to Spanish-English bilingual adults, identified based on Latin American immigrants' self-reported ability to understand and speak English. Thus, the total subsample size is 632 self-identified bilinguals, 261 of which chose to be interviewed in English. Almost half of this subsample were respondents born in Mexico (n=301), while the rest came from Colombia, Cuba, Dominican Republic, El Salvador, Guatemala and Peru.

2.2 Causal Inference in the Absence of Random Assignment

In order to study the effect of language on survey responding in the absence of random assignment to a language of interview, we need to account for the self-selection of respondents to a particular language by controlling for many background characteristics. Respondents could be grouped using correlates of language selection, so within each group they are equally likely to select a given language. The propensity score technique (Rosenbaum and Rubin, 1983) addresses this issue by replacing the collection of correlates with a single composite characteristic (a propensity score). This approach enables us to assess whether the two language groups overlap enough in terms of respondent characteristics to allow estimation of language effects from the data set. The propensity score is estimated by predicting group membership (i.e., being interviewed in English versus Spanish) from the set of confounding covariates (e.g.,

education, ability to speak and understand English, language used at home, etc.) using logistic regression. Each respondent has an estimated probability (determined by the covariate values) of being interviewed in English rather than Spanish, regardless of the language he/she actually selected in this survey. The region of overlap of the propensity scores for those interviewed in English and those interviewed in Spanish is an important indicator for the usefulness of this approach. When the overlap is limited, causal conclusions about the differential effect of language cannot be supported.

A fundamental assumption in the use of propensity scores to remove bias from self-selection is that they include correlates of the (language) selection process. For language selection, the literature would suggest the inclusion of measures of acculturation, reasons for immigration, language use and demographic characteristics (e.g., Laroche et al., 1998).

After the model is built, subclassification of respondents into about five groups (usually, quintiles) based on their individual propensity scores is performed (Rosenbaum and Rubin, 1983). Within a group of respondents with the same or similar propensity scores, those interviewed in English and Spanish essentially have the same likelihood of choosing English regardless of the language in which they were interviewed, that is, within each propensity score group, we simulate random assignment conditional on the propensity model specification. Before examining any causal relations between language of interview and any outcome measures of

interest, the subclasses are examined for balance with respect to the covariates. If the propensity scores are relatively constant within each stratum, then the distribution of covariates should also be approximately the same in both language groups within a stratum (Rosenbaum and Rubin, 1984). If significant within-stratum differences are found on some covariates, then the propensity score prediction model has to be reparameterized or it has to be concluded that the covariates do not overlap sufficiently, so the subclassification does not have the property of adjusting for these covariates (for an example of this cycling process, see Rosenbaum and Rubin, 1984).

Finally, analyses of residuals can reveal omitted model interactions. Separate models for each covariate as a dependent variable and the raw propensity score as the predictor are run and the Pearson residuals plotted for those interviewed in English and those interviewed in Spanish. If the residuals for the two groups do not overlap, further reparameterization of the propensity model is needed.

The outcome measures of interest (reports of mental and health problems, alcohol consumption, marital status, living situation and number of biological children, see Figure 3, Appendix) are then compared for the language groups, accounting for stratification by propensity score. Most outcome measures in this investigation are dichotomous. An indicator for reported health problems was constructed based on multiple questions that asked about diagnoses of high blood pressure, diabetes, cancer, lung

disease, heart disease, stroke, arthritis, asthma and whether the respondent was often troubled by pain.

When interpreting the results, it is important to keep in mind that propensity score methods can only adjust for the observed confounding covariates. In observational studies, confidence in the causal conclusions is often based on replication and how sensitive the conclusions are to deviations from the model assumptions.

3. Results

3.1 Propensity Score Model

A logistic model predicting the probability to select English over Spanish among bilingual immigrants was fit as a function of available correlates of language selection. Bilingualism was determined based on self-reported ability to understand and speak English on a four point scale (very well, well, not well, not at all) – only those who identified themselves as able to understand and speak English “very well” and “well” were included in the analyses. Those unable to understand and speak English well were not of interest for the current investigation because of the possible different mechanism related to their second language use⁸. One possible weakness of this approach is the potential for measurement error

⁸ According to the Revised Hierarchical Models for representation of bilingual's languages (Kroll and Sholl, 1992; Kroll and Steward, 1994; Dufour and Kroll, 1995), novice bilinguals access semantic referents through their first-language lexicon – a “think-then-translate-then-speak process”. In contrast, expert bilinguals think in their second language.

in self-reports – to the extent to which ability to speak and understand English is a socially desirable characteristic of immigrants to the U.S. and reports of such ability are influenced by the language of interview (completed in the self-selected language), the sample of self-determined bilingual respondents may not be ideal for testing language influences.

The predictors in the logistic regression model included immigrant status (new arrival or adjusted status from another immigrant or non-immigrant category), visa type related to reason for immigration (employment, diversity [green-card] lottery, married to U.S. citizen, other), whether the respondent took English classes before coming to the U.S., whether the respondent was enrolled in English classes, current language at home, at work and with friends, ability to understand and speak English, mode of the interview (face-to-face or telephone), whether the respondent had plans to travel home within the next 12 months, and demographic characteristics such as age, country of birth, whether the respondent spent his/her childhood in a Spanish-speaking country, whether the respondent grew up in a rural area, gender, education, years spent in school in the U.S., whether the respondent has been employed since coming to the U.S., current employment status and whether the respondent belonged to a church⁹ (Table 1). Ethnic identification (Hispanic or not) was considered, but not included in the model due to low variability (97% of respondents

⁹ Income was not included in the propensity model as the various questions that collected information on income were plagued by more than 20% item missing data (e.g., data on *income from wages and salaries* was missing for 26% of the respondents eligible to answer this question).

identified themselves as Hispanic) and concerns that responses to this question may be affected by language of interview (i.e., those who chose to be interviewed in Spanish were more likely to identify themselves as Hispanics). Similarly, interviewer observations, such as respondent's understanding of the questions were bound to be influenced by the initial choice of interview language, thus were not included in the propensity model. The pseudo R-square for the model was 0.24 and the Hosmer and Lemeshow goodness of fit test was $\chi^2_8=8.39$. The strongest predictors of likelihood to select English as the language of interview were *number of years in a U.S. school*, respondent's self-evaluation of *ability to speak English*, *language spoken at home* and *country of birth* ($p<0.001$).

Table 1. Logistic Regression Coefficients for Likelihood of Selecting English Language for the Survey Interview

Predictor	Coefficient	Standard Error
Intercept	-0.64	0.56
Adjustment status: New arrival	0.15	0.41
Adjusted immigrant	-	-
Visa status: US citizen spouse	0.31	0.25
Employment	0.76**	0.34
Green card lottery	2.45	1.32
Other	-	-
Gender: Male	0.11	0.21
Age: 59+	-0.27	1.22
49-58	1.085	0.56
44-48	-0.28	0.58
39-43	-0.046	0.41
34-38	0.16	0.37
29-33	-0.34	0.33
24-28	-0.0039	0.32
18-23	-	-
Years spent in school	0.020	0.031
Years in school in the US	0.16***	0.038
Enrolled in English Classes	0.098	0.30
Current language at home: Not English	-0.60***	0.21
Language at work: Not English	-0.42	0.44
Language with friends: Not English	-0.53**	0.25
Country of birth: Not Mexico	-0.56***	0.21
Country of childhood: Not Spanish speaking	0.018	0.38
Country of childhood: Rural	-0.010	0.22
Not worked since moved to the US	-0.64**	0.29
Employment: temporarily out of job	0.92	0.52
Retired/disabled	-0.019	1.10
Homemaker	1.28**	0.57
Student	0.27	0.71
Employed	-	-
Understands Spoken English Well	0.28	0.28
Understands Spoken English Very Well	-	-
Speaks English Well	-0.97***	0.28
Speaks English Very Well	-	-
Took English Classes before coming to the US	0.52**	0.22
Member of a church	0.053	0.21
Does not plan to Travel Home Next 12 months	0.17	0.20
Telephone mode of administration	0.0069	0.20

*** p<0.001 ** p<0.05

Propensity score methods rely on sufficient overlap between the groups defined on the independent variables; if no overlap exists, there are no comparable respondents across the two selected languages. There was a large enough region of overlap of raw propensity scores for those interviewed in English and those interviewed in Spanish to suggest we can sensibly estimate the effect of language from these data (Figure 2).

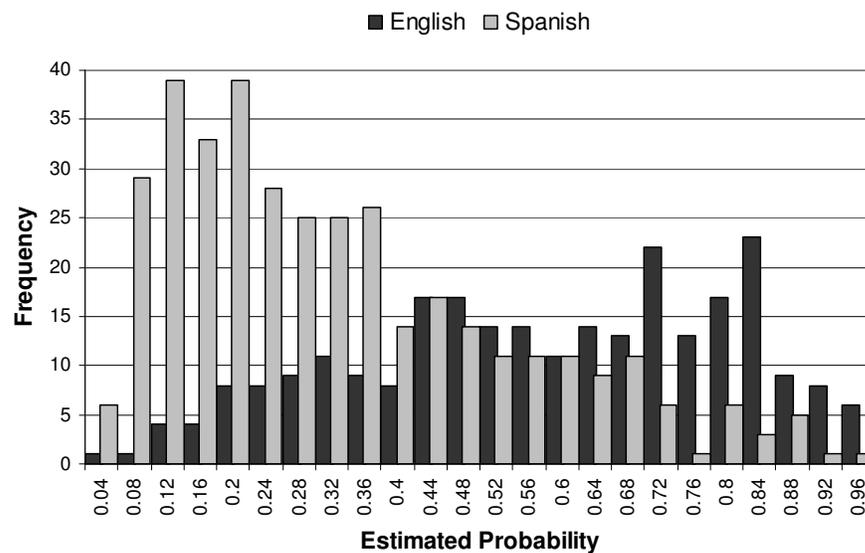


Figure 2. Overlap of Estimated Propensity Scores for Those Interviewed in English and Those Interviewed in Spanish

The raw scores were divided into quintiles (propensity strata) as shown in Table 2. The distributions of covariates within each propensity stratum were not significantly different (based on Chi-square and t-tests) for the two language groups (Table 6, Appendix). Finally, the overlap of

residuals for the two language group was examined and no need for model reparameterization was discovered (results not presented)¹⁰.

Table 2. Number of Cases in Propensity Strata by Language of Survey Administration

Propensity Stratum	Language of Interview		Total
	English	Spanish	
1	11	115	126
2	30	97	127
3	51	75	126
4	68	59	127
5	101	25	126
Total	261	371	632

3.2 Hypotheses Tests

Table 3 presents the results of language comparisons across measures of physical health, mental health and alcohol consumption, marital status and living situation, accounting for stratification by propensity class. Recall that we expected to find significantly higher rates of diagnosed physical health problems, higher reports of depression and psychiatric problems and alcohol consumption when respondents chose to

¹⁰ Despite the well-balanced strata that resulted from the main effects model, several interactions (whether the *respondent took English language classes with adjustment status* and with *years in school*; *country of birth with ability to understand English* and with *gender*) were also tested. Even though some of the interactions were significant, we were not able to achieve good balancing across propensity strata and the contribution of these terms to the overall model fit was negligible (pseudo R-square of 0.27). Any attempts to introduce additional interactions to improve balancing resulted in model failure to converge and reach a maximum likelihood estimate. Thus, the model presented in Table 1 was chosen as the final.

respond in English. In addition, we expected to find no difference for reports on marital status, living situation and number of biological children.

Table 3. Relative Risk for the English Language Group versus Spanish across Measures by Expectations of Language Effect

Outcome Measure		Mantel-Haenszel Relative Risk for English	95% Confidence Interval [†]	
Language Effects Expected				
Emotional, nervous or psychiatric problems		0.90	0.54	1.50
Sad, depressed for two weeks or more		0.80	0.63	1.01
At least one diagnosed health problem		1.07	0.87	1.32
Respondent drinks alcoholic beverages		1.24 ^{***}	1.05	1.46
Language Effects not Expected				
Marital Status	Married vs. Divorced/Widowed	1.02	0.64	1.65
	Married vs. Never married	1.12	0.94	1.33
Living alone		0.81	0.51	1.29

*** p<0.001 **p<0.05

[†] Standard errors reflect stratification by propensity strata

We examine the probability that those who were interviewed in English report more physical health problems, psychiatric problems, depression and use of alcohol relative to those interviewed in Spanish. Mantel-Haenszel test for stratified contingency tables allows us to avoid the Simpson paradox and to test for tendencies of a particular cell to be higher or lower in series of tables. The results are rather surprising – we found significant effect of language of survey administration only on reports of current alcohol consumption i.e., those who were interviewed in English were 1.24 times more likely to report current alcohol use relative to those

interviewed in Spanish. The magnitude of this effect varied across propensity strata and was strongest in the stratum of respondents with medium propensity to choose English as the language of interview (Table 4).

Table 4. Distribution of Outcome Variables by Language of Administration by Propensity Strata

Had/diagnosed with emotional, nervous or psychiatric problems (expected higher reports in English)							
Stratum	English			Spanish			Difference
	n	Percent	SE	n	Percent	SE	
1	11	0.0%	0.0%	115	4.4%	1.9%	-4.4%
2	30	0.0%	0.0%	97	2.1%	1.4%	-2.1%
3	51	5.9%	3.3%	75	5.3%	2.6%	0.6%
4	68	8.8%	3.4%	59	1.7%	1.7%	7.1%
5	101	0.0%	0.0%	25	12.0%	6.5%	-12.0%
Total † $\chi^2_1=0.15$	261	1.4%	0.5%	371	2.4%	0.6%	-1.0%
Felt Depressed, Sad (expected higher reports in English)							
Stratum	English			Spanish			Difference
	n	Percent	SE	n	Percent	SE	
1	11	0.0%	0.0%	115	13.9%	3.2%	-13.9%
2	30	10.0%	5.5%	97	19.6%	4.0%	-9.6%
3	51	11.8%	4.5%	75	20.0%	4.6%	-8.2%
4	68	17.7%	4.6%	59	11.9%	4.2%	5.8%
5	101	17.8%	3.8%	25	36.0%	9.6%	-18.2%
Total † $\chi^2_1=0.89$	261	6.2%	0.9%	371	10.4%	1.2%	-4.3%
At Least One Diagnosed Health problem (expected higher reports in English)							
Stratum	English			Spanish			Difference
	n	Percent	SE	n	Percent	SE	
1	11	0.0%	0.0%	115	24.4%	4.0%	-24.4%
2	30	13.3%	6.2%	97	21.7%	4.2%	-8.3%
3	51	23.5%	5.9%	75	16.0%	4.2%	7.5%
4	68	26.5%	5.4%	59	13.6%	4.5%	12.9%
5	101	15.8%	3.6%	25	40.0%	9.8%	-24.2%
Total † $\chi^2_1=0.43$	261	7.9%	1.1%	371	12.5%	1.3%	-4.6%
Consume Alcoholic Beverages such as wine, beer and liquor (expected higher reports in English)							
Stratum	English			Spanish			Difference
	n	Percent	SE	n	Percent	SE	
1	11	55%	15%	115	37%	5%	17%
2	30	53%	9%	97	45%	5%	8%
3	51	69%	6%	75	39%	6%	30%
4	68	46%	6%	59	41%	6%	5%
5	101	54%	5%	25	56%	10%	-2%
Total † $\chi^2_1=10.78^{***}$	261	22.6%	1.6%	371	24.4%	1.7%	-1.7%

*** p<0.001 ** p<0.05

† Unweighted totals, standard errors reflect stratification by propensity

Language was not a significant predictor for reports of diagnosed health and psychiatric problems. Interestingly, language effects varied in both magnitude and direction across strata (Table 4) – a meek suggestion that language may have a differential impact on survey answers depending on the respondents' propensities to choose English over their mother-tongue and a likely explanation for the nonsignificant results.

It is possible that a larger effect of language would emerge if we could control on covariates that explain some of the variation in the outcome measures of interest. Table 5 presents the results of logistic regression models where diagnosed emotional and psychiatric problems, depression, diagnosed health problems and alcohol consumption were regressed on major correlates of these conditions among Hispanics and language of survey interview (for comparison purposes, the second column in Table 5 presents the results of regression on language only). Indeed, the coefficient for language of interview was slightly larger (0.63 versus 0.54) after controlling for the major predictors of alcohol use among Hispanics, namely, gender, depression and emotional problems (Caetano, 1987). However, controlling for the main correlates of diagnosed health, emotional problems and depression among Hispanics, namely, ethnic group membership (Alva, 1995; Portes, 1992; Collins et al., 2002; Harris, Edlund and Larson, 2005), English language proficiency (Padilla, 1986; Padilla, Cervantes and Maldonado, 1988; Alva, 1991; Brach, Fraser and Paez, 2005) and gender (Caetano, 1987; Mazzone, Boiko, Katon and Russo,

2007), the coefficients for language did not reach statistical significance. Thus, trying to find expected effects by controlling on confounding covariates did not increase support for the hypothesized language effects.

Table 5. Coefficients for Logistic Regression Models for Being Diagnosed with Emotional and Health Problems, Being Depressed and Alcohol Consumption

Outcome Measure		Language Only Model		Language and Covariates Model	
		Coefficient	SE [†]	Coefficient	SE [†]
Emotional, Nervous or Psychiatric Problems	Intercept	-3.17***	0.26	-1.77	0.92
	English	-0.17	0.43	-0.49	0.47
	Hispanic			-1.15	0.77
	Male			-0.74	0.45
	Current language at home: Not English			-0.59	0.52
	Language with friends: Not English			0.15	0.60
	Well Understands Spoken English			0.18	0.60
	Well Speaks English			0.51	0.60
Felt Sad, Depressed	Intercept	-1.53***	0.14	-0.23	0.58
	English	-0.21	0.22	-0.38	0.23
	Hispanic			-0.68	0.53
	Male			-0.71***	0.22
	Current language at home: Not English			0.074	0.25
	Language with friends: Not English			-0.62**	0.29
	Well Understands Spoken English			-0.50	0.32
	Well Speaks English			0.42	0.31
Health Problems	Intercept	1.24***	0.12	-1.087	0.68
	English	0.10	0.20	-0.16	0.22
	Hispanic			0.22	0.66
	Male			-0.96***	0.21
	Current language at home: Not English			-0.15	0.22
	Language with friends: Not English			0.27	0.25
	Well Understands Spoken English			0.10	0.30
	Well Speaks English			-0.10	0.31
Drinks Alcoholic Beverages	Intercept	-0.34***	0.11	-1.054***	0.16
	English	0.54***	0.16	0.63***	0.17
	Male			1.0069***	0.17
	Felt Sad, depressed			0.64***	0.23
	Have emotional problems			-0.079	0.46

*** p<0.001 ** p<0.05

[†]Standard errors reflect stratification by propensity strata

We also hypothesized what types of questions should be immune to the expected language effects. As predicted, language of interview did not affect responses to well established and readily accessible facts such as marital status and living situation (see Table 3). Surprisingly however, language was a significant predictor for number of biological children the respondent had given birth to or fathered. The mean for those interviewed in English was 1.13 (0.08)¹¹, while it was significantly higher for those interviewed in Spanish – 1.38 (0.07). A possible post-hoc explanation of this difference brings up the value of familism in the Hispanic culture and the possibility that it is socially desirable to report having a large family. Thus, the question related to number of biological children, even though initially considered simply factual, may hold different affective characteristics in the Hispanic and American cultures. An alternative hypothesis is that sample persons with large families disproportionately chose to be interviewed in Spanish, in a way not reflected in the propensity model.

4. Discussion and Conclusions

The results from the NIS data failed to demonstrate strong support for the effect of language of survey administration on the question answering process. We found partial support for the theory that language would influence responses to questions that differ in affective characteristics

¹¹ Standard error, reflecting stratification by propensity strata.

across the two cultures. Specifically, we found support for the hypothesis that Hispanic bilingual bicultural respondents interviewed in Spanish will report lower rates of alcohol consumption relative to those interviewed in English. However, we did not find significant effect of language on questions related to mental and health problems. Even though the topic of the questions should have induced cultural frame switching, it is possible that the question wording and/or translational issues prevented us from detecting such effects. It is also possible that the social desirability effects are not that different across the two cultures. Furthermore, it is unclear whether bilingual interviewers were able and allowed to go back and forth between language versions of the survey, possibly diminishing the effect of language priming a particular mind frame within which respondents interpreted the questions, accessed relevant information and provided a response consistent with their activated cultural identity.

We found support for the hypothesis that language will not affect all types of survey questions and specifically, that responses to questions related to well established and readily accessible autobiographical facts, such as marital status or living conditions, will not be influenced by language of interview. The significant effect of language on reports of number of children was surprising, but could possibly be related to the central value of familism in the Hispanic culture; thus, induced by the socially desirable standard of having a large family. An alternative explanation for the unexpected result, however, is that the propensity strata

were not sufficiently homogenous with respect to unobserved confounding covariates (e.g., family size, presence of family members during the interview). Despite their promise in addressing causal questions in the absence of randomization, propensity score methods can only adjust for observed confounding covariates; therefore, any unobserved covariates may challenge the interpretation of results. Thus, an examination of the effects of language on survey responding should also be conducted on the basis of random assignment. Chapter IV attempts to achieve this goal.

The results of this investigation have implications for surveys of immigrants and ethnic minorities – they suggest that leaving the choice of language to the bilingual respondent may not be a good practice. Ideally, researchers would be able to inform language assignment based on knowledge about domains where cultural differences and the direction of such differences may be expected, or depending on what respondent cultural identity is of interest. If such knowledge is not available, random assignment of bilingual respondents to a language would at least allow estimation of language effects.

This study was a necessary first step in examining the effect of language on survey responding; however, it reveals many additional avenues for research. First, it is important to simultaneously examine language and interviewer effects (for example, observable characteristics such as ethnic group membership, accent, gender). It can be hypothesized that when interviewer physical characteristics and accent do not match the

physical characteristics and accent of the culture associated with the language of interview, language effects may be dampened. Second, it is essential to explore how language can be used to reduce measurement error in recall – several studies have demonstrated that the language in which mental activity is carried out during information encoding creates an internal context analogous to a mental state and can serve as a retrieval cue during information recall. Similarly, the language spoken aloud during an event creates an external context analogous to a physical context and can serve as a situational cue during event recall (Schrauf and Rubin, 1998; Schrauf and Rubin, 2000; Marian and Neisser, 2000). This implies that language may facilitate the recall of information from life-periods or domains in which the same language is spoken. Third, language may be used to manipulate response styles associated with particular cultures. For example, there is a popular belief in the cross-cultural survey world that Asian respondents avoid extreme responses (Zax and Takahashi, 1967; Chun et al., 1974; Stening and Everett, 1984; Hayashi, 1992). By evoking a different cultural mind set in bilingual respondents, we may be able to attenuate response bias. Finally, to fully understand the effect of language on survey responding, it is important to disentangle the various mechanisms that may produce such effects – for example, cultural-frame switching from language dependent recall (see Chapter II for detailed discussion on these influences).

5. Appendix

Table 6. Distributions of Correlates of Language Selection and Demographic Characteristics by Propensity Strata

Covariate	Stratum 1		Stratum 2		Stratum 3		Stratum 4		Stratum 5	
	English	Spanish								
New Arrival	18%	9%	0%	10%	4%	3%	10%	10%	6%	4%
Adjusted Immigrant	82	91	100	90	96	97	90	90	94	96
<i>Visa</i>										
US citizen spouse	0	13	37	24	39	28	28	41	26	32
Employment	0	3	13	7	14	12	13	17	13	16
Diversity	0	0	0	0	0	0	1	2	1	0
Other	100	83	50	69	47	60	57	41	60	52
Male	73	57	63	66	53	51	52	53	50	56
Female	27	43	37	34	47	49	48	47	50	44
<i>Enrolled in English Classes</i>										
Yes	18	11	7	11	20	16	12	12	8	12
No	82	89	93	89	80	84	88	88	92	88
<i>Language at home English</i>										
Yes	27	24	53	43	69	68	79	78	82	92
No	73	76	47	57%	31	32	21	22	18	8
<i>Language at work English</i>										
Yes	64	68	90	73	67	71	72	76	69	76
No	36	32	10	27	33	29	28	24	31	24
<i>Language with friends English</i>										
Yes	27	32	57	64	80	80	94	88	98	92
No	73	68	43	36	20	20	6	12	2	8
<i>Country of Birth</i>										
Other Latin American	91	77	53	61	49	51	50	46	25	32
Mexico	9	22	47	39	51	49	50	54	75	68
<i>Spanish-speaking childhood country</i>										
Not Spanish-speaking	100	97	97	100	88	89	74	80	42	40
	0	3	3%	0	12	11	26	20	58	60
Rural	27	40	40	42	27	24	16	19	23	16
Not Rural	73	60	60	58	73	76	84	81	77	84
<i>Worked in the U.S. since Move</i>										
Yes	82	70	97	87	84	84	78	85	91	96
No	18	30	3	13	16	16	22	15	9	4
<i>Very well understands spoken English</i>										
Well understands spoken English	27	25	23	32	55	48	63	66	87	80
Very well speaks English	73	75	77	68	45	52	37	34	13	20
Well speaks English	0	4	0	8	37	29	60	58	79	76
	100	96	100	92	63	71	40	42	21	24
<i>Took English classes before coming to the U.S.</i>										
Did not Take English classes	18	27	33	48	55	48	50	53	33	24
Church member	82	73	67	52	45	52	50	47	67	76
Not church member	27	17	37	31	27	36	37	34	27	28
	73	83	63	69	73	64	63	66	73	72
<i>Plans to travel home next 12 months</i>										
No plans to travel	64	54	43	57	67	52	62	59	42	56
	36	46	60	40	33	48	38	41	58	44
<i>Face-to-face interview</i>										
Telephone interview	73	52	63	72	63	68	56	61	57	44
	27	48	37	28	37	32	44	39	43	56

Covariate	Stratum 1		Stratum 2		Stratum 3		Stratum 4		Stratum 5	
	English	Spanish								
<i>Employment</i>										
Working	73%	88%	93%	79%	75%	75%	76	78%	72%	80%
Temporarily out of work	9	3	3	10	12	15	12	8	14	12
Retired/disabled	0	2	0	1	0	1	3	0	0	0
Homemaker	18	6	3	8	12	4	3	10	10	8
Student	0	2	0	1	2	5	6	3	4	0
<i>Age</i>										
59+	0	3	0	1	2	0	0	0	0	0
49-58	0	1	3	3	4	3	6	3	7	12
44-48	18	9	7	5	0	4	4	5	0	0
39-43	27	18	10	12	14	11	7	8	2	8
34-38	18	20	37	31	18	16	13	19	5	4
29-33	27	31	27	25	25	28	19	17	10	16
24-28	9	10	13	13	18	16	21	24	28	24
18-23	0	7	3	9	20	23	29	24	49	36
Means (standard error) for continuous variables										
Years in school	13 (1.3)	11 (0.3)	13 (0.8)	13 (0.5)	13 (0.5)	13 (0.4)	13 (0.4)	14 (0.4)	13 (0.3)	15 (0.7)
Years in U.S. school	0.5(0.2)	1 (0.1)	1 (0.5)	1 (0.2)	2 (0.4)	3 (0.4)	5 (0.5)	4 (0.6)	9 (0.4)	9 (0.9)

Questions where Language effects are Expected

General Health Questions

D4 Has a doctor ever told you that you have high blood pressure or hypertension?

(Interviewer Instruction: It must be a diagnosed condition.)

- 1. YES
 - 2. NO
 - 2. DK
 - 1. RF
-

D8 Has a doctor ever told you that you have diabetes or high blood sugar?

- 1. YES
 - 2. NO
 - 2. DK
 - 1. RF
-

D13 Has a doctor ever told you that you have cancer or a malignant tumor, excluding minor skin cancers?

- 1. YES
 - 2. NO
 - 2. DK
 - 1. RF
-

D19 Has a doctor ever told you that you have chronic lung disease such as chronic bronchitis or emphysema?

[INTERVIEWER INSTRUCTION: DO NOT INCLUDE ASTHMA]

- 1. YES
 - 2. NO
 - 2. DK
 - 1. RF
-

D24 Has a doctor ever told you that you had a heart attack, coronary heart disease, angina, congestive heart failure, or other heart problems? [INTERVIEWER INSTRUCTION: READ SLOWLY]

- 1. YES
 - 2. NO
 - 2. DK
 - 1. RF
-

D43 Has a doctor ever told you that you had a stroke?

- 1. YES
 - 2. NO
 - 2. DK
 - 1. RF
-

D55 Have you ever had, or has a doctor ever told you that you have arthritis or rheumatism?

- 1. YES [D56]
 - 2. NO [D59]
 - 2. DK [D59]
 - 1. RF [D59]
-

D59 Have you ever had, or has a doctor ever told you that you have asthma?

- 1. YES
 - 2. NO
 - 2. DK
 - 1. RF
-

D67 Are you often troubled with pain?

- 1. YES
 - 2. NO
 - 2. DK
 - 1. RF
-

(Continued)

Questions where Language effects are Expected

Mental Health Questions

D50 Have you ever had or has a doctor ever told you that you have any emotional, nervous, or psychiatric problems?

- 1. YES
 - 2. NO]
 - 2. DK
 - 1. RF
-

D88 During the past 12 months, was there ever a time when you felt sad, blue, or depressed for two weeks or more in a row?

- 1. YES
 - 2. NO
 - 3. [VOL] DID NOT FEEL DEPRESSED BECAUSE ON ANTI-DEPRESSANT MEDICATION
 - 2. DK
 - 1. RF
-

Alcohol Consumption

D81 Do you ever drink any alcoholic beverages such as beer, wine, or liquor?
(INTERVIEWER INSTRUCTION: REFERS TO CURRENT ALCOHOL CONSUMPTION)

- 1. YES
 - 2. NO
 - 2. DK
 - 1. RF
-

Questions where Language Effects are Not Expected

A10. Are you currently living by yourself or with other people? By living with other people I mean that there are, in your current residence, family members, housemates or roommates, foster children, roomers, boarders, or live-in employees, or persons who are temporarily away on a business trip, on vacation, or in a general hospital. (QxQ: DO NOT INCLUDE: college students living away while attending college; people in a correctional facility, nursing home, or mental hospital on the date of the interview; Armed Forces personnel living somewhere else; or people who live or stay at another place most of the time;

DO INCLUDE: people staying here on date of interview who have no other permanent place to stay; people living here most of the time while working, even if they have another place to live; college students who stay here while attending college; persons in the Armed Forces who live here; newborn babies still in the hospital; children in boarding schools below the college level)

- ALONE..... 1
 - WITH OTHER PEOPLE..... 2
 - REFUSED.....-1
 - DON'T KNOW.....-2
-

A52. Are you now: [IWER: IF R IS MARRIED **AND ALSO** LIVING TOGETHER WITH SOMEONE ELSE IN A MARRIAGE-LIKE RELATIONSHIP, CODE "MARRIED" HERE.]

- Married..... 1
 - Living together in a marriage-like relationship but not married 2
 - Separated.....3
 - Divorced.....4
 - Widowed.....5
 - Never married,not living with someone in a marriage like relationship 6
 - REFUSED.....-1
 - DON'T KNOW.....-2
-

A232. Now, we would like to ask about births of children. How many children [IF A6=1 / have you ever fathered; IF A6=2 / have you yourself ever given birth to; IF A6=-1,-2 / have you ever fathered or have you yourself ever given birth to]? DO NOT COUNT STILLBIRTHS, STEPCHILDREN, OR CHILDREN RESPONDENT HAS ADOPTED

- ENTER NUMBER: ## [SOFT RANGE CHECK = 0 TO 35]
 - REFUSED.....
 - DON'T KNOW.....
-

Figure 3. Outcome Measures of Interest from the NIS Questionnaire

Chapter IV. Does Language Matter: An Investigation of the Effect of Language Assignment on Answering Behavior among Hispanic Bilingual Survey Respondents

1. Introduction

Instrument translation that ensures measurement equivalence and concept relevance across ethnic groups has long been of importance to cross-cultural research (for example, Canino and Bravo, 1994; Harkness, 2003; Allegria et al., 2004). However, recent work in psycholinguistics suggests there is another aspect of language that deserves attention – its interrelation to thought processes.

Various mechanisms through which language influences cognition have been identified in psychology, psycholinguistics and linguistics. That growing body of literature suggests that language can influence cognition through its syntactic properties (for example, Briere and Lanktree, 1983; Ng, 1990; Khosroshashi, 1989), grammatical structure (for example, Bickel, 1997; Widlock, 1997; Majid et al., 2004;) and semantic categories (for example, Hunt and Agnoli, 1991). More importantly for the present research, language can serve as a powerful cue that primes the associated cultural meaning system in bicultural bilingual respondents (e.g., Trafimow et al., 1997; Ross et al., 2002) and can facilitate recall when the initial information encoding was experienced in the same language (e.g.,

Schrauf and Rubin, 1998; Marian and Neisser, 2000; Schrauf and Rubin, 2000). Specifically, Ross et al., (2002) used language to shift the cultural frame adopted by Chinese-born students in Canada in a task of self and mood description. As expected, those interviewed in Chinese provided more collective statements and reported higher endorsement of Chinese views. These respondents also reported equal levels of positive and negative moods, whereas participants who were asked in English, reported a preponderance of positive moods.

Language dependent recall has been demonstrated in several bilingual groups in terms of number of recalled memories and time in life when the recalled events took place. Bugelski (1977) found that when Spanish-English bilingual immigrants were cued with Spanish words, 45% of their thoughts were related to their childhood, but when cued with English words, only 30% of their thoughts were related to pre-immigration events. Similarly, Schrauf and Rubin (1998) found that elderly Spanish-English bilingual immigrants reported more memories from ages 5-20 in response to Spanish rather than English prompts, but more memories from ages 20-55 in response to English rather than Spanish prompts.

The effect of language of administration on survey responding remains largely unexplored. To our knowledge, there are only two published survey-related articles. Botha (1968) examined the effect of language on attitudes among bilingual students in Lebanon and found a significant difference only for students tested in Arabic and French, but not

for students tested in Arabic and English. Feldman (1975) examined the effects of language (English or Gussi) and national origin of the administrator (American or Gussi) on traditional and modern attitudes among students in Kenya in a fully crossed design. He found no effect for administrator, but an interaction effect of language, school level and gender – female upper-level secondary students who answered the questionnaire in English expressed more modern views than those who answered in Gussi. Currently in press, an article by the principal investigators of the National Latino and Asian American study compared the English and Spanish versions of the World Mental Health Survey Composite International Diagnostic Interview among respondents randomly assigned to one of two languages (Shrout et al., 2008). The authors found no differences in rates for 10 of the 11 diagnoses between the two language groups, but stressed the need for randomized language experiments in health surveys of multilingual populations as a step allowing tests of the translational equivalence of the instruments and comparability of prevalence rates.

In Chapter III we reported language effects in reports of alcohol consumption among Spanish-English bilingual immigrants and the surprising finding that reports of number of biological children were language dependent – when asked in Spanish, respondents were 44% more likely to report one more child relative to when asked in English. Chapter III demonstrated one of two approaches to examining language

effects. Because in most surveys bilingual respondents self-select themselves into a given language version of the survey instrument, it is necessary to employ propensity score methods to “correct” for the absence of random assignment. Despite its promise, this approach is sensitive to propensity model specification and the omission of unobserved covariates. In order to build confidence in the causal conclusions, replication and alternative approaches are needed.

The ideal approach to studying language as a cause of measurement differences across language versions of a survey instrument is to randomly assign bilingual respondents to language of administration. Such experimental data are hard to locate as often random assignment to a language is not a feasible option due to cost constraints and/or inability to identify respondents who speak and understand both languages equally well. The National Latino and Asian American Study (NLAAS) presents a unique opportunity – respondents who self-identified themselves as Spanish-English bilinguals were not given a choice of language, but were randomly assigned to a language of interview. In addition, the questionnaire covered similar topics to the NIS, used in Chapter III, allowing us to attempt to replicate some of the results through an experimental design.

As discussed in Chapter III, general and mental health issues and alcohol consumption are associated with different stigma in the American and Hispanic cultures (e.g., Silva de Crane and Spielberger, 1981; Caetano

and Clark, 1999). Specifically, recent immigrants, despite poverty and more difficult access to health care, report better physical and mental health than their more acculturated or U.S.-born counterparts (Vega et al., 1998; Ortega et al., 2000; Grant et al., 2004). Furthermore, Hispanics have been reported to have more conservative alcohol norms and attitudes than Whites (Caetano and Clark 1999) and report lower levels of alcohol consumption (Dawson et al., 1995). Thus, similar to the hypotheses in Chapter III, we would expect bilingual bicultural respondents to provide more positive evaluations of their physical and mental health when the interview is in Spanish than in English. Likewise, we would expect to see reports of first alcohol consumption to be associated with older age when respondents are interviewed in Spanish rather than English.

Additionally, NLAAS contains rich measures of respondents' family values and importance of family, allowing us to test additional hypotheses related to *familism*. The importance and strong connection to family is a central value in the Latino culture (Lisansky, 1981; Triandis et al., 1984; Sabogal et al., 1987; Marin and Marin, 1991; Vasquez, 1994; Levine et al., 2001; De Las Fuentes et al., 2003). Analogous to findings by Trafimow et al. (1997) and Ross et al. (2002) where language is used to shift the cultural frame adopted by respondents, it can be expected that when bilingual Hispanic respondents are primed with Spanish language, they will exhibit a higher tendency to endorse Hispanic values relative to when asked in English. Thus, we would expect lower scores on any of the familism scales

in the NLAAS for respondents interviewed in Spanish relative to those interviewed in English (where lower scores represent higher levels of family pride, cohesion and lower levels of family cultural conflict).

As in Chapter III, we would not expect any language effects on demographic questions that ask for highly accessible and well established facts, such as household size, marital status, or the respondent's number of biological children. Surprisingly, the latter yielded significant differences between the English and Spanish language administrations of NIS, so this investigation will give us a chance to see if the result will replicate in an experimental setting.

2. Data and Methods

The NLAAS is part of the national Collaborative Psychiatric Epidemiological Studies (along with the National Co-morbidity Study-Replication and the National Survey of American Life), designed to collect psychiatric epidemiological data on various populations. It was administered to a sample of non-institutionalized Latino and Asian American adults ages 18 or older, residing in households in the conterminous U.S. The sample targeted four Latino populations (Cuban, Mexican, Puerto Rican and other adults of Latino descent), four Asian populations (Chinese, Filipino, Vietnamese and other adults of Asian descent) and a small control group of white non-Hispanic, non-Asian respondents. The Latino-American sample

selection followed a four-step process, selecting: 1) U.S. Census Metropolitan Statistical Areas and counties, 2) area segments, 3) housing units, and 4) eligible respondents from the selected housing units (for detailed outline of the sampling design see Heeringa et al., 2004). The interview was administered using computer assisted personal interview methods at the respondents' homes. The survey yielded 2,554 interviews with Latino respondents (75.5% weighted response rate), 2,095 interviews with Asian respondents (65.6% weighted response rate) and 215 control interviews with white non-Hispanic, non-Asian respondents, that were discontinued after the first months of the field period for cost reasons.

The NLAAS questionnaire measures psychiatric illness, service use and impairment, as well as neighborhood safety, social cohesion, migration status and history, discrimination, familism, acculturation, language proficiency, gender roles, social ties and others. In recognition that participation of minorities who do not speak English was very desirable, the survey instrument was translated into four languages – Spanish, Chinese, Vietnamese and Tagalog. Special effort during questionnaire development, translation and adaptation was placed on attaining cultural relevance as well as measurement equivalence across languages. During this process, three goals were pursued: a) attaining cultural relevance by formulating the research problem with attention to cultural and contextual differences of Latinos and Asians; b) achieving cultural equivalence in standardized instruments, and c) achieving measures that do not fluctuate depending on

culture and translation (Allegria et al., 2004). The Spanish adaptation and translation was a deliberate process that integrated both the emic perspective (starts with concepts within a culture and works out associations and meanings within the cultural frame) and the etic perspective (evaluates a phenomenon using culture-free and objective constructs) for conducting cross-cultural research (for applications of this approach in mental health research, see Canino, Lewis-Fernandez and Bravo, 1997; Lopez and Guarnaccia, 2000). The instrument was sent for professional translation and back translation and then was reviewed by a multinational bilingual committee that had to evaluate whether the translation was culturally relevant for different Latino groups. Twelve focus groups were conducted with Spanish-speaking respondents from various cultural backgrounds to ensure the instrument was adequately translated and adapted. The process sought to create a Spanish instrument that maintained semantic, content and technical equivalence to the English instrument (for a complete description of the steps involved, see Allegria et al., 2004).

For most respondents, the survey instrument was administered in different languages by bilingual interviewers. Before the beginning of the survey interview, respondents from the Hispanic sample were asked to report on whether they speak only Spanish, mostly Spanish, Spanish and English about the same, mostly English, or only English. Those who reported speaking Spanish and English about the same were randomly

assigned to the Spanish or English version of the instrument by the interviewer's computer. Interviewers were instructed to maintain the interview in the language of the random assignment, unless problems related to question understanding or refusal to conduct the interview in the assigned language occurred. In such cases interviewers were instructed to make an explicit note and allowed to switch languages. However, there were no such cases reported (Shrout et al., 2008). Thus, the subsample of 332 bilingual respondents is the casebase for the current investigation of the effect of language on survey responding.

We selected questions that measured concepts similar to the ones discussed in Chapter III; namely, general and mental health self-evaluation, report on age of first alcohol consumption, marital status, number of biological children (see Figure 4, Appendix). However, the different operationalizations of these concepts do not allow direct comparison of results. In addition, we looked at measures of familism – a central value in the Latino cultural system. NLAAS contains subscales on family pride and family cohesion, taken from the Family Environment Scale (Olson, 1986; Olson, 1989) that are congruent with the values of familism reported in Latino cultures (Vega, 1990; Ortiz, 1995) and used in many studies of various Latino subgroups (Vega et al., 1993; Gil et al., 1994; Gil and Vega, 1996). The scale on family pride contains measures of trust and respect between family members, shared values, loyalty to the family, family pride and general family orientation and functioning. The scale on family

cohesion contains measures on family closeness and togetherness and importance of time spent together. Questions in both scales have four response options, ranging from strongly agree to strongly disagree. In addition, a family cultural conflict scale that addresses issues of cultural and intergenerational conflict between respondents and family members was examined. The family cultural conflict scale is a subscale of the Hispanic Stress Inventory (Cervantes, Padilla and Salgado de Snyder, 1991) and measures frequency of incidents of cultural conflict between respondents and family members, interference with personal goals, arguments due to different belief systems, breakdown of family unity and comparison of respondent's family closeness to friends' families. The three response categories are "hardly ever", "sometimes", and "often".

3. Results

3.1 Hypotheses Tests

Table 7 and Table 8 present the results of language comparisons across measures of physical and mental health, age of first alcohol consumption, family pride, cohesion and conflict. These are measures where we expected the effect of language to be the strongest. Specifically, we expected to find significantly lower reports of excellent physical and mental health, higher rates of depression, higher scores on the familism scales (where higher scores represent lower levels of family pride, cohesion

and higher levels of family cultural conflict) and lower age of self-report of first alcohol consumption when respondents were assigned to the English version of the questionnaire. In addition, we looked at reports on marital status, household size and number of biological children, where no differences between the English and Spanish administrations were expected. The first columns in Table 7 and Table 8 present the overall comparison between English and Spanish¹², followed by analyses on a subset of respondents born outside of the U.S. The results are rather surprising – the probability of reporting excellent mental and physical health or depression was not different for those assigned to English language administration relative to those assigned to Spanish. Similarly, the mean differences for reported age of first alcohol consumption and scores on the familism scales were not significant across the two language groups at the traditional $p < 0.05$ level of statistical significance. The results remained unchanged when the respondent casebase was subset to immigrants.

¹² The two randomly assigned language groups were balanced on gender, age, education, work status, age at immigration, language spoken with friends, language spoken with family members, language in which respondent thinks, self-evaluation of ability to read, write and speak English.

Table 7. Relative Risk for the English Language Assignment Group across All Bilinguals and Immigrants

Outcome Measure		All Bilinguals (n=332)			Born Outside the U.S. Bilinguals (n=176)		
		Mantel-Haenszel Relative Risk for English	95% Confidence Interval		Mantel-Haenszel Relative Risk for English	95% Confidence Interval	
Language Effects Expected							
Physical Health	Excellent vs. Very Good	0.94	0.72	1.24	0.74	0.47	1.16
	Excellent vs. Good	0.91	0.69	1.19	0.78	0.49	1.25
	Excellent vs. Fair or Poor	1.20	0.83	1.74	0.88	0.49	1.59
Mental Health	Excellent vs. Very Good	0.96	0.78	1.20	0.83	0.62	1.12
	Excellent vs. Good	1.16	0.87	1.53	1.77	0.98	3.20
	Excellent vs. Fair or Poor	1.45	0.86	2.47	1.41	0.63	3.18
Sad, Depressed for several days		0.93	0.76	1.13	0.93	0.70	1.26
Language Effects not Expected							
Marital Status	Married vs. Divorced/Widowed	1.12	0.82	1.53	1.43	0.87	2.34
	Married vs. Never married	0.81**	0.66	1.00	0.79	0.58	1.09

** p<0.05

Table 8. Means for the English and Spanish Language Assignment Groups across All Bilinguals and Immigrants

Outcome Measure	All Bilinguals (n=332)					Born Outside the U.S. Bilinguals (n=176)						
	English		Spanish		Mean Difference SE	English		Spanish		Mean Difference SE		
Language Effects Expected	Mean	SE	Mean	SE		Mean	SE	Mean	SE			
Age first drank alcohol	15.71	0.40	16.33	0.39	-0.62	0.57	16.27	0.56	16.68	0.56	-0.41	0.80
Family pride	10.22	0.29	9.46	0.29	0.77	0.41	9.20	0.32	9.21	0.36	-0.002	0.48
Family cohesion	4.46	0.14	4.11	0.14	0.35	0.20	4.08	0.16	4.07	0.19	0.01	0.25
Family conflict	6.77	0.17	6.34	0.16	0.42	0.24	6.70	0.23	6.21	0.18	0.50	0.29
Language Effects Not Expected												
Number of biological children	1.53	0.11	1.91	0.12	-0.38**	0.16	1.36	0.13	2.00	0.16	-0.64**	0.21
Household size	2.82	0.11	2.81	0.13	0.01	0.17	2.77	0.17	2.85	0.17	-0.08	0.24

** p<0.05

An important factor to consider in the analyses of immigrants is age at immigration. Age at immigration may play a significant role in the acculturation processes – respondents who learned the second language as children in the same context in which they learned their first language are more likely to have one cognitive representation for a concept and its translation equivalent than others. Conversely, respondents who learned the second language in a context different from the context of their mother-tongue are likely to have separate representations for two translation equivalents (Ervin and Osgood, 1954; Lambert, Havelka and Crosby, 1958). Such hypotheses are supported by neurological evidence that shows that in bilinguals who acquired their second language early in childhood, the two languages are represented in common frontal cortical

areas, while in late bilinguals, the two languages are spatially separated in the cortex (Kim, Relkin, Lee and Hirsch, 1997). A study by Chee et al., (1999) used fMRI to demonstrate that Singapore Chinese bilinguals who learned both English and Mandarin before age 6 used common neuroanatomical regions during conceptual and syntactic processing of visually presented sentences, regardless of testing language. Such findings suggest, that not all immigrants would be equally likely to undergo cultural frame switching when primed with a language – those who have one cognitive representation will not exhibit any language effects. Thus, one additional constraint was imposed on the respondent casebase and we repeated the analyses presented in Table 7 and Table 8 with respondents who came to the U.S. as teenagers or older (n=76). Surprisingly, again there were no significant differences between the two language groups (see Table 12 and Table 13, Appendix).

Table 7 and Table 8 also present the results of language comparisons across several factual questions where no language effects were expected. Interestingly, we found lower probability for respondents interviewed in English to be married rather than be single relative to those interviewed in Spanish, but the effect disappeared when restricted only to immigrants. Consistent with results reported in Chapter III, we also found significantly higher reports of biological children when respondents were interviewed in Spanish rather than in English.

Finding differences where none were expected and failing to find them where expected, argues against the original hypotheses. However, it may also be the result of failure in the experimental design, such as noncompliance with the randomization. A final set of analyses addresses this possibility. Even though not officially documented, interviewers could have been making the decision to switch the originally assigned languages during the interview for unknown reasons. Similar problems often occur in clinical trials, where patients randomly assigned to a drug treatment fail to comply and take the prescribed dose. In such cases, “intent-to-treat” analysis (analysis where the compared groups are determined by the randomization procedure) is not reasonable and often the groups to be compared are determined by an algorithm based on the participants’ compliance during the trial (“as-treated” analysis). Analogously, we can treat the results presented so far as “intention-to-treat” analysis and check whether they are threatened by selection bias as a result of randomization breakdown.

3.2 As-treated Analyses and Propensity Model

We are likely to detect language effects when language induces a particular mind frame; however, it is unclear whether this can occur if languages are switched back and forth during the interview. Thus, we define noncompliance with the randomization as any case where at least one language switch occurred at some point during the interview. Such information is available from the key stroke files – in order to switch from

one language to another, interviewers had to press “CRTL+L” on their laptops every time. Examination of these files showed that thirty-three percent of the 332 bilingual respondents who were initially assigned to English or Spanish were switched to the other language at some point during the interview. The majority of language switches occurred only once, but there were 22 respondents who were switched several times between languages. Overall, 72% of the language switches occurred in the originally assigned Spanish language group. Thus, depending on compliance with the initial randomization, we formed four language groups – English or Spanish compliant groups (respondents who were interviewed in the initially assigned language) and English or Spanish non-compliant groups (respondents, initially assigned to English or Spanish, but were switched to the other language at some point during the interview). Possible bias from nonrandom language switch may be affecting the results reported in Table 7 and Table 8.

First, we examined whether the expected balance in the two initial random language assignments was affected (Table 14, Appendix). We compared demographic characteristics and correlates of acculturation related to the likelihood the respondent will or will not experience any problems with the initially assigned language and will remain in his/her group. Two factors, the country in which respondent received their education before the age of 16 and the language spoken with friends, varied significantly between the English compliant and Spanish compliant

groups, suggesting the results presented in Table 7 and Table 8 could be biased.

As in Chapter III, propensity score methods were used to overcome the breakdown of the experimental assignment to a language of interview (see Chapter III for an overview of the method). We decided to exclude respondents from the noncompliant groups from the investigation as it was impossible to detect the reason for the language switch – whether it was related to respondent difficulties with the initially assigned language, or whether it was interviewer related (especially in cases of multiple switches, where one possibility is that the bilingual interviewers toggled between language versions to see the translation equivalent with which they were more comfortable). Thus, we modeled the likelihood to be in the English compliant group rather than the Spanish group as a function of demographic characteristics and covariates of acculturation such as language spoken with friends and family, language in which the respondent thinks, self-evaluation of written English skills, how important is marriage within the same ethnicity and how close the respondent feels in ideas with people from the same ethnicity (Table 9). The pseudo R-square for the model was 0.10 and the model fit reasonably well, as indicated by the Hosmer and Lemeshow goodness of fit test ($\chi^2_8=5.89$). None of the predictors reached statistical significance at the traditional $p<0.05$ significance level.

To evaluate the adequacy of the propensity model for estimating the effects of language from these data, we examined the region of overlap of raw propensity scores for those who remained in the English compliant group and those who remained in the Spanish compliant group (Figure 5, Appendix). There was a large enough region of overlap to conclude we could sensibly estimate the effect of language in the compliant subsample.

Table 9. Logistic Regression Coefficients for Likelihood of Being in the English Language Compliant Group

Predictor	Coefficient	Standard Error
Intercept	1.10	1.32
Gender:		
Male	0.15	0.34
Female	-	-
Age	-0.011	0.014
Household income in thousands	0.00013	0.0034
Years of education	0.014	0.073
Work Status:		
Unemployed	-0.14	0.72
Not in the work force	0.58	0.42
Employed	-	-
At least one U.S. born parent:		
Yes	-0.034	0.41
No	-	-
Years in the U.S.:		
Less than 20	0.41	0.59
More than 20	-0.029	0.42
U.S. Born	-	-
Country received education before 16:		
U.S.	0.45	0.47
Other	-	-
Language with Friends:		
English	0.22	0.39
Spanish	-0.47	0.48
Both English and Spanish	-	-
Language with Family:		
English	0.020	0.50
Spanish	0.047	0.40
Both English and Spanish	-	-
Language in which Thinking:		
English	0.29	0.38
Spanish	0.36	0.51
Both English and Spanish	-	-
Written English:		
Poor	-1.54	1.35
Fair	0.78	0.62
Good	0.30	0.37
Excellent	-	-
Close in ideas with people from same ethnicity:		
Very	-1.016	0.72
Somewhat	-0.99	0.71
Not Close	-	-
Marriage within same ethnicity is important:		
Very	-0.77	0.52
Somewhat	-0.20	0.41
Not very	0.036	0.45
Not at all	-	-

The raw propensity scores were divided into quartiles (propensity strata) as shown in Table 15 (Appendix), with the first quartile corresponding to those with the lowest propensity to be in the English compliant group. The distributions of the covariates within each propensity stratum was not significantly different (based on Chi-square and t-tests) for the two language groups (Table 16, Appendix), indicating that the expected balance by the initial randomization across known factors was finally established.

3.3 Hypotheses Tests Revisited

Table 10 and Table 11 present analyses analogous to the ones presented in Table 7 and Table 8, accounting for stratification by propensity class¹³. With proper balancing of respondents across language groups, we observed some evidence that the language in which survey questions are asked can affect responses. As predicted, when asked in Spanish, respondents endorsed more the Hispanic value of familism, resulting in significantly lower scores on the family pride scale (Table 11). Even though differences between the two language groups on the rest of the familism scales did not reach significance, they were in the expected direction – respondents interviewed in Spanish had lower scores. Similarly, the reported age of first alcohol consumption was higher for the respondents interviewed in Spanish, consistent with more conservative alcohol norms

¹³ Because of sample size restrictions across contrast cells, we were unable to repeat the presented analyses only for the respondents born outside the United States. However, the propensity model ensured balance by immigration status.

and attitudes among Hispanics relative to Whites (Cateano and Clark, 1999), but failed to reach statistical significance.

As expected, language of interview did not affect responses to well-established and highly accessible facts, such as respondent's marital status, household size, and number of biological children.

Table 10. Relative Risk for the English Language Compliant Group versus Spanish across Measures with Different Language Effect Expectations Reflecting Stratification by Propensity Strata

Outcome Measure		Mantel-Haenszel Relative Risk for English	95% Confidence Interval [†]	
Language Effects Expected				
Physical Health	Excellent vs. Very Good	0.91	0.72	1.15
	Excellent vs. Good	0.90	0.71	1.15
	Excellent vs. Fair or Poor	1.09	0.73	1.60
Mental Health	Excellent vs. Very Good	1.04	0.86	1.27
	Excellent vs. Good	1.01	0.79	1.31
	Excellent vs. Fair or Poor	0.93	0.65	1.34
Sad, Depressed for several days		1.05	0.88	1.25
Language Effects not Expected				
Marital Status	Married vs. Divorced/Widowed	1.03	0.76	1.40
	Married vs. Never married	0.88	0.73	1.06

[†] Standard errors reflect stratification by propensity strata

Table 11. Means for the English and Spanish Language Compliant Groups across Measures with Different Language Effect Expectations Reflecting Stratification by Propensity Strata

Outcome Measure	English		Spanish		Mean Difference	SE [†]
	Mean	SE [†]	Mean	SE [†]		
Language Effects Expected						
Age first drank alcohol	15.54	0.42	16.69	0.61	-1.15	0.74
Family pride	10.46	0.32	9.25	0.37	1.21**	0.50
Family cohesion	4.48	0.15	4.11	0.2	0.37	0.26
Family cultural conflict	6.83	0.19	6.54	0.24	0.30	0.30
Language Effects Not Expected						
Number of biological children	1.49	0.11	1.88	0.17	-0.39	0.21
Household size	2.84	0.12	2.57	0.16	0.28	0.20

** p<0.05

[†] Standard errors reflect stratification by propensity strata

4. Discussion and Conclusions

Overall, the present analyses failed to identify strong effects of the language of interview on survey response. Consistent with findings in other fields, we expected to find big language effects on responses to questions that differed in affective characteristics across two cultures. Specifically, we predicted that bilingual bicultural respondents will provide more positive evaluations of their physical and mental health when the interview was in Spanish than in English. Likewise, we expected to see reports of first alcohol consumption to be associated with older age in life when respondents were interviewed in Spanish. Furthermore, we predicted that bilingual Hispanic respondents primed with Spanish language would exhibit a higher tendency to endorse Hispanic values that would manifest in lower scores on any of the familism subscales. Instead, the observed language effects were limited – we found support for the hypothesis that Hispanic bilingual bicultural respondents interviewed in Spanish have lower scores on a family pride scale relative to those interviewed in English. We also found support for the hypothesis that language will not affect all types of survey questions as responses to factual questions such as marital status, number of biological children, or number of household members do not exhibit language influences.

We found no significant effects of language on questions related to mental and physical health (consistent with findings reported in Chapter III) and the familism subscales related to family cohesion and conflict (even

though the observed differences were in the anticipated direction). A possible explanation for these null findings is the level of acculturation of the Hispanic sample. Only 23% of the total bilingual sample consisted of respondents who came to the U.S. in their teens or later. This percent became even smaller when restricting the sample to the compliant with the randomization process language groups, not allowing us to present analyses for the immigrants only. In addition, we were unable to differentiate among respondents born in different Latino countries (a significant correlate of acculturation in Chapter III) as country of birth was dichotomized to “U.S.” and “Other” in the data.

We also failed to replicate the significant language effect related to alcohol consumption reported in Chapter III. This can be related to the parameterization of the concept – while in the NIS respondents were asked to report current alcohol consumption, in the NLAAS they were asked to recall their age of first alcohol consumption. To the extent to which language dependent recall plays an important role in survey responding, the reports may be influenced by the match between language spoken during the event and language of interview in addition to social desirability. We do not have sufficient information to test such a hypothesis and predict the direction of the effect given interaction between cultural frame switching and language dependent recall.

Another unexplored possibility for the observed null findings is the actual translation of the examined questions – despite the deliberate effort

to ensure cultural relevance, semantic, context and technical equivalence, it is possible that the wording of the questions of interest in Spanish played down the social desirability effect of the topics, suppressing language influences that would otherwise have been visible. Yet another possibility for the nonsignificant results is that different mechanisms affected (even simultaneously) the examined questions. We are unable to explore further the likely influences with these data, but future work will focus on disentangling potential causes of language effects across various question types and possibly, across stages of the response formation process.

In sum, the results of this investigation suggest that the language in which survey questions are asked may pose less of a threat to data quality in surveys of immigrants and ethnic minorities. However, before we can comfortably accept such a conclusion, we need a carefully designed and more detailed investigation of the mechanisms that induce different answers across languages. The implication of such results (along with the findings reported in Chapter III) for current national surveys that sample ethnic minorities and immigrants is that language assignment should be informed by the goals of the survey questions and leaving the choice of language to a bilingual bicultural respondent may affect data quality.

Other factors may play a role in this process and deserve further examination – for example, it is important to understand the interplay between language of survey administration and interviewers' observable cultural characteristics. It can be speculated that when interviewer physical

characteristics and accent do not match the physical characteristics and accent of the culture associated with the language of interview, language effects may be dampened. Furthermore, language may interact with questions' response scales as different cultures and ethnic minorities are associated with different response styles (Zax and Takahashi, 1967; Chun et al., 1974; Ross and Mirowsky, 1984; Hui and Triandis, 1989); thus, the choice of a particular response option may be influenced by the cultural frame primed by language, but further enhanced by the response style associated with the culture. Exploring such research avenues will help us understand how language of survey administration affects respondents' answers and enable us to collect accurate data. At present, little is known about these influences despite the considerable interest and resources invested in surveys of ethnic minorities and immigrants.

5. Appendix

Table 12. Relative Risk for the English Language Group versus Spanish of Immigrants who Came to the U.S. at Age 13 or Older across Measures where Language Effects are Expected (n=76)

Outcome Measure		Mantel-Haenszel Relative Risk for English	95% Confidence Interval	
Physical Health	Excellent vs. Very Good	0.81	0.42	1.57
	Excellent vs. Good	0.92	0.45	1.92
	Excellent vs. Fair or Poor	1.23	0.50	3.00
Mental Health	Excellent vs. Very Good	1.12	0.64	1.94
	Excellent vs. Good	1.21	0.64	2.31
	Excellent vs. Fair or Poor	1.86	0.55	6.27
Sad, Depressed for several days		1.00	0.62	1.61

Table 13. Means for Respondents who Came to the U.S. at Age 13 or Older across Measures where Language Effects are Expected (n=76)

Outcome Measure	English		Spanish		Mean Difference	SE
	Mean	SE	Mean	SE		
Language Effects Expected						
Age first drank alcohol	15.71	0.75	17.25	0.72	-3.61	1.04
Family pride	8.58	0.46	8.75	0.47	-0.17	0.66
Family cohesion	3.72	0.24	3.95	0.25	-0.23	0.35
Family cultural conflict	6.78	0.43	5.85	0.18	0.93	0.45

Table 14. Demographic Characteristics and Correlates of Acculturation by Compliance Groups

Variable		English Compliant n=152	English Noncompliant n=30	Spanish Compliant n=72	Spanish Noncompliant n=78	p-value for the English-Spanish Compliant Contrast
Age at immigration	US Born	54%	40%	49%	34%	0.21
	Less than 12	30	20	29	35	
	13-17	9	13	8	12	
	18-34	7	23	11	19	
	35 and older	0	3	3	0	
Years in the US	US born	54	40	49	34	0.43
	Less than 20	17	23	14	25	
	More than 20	29	37	38	42	
US citizen	Yes	84	83	86	83	0.62
	No	16	17	14	17	
Gender	Male	44	40	39	53	0.46
	Female	56	60	61	47	
Country of Education before 16	US	76	57	63	61	0.04
	Other	24	43	38	39	
Work Status	Employed	71	83	76	74	0.64
	Unemployed	5	0	6	4	
	Not in Labor Force	24	17	18	22	
Marital Status	Married, cohabiting	61	53	65	63	0.07
	Divorced, widowed	12	23	19	18	
	Never married	28	23	15	19	
US born parent	Yes	28	13	26	18	0.85
	No	72	87	74	82	
Language with Friends	English	47	7	33	35	0.05
	Spanish	11	37	21	13	
	Equally	42	57	46	52	
Language with Family	English	15	10	17	9	0.96
	Spanish	47	57	46	55	
	Equally	38	33	38	36	
Language in which Respondent Thinks	English	48	17	38	28	0.36
	Spanish	15	33	18	28	
	Equally	38	50	44	44	
Written English	Poor	1	10	6	3	0.13
	Fair	11	23	8	18	
	Good	39	33	36	53	
	Excellent	49	33	50	26	
Read English	Poor	0	3	4	1	0.09
	Fair	7	23	6	14	
	Good	38	40	36	53	
	Excellent	55	33	54	31	
Spoken English	Poor	0	0	3	1	0.13
	Fair	7	33	11	13	
	Good	39	33	36	61	
	Excellent	54	33	50	25	

Variable		English Compliant n=152	English Noncompliant n=30	Spanish Compliant n=72	Spanish Noncompliant n=78	p-value for the English- Spanish Compliant Contrast
Identify with Others of Same Ethnicity	Very Closely	65%	63%	64%	59%	0.80
	Somewhat Closely	29	33	28	29	
	Not at all/Not very closely	6	3	8	12	
Feel Close in Ideas with people of same descent	Very Close	39	57	46	46	0.23
	Somewhat Close	50	37	50	38	
	Not at all/Not very close	11	7	4	16	
Time would spend with people from same ethnicity	A Lot	36	43	32	35	0.44
	Some	52	47	50	52	
	None or little	12	10	18	13	
Importance of marriage within same ethnicity	Very important	9	20	17	15	0.29
	Somewhat important	24	30	28	19	
	Not very important	21	10	17	22	
	Not at all important	46	40	39	44	
Means for Continuous Measures						
	Age	35.36	39.67	38.50	38.59	0.11
	Income	64,874	58,446	68,033	68,801	0.69
	Years in school	13.07	12.40	12.96	12.77	0.75

Table 15. Language of Survey Administration by Propensity Strata

Propensity Stratum	Language of Interview		Total
	English	Spanish	
1	25	30	55
2	37	18	55
3	39	16	55
4	47	8	55
Total	148	72	220

Table 16. Balancing of Propensity Model Predictors across Propensity Strata

Covariate		Stratum 1		Stratum 2		Stratum 3		Stratum 4	
		English	Spanish	English	Spanish	English	Spanish	English	Spanish
Country Received Education before 16	US	48%	37%	78%	78%	77%	88%	85%	75%
	Other	52	63	22	22	23	13	15	25
Language with Friends	English	20	10	38	22	46	69	70	75
	Spanish	32	50	14	0	5	0	2	0
	English and Spanish equally	48	40	49	78	49	31	28	25
Language in which Respondent Thinks	English	32	20	43	44	46	50	62	68
	Spanish	20	30	8	11	21	6	13	13
	English and Spanish equally	48	50	49	44	33	44	26	25
Written English	Poor	4	13	0	0	0	0	0	0
	Fair	4	13	3	6	10	0	24	13
	Good	36	30	43	22	38	50	38	63
	Excellent	56	43	54	72	51	50	38	25
Gender	Male	36	30	32	44	49	44	53	50
	Female	65	70	68	56	51	56	47	50
Work Status	Employed	84	90	78	78	72	63	55	50
	Unemployed	8	0	8	11	5	13	2	0
	Not in Labor Force	8	10	14	11	23	25	43	50
Language with Family	English	12	23	14	11	15	19	19	0
	Spanish	44	53	41	33	51	44	51	50
	English and Spanish equally	44	23	46	56	33	38	30	50
Feel Close in Ideas with people of same descent	Very Close	56	60	43	61	38	13	23	25
	Somewhat Close	44	40	54	33	59	88	47	50
	Not at all or Not very close	0	0	3	6	3	0	30	25
US born parent	Yes	80	83	62	56	69	75	79	75
	No	20	17	38	44	31	25	21	25
Years in the US	US born	48	30	65	61	51	63	49	63
	less than 20	8	13	8	11	10	13	36	25
	more than 20	44	57	27	28	38	25	15	13
Importance of marriage within same ethnicity	Very important	32	30	5	11	0	6	4	0
	Somewhat important	28	37	35	33	21	6	17	25
	Not very important	0	10	24	11	28	31	26	25
	Not at all important	40	23	35	44	51	56	53	50
Means (standard error) for continuous variables									
Age		42	43	39	38	35	33	28	32
		(2.5)	(2.5)	(2.3)	(3.1)	(2.1)	(3.2)	(1.6)	(5.0)
Income in \$1000		67	80	58	70	75	54	61	45
		(11.2)	(11.6)	(7.3)	(10.6)	(10.1)	(12.9)	(8.0)	(15.5)
Years in school		13	13	13	13	13	13	13	13
		(0.5)	(0.6)	(0.4)	(0.5)	(0.4)	(0.7)	(0.4)	(0.6)

Questions where Language effects are Expected

SC8.1. How would you rate your overall physical health – excellent, very good, good, fair, or poor?

EXCELLENT1
VERY GOOD.....2
GOOD.....3
FAIR.....4
POOR.....5
DON'T KNOW.....8
REFUSED.....9

SC8.2. How would you rate your overall mental health – excellent, very good, good, fair, or poor?

EXCELLENT1
VERY GOOD.....2
GOOD.....3
FAIR.....4
POOR.....5
DON'T KNOW.....8
REFUSED.....9

SC21. Have you ever in your life had a period of time lasting several days or longer when most of the day you felt sad, empty or depressed?

1. YES
 5. NO
 8. DONT KNOW
 9. REFUSED
-

SU1. The next questions are about your use of alcohol. How old were you the very first time you ever drank an alcoholic beverage – including either beer, wine, a wine cooler, or hard liquor?

_____ YEARS OLD

(IF VOL): "NEVER".....997
DON'T KNOW.....998
REFUSED.....999

Family Pride Subscale

FC1. Now I'd like to know how strongly you agree or disagree with the following statements about your family. Family members respect one another.

- 1 STRONGLY AGREE
 - 2 SOMEWHAT AGREE
 - 3 SOMEWHAT DISAGREE
 - 4 STRONGLY DISAGREE
 - 9 (M) REFUSED
 - 8 (M) DONT KNOW
-

FC2. We share similar values and beliefs as a family.

- 1 STRONGLY AGREE
 - 2 SOMEWHAT AGREE
 - 3 SOMEWHAT DISAGREE
 - 4 STRONGLY DISAGREE
 - 9 (M) REFUSED
 - 8 (M) DONT KNOW
-

FC3. Things work well for us as a family.

- 1 STRONGLY AGREE
 - 2 SOMEWHAT AGREE
 - 3 SOMEWHAT DISAGREE
 - 4 STRONGLY DISAGREE
 - 9 (M) REFUSED
 - 8 (M) DONT KNOW
-

Questions where Language effects are Expected

FC4. We really do trust and confide in each other.

- 1 STRONGLY AGREE
 - 2 SOMEWHAT AGREE
 - 3 SOMEWHAT DISAGREE
 - 4 STRONGLY DISAGREE
 - 9 (M) REFUSED
 - 8 (M) DONT KNOW
-

FC5. Family members feel loyal to the family.

- 1 STRONGLY AGREE
 - 2 SOMEWHAT AGREE
 - 3 SOMEWHAT DISAGREE
 - 4 STRONGLY DISAGREE
 - 9 (M) REFUSED
 - 8 (M) DONT KNOW
-

FC6. We are proud of our family.

- 1 STRONGLY AGREE
 - 2 SOMEWHAT AGREE
 - 3 SOMEWHAT DISAGREE
 - 4 STRONGLY DISAGREE
 - 9 (M) REFUSED
 - 8 (M) DONT KNOW
-

FC7. We can express our feelings with our family.

- 1 STRONGLY AGREE
 - 2 SOMEWHAT AGREE
 - 3 SOMEWHAT DISAGREE
 - 4 STRONGLY DISAGREE
 - 9 (M) REFUSED
 - 8 (M) DONT KNOW
-

Family Cohesion Subscale

FC8. Family members like to spend free time with each other.

- 1 STRONGLY AGREE
 - 2 SOMEWHAT AGREE
 - 3 SOMEWHAT DISAGREE
 - 4 STRONGLY DISAGREE
 - 9 (M) REFUSED
 - 8 (M) DONT KNOW
-

FC9. Family members feel very close to each other.

- 1 STRONGLY AGREE
 - 2 SOMEWHAT AGREE
 - 3 SOMEWHAT DISAGREE
 - 4 STRONGLY DISAGREE
 - 9 (M) REFUSED
 - 8 (M) DONT KNOW
-

FC10. Family togetherness is very important.

- 1 STRONGLY AGREE
 - 2 SOMEWHAT AGREE
 - 3 SOMEWHAT DISAGREE
 - 4 STRONGLY DISAGREE
 - 9 (M) REFUSED
 - 8 (M) DONT KNOW
-

Family Cultural Conflict Subscale

FC11a. Please tell me how frequently the following situations have occurred to you. You have felt that being too close to your family interfered with your own goals.

- 1 HARDLY EVER OR NEVER
- 2 SOMETIMES
- 3 OFTEN
- 9 (M) REFUSED
- 8 (M) DONT KNOW

FC11c. Because you have different customs, you have had arguments with other members of your family.

- 1 HARDLY EVER OR NEVER
- 2 SOMETIMES
- 3 OFTEN
- 9 (M) REFUSED
- 8 (M) DONT KNOW

FC11d. Because of the lack of family unity, you have felt lonely and isolated.

- 1 HARDLY EVER OR NEVER
- 2 SOMETIMES
- 3 OFTEN
- 9 (M) REFUSED
- 8 (M) DONT KNOW

FC11f. You have felt that family relations are becoming less important for people that you are close to.

- 1 HARDLY EVER OR NEVER
- 2 SOMETIMES
- 3 OFTEN
- 9 (M) REFUSED
- 8 (M) DONT KNOW

FC11g. Your personal goals have been in conflict with your family.

- 1 HARDLY EVER OR NEVER
- 2 SOMETIMES
- 3 OFTEN
- 9 (M) REFUSED
- 8 (M) DONT KNOW

Questions where Language Effects are Not Expected

MAR3CAT. Are you currently married, separated, divorced, widowed, or never married?

- 1 MARRIED/COHABITING
- 2 DIVORCED/SEPARATED/WIDOWED
- 3 NEVER MARRIED
- 9 (M) REFUSED
- 8 (M) DONT KNOW

CN1. The next questions are about children. How many living biological children do you have, not counting stepchildren, adopted children, or foster children?

_____ BIOLOGICAL CHILDREN

NONE.....00
DON'T KNOW.....98
REFUSED.....99

Figure 4. Outcome Measures of Interest from the NLAAS Questionnaire

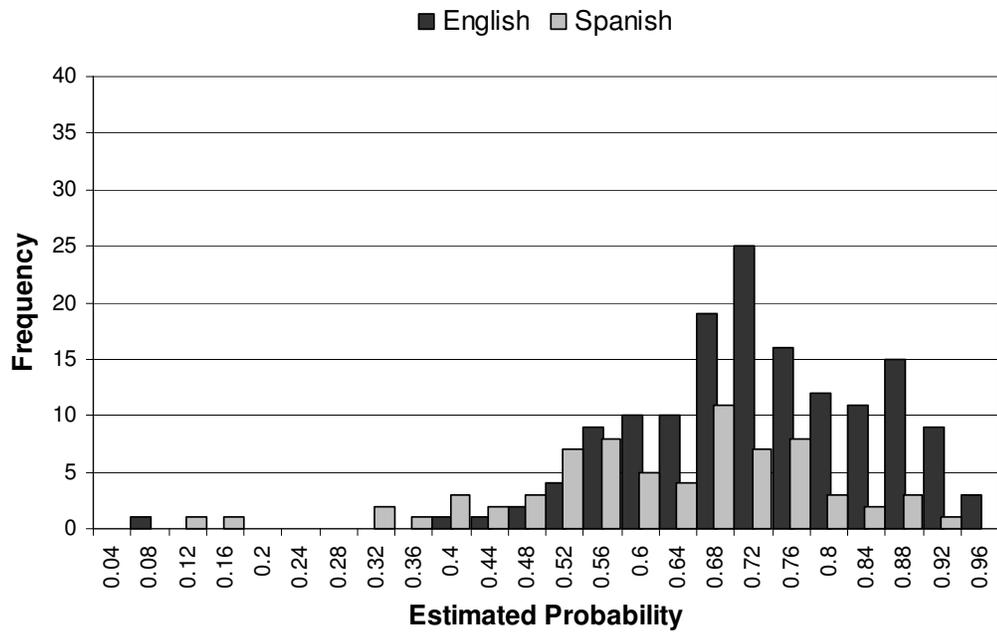


Figure 5. Overlap of Estimated Propensity Scores for Those Interviewed in English and Those Interviewed in Spanish

Chapter V. Conclusions

This dissertation addressed questions about the existence and practical importance of language effects in surveys of bilingual bicultural respondents. It presented three related analyses – the first one proposed a framework for the possible language influences at each step of the survey response formation process, based on findings from psycholinguistics and cross-cultural psychology. The second examined the existence of language effects in observational data from a national survey of immigrants. It compared groups of questions where language effects were expected based on divergent norms of the two cultures under study (Hispanic vs. American) and groups of questions where no such effects were expected. The last presented analyses of an experimental assignment of bilingual respondents to language of survey administration in a national survey of ethnic minorities, comparing questions on similar topics as the ones examined in the observational study.

1. What have we learned?

The presented theoretical framework outlined several possible mechanisms that can induce language effects at each stage of the response formation process, independent of translation. However, at the current stage of

development of the cross-cultural survey field and the existing data we can not further evaluate the hypotheses motivated by the framework. Systematic investigation of the expected language effects is necessary to determine the prevailing mechanisms inducing such effects and the questions most likely to be affected by them.

Overall, the empirical investigation found limited support for language influences in surveys. We found significant effects of language only for questions related to current alcohol use, family pride and number of biological children. Specifically, consistent with more conservative alcohol norms and attitudes among Hispanics (Cateano and Clark, 1999), Hispanic bilingual bicultural respondents interviewed in Spanish reported lower current rates of alcohol consumption relative to those interviewed in English. Similarly, consistent with the importance of familism in the Hispanic culture, Hispanic bilingual respondents reported higher levels of family pride when interviewed in Spanish relative to English. However, even though in the expected direction, the difference between the two language groups on the rest of familism subscales failed to reach statistical significance. Surprisingly, we found significant differences in reports of number of children in Spanish versus English in the NIS data. The results did not replicate with the NLAAS data, possibly suggesting the propensity model used in NIS analyses failed to account for unobserved confounding covariates. We failed to detect language effects in responses to questions related to respondent's physical and mental health. One possible explanation is that the social desirability

effects were not that different between the two cultures. As expected, responses to factual questions such as marital status, living situation, and number of household members did not exhibit language influences.

A comforting, but dangerous conclusion based on such results will be that the psycholinguistic and psychological theories are irrelevant for survey methodologists. However, various factors may explain the lack of strong language effects in this investigation. First, the level of acculturation in both samples may not have presented ideal conditions for detection of language influences – the NIS sample consisted of legal immigrants interviewed 2-3 years after the immigration process was completed¹⁴, while only 23% of the NLAAS sample of bilingual respondents that were originally randomly assigned to a language came to the U.S. in their teens or older¹⁵. We suspect that the longer one has been living in the new culture, the more familiar one is with nuances and differences between that culture and one's original cultural system, with what is acceptable and appropriate in one versus the other. Furthermore, the younger one was when first exposed to the new culture, the less sensitive to differences in the two systems one would be as he/she would have had less exposure to the culture-specific values and norms of the culture of origin. In addition, individuals who learned the second language in a context different than the context of their mother-

¹⁴ Even though the propensity model controlled for whether the respondent was a new immigrant or adjusted his/her status, we were unable to further differentiate among respondents in terms of their exposure and learning of formal and informal cultural images, practice of different culture's traditions and norms.

¹⁵ This percentage became even smaller when restricting the sample to only those who complied with the language assignment, preventing us to test the hypotheses on the further subset of immigrants.

tongue are likely to have separate cognitive representations of two translation equivalents (Ervin and Osgood, 1954; Lambert et al., 1958), while individuals who learned both languages in parallel share the same representational system. This suggests that language may not have the same priming effects across all bilinguals. Lack of precise data and sample size restrictions prevented us from further examining this possibility.

Second, in both surveys, levels of bilingualism were determined based on self-reports. In contrast, in most of the presented studies, bilingualism was tested during the recruitment process (e.g., Marin et al., 1983), or participants were students who attended school in the country of their second language (e.g., Marian and Neisser, 2000) and were selected based on their scores on the Test of English as a Foreign Language (e.g., Matsumoto and Stanny, 2006). To the extent that bilingualism is a socially desirable characteristic and its reports are influenced by the language of interview (completed in the self-selected language in NIS, and in English in NLAAS), the sample of self-determined bilingual respondents may not be ideal for testing language influences. That is, we selected respondents who reported equal familiarity with both languages, but there may be measurement error bias in these self-reports. If respondents truly overestimated their level of bilingualism, then we would not expect to see strong language influences because of a different mechanism related to language use in novice bilinguals – studies by Kroll and Sholl, 1992; Kroll and Steward, 1994; Dufour and Kroll, 1995 suggest that novice bilinguals employ a “think-then-translate-then-speak” process when

using their second language, rather than think in the second language. If this is the case, then the link between language as a culture cuing factor and cognition will be weakened or lost. Additionally, the information we have about language competence based on the self-report is insufficient to get an idea how diverse are respondents in their bilingual fluency – some may still be acquiring their second language, some may be dominant in one language, while others may be balanced on both languages. Such variability can reduce the language effects that are sought.

Third, the absence and breakdown of random assignment in the two datasets necessitated the use of propensity score methods. Despite their promise in addressing causal questions in the absence of randomization, propensity score methods can only adjust for observed confounding covariates, thus any unobserved covariates may challenge the interpretation of results. Therefore, in contrast to the presented carefully controlled randomized psycholinguistic and psychological studies, the expected balance between unknown factors in the two language groups may have been lost. Such confounding variables can be uncaptured nuances of acculturation (e.g., whether one identifies more with the Hispanic or American culture), nuances of bilingualism (e.g., how the second language was acquired, whether there is a dominant language, everyday use of each language, domains in which each language is used), personal characteristics related to the need to provide socially acceptable responses, presence of others during

the interview, display of cultural icons, etc. At this point it is unclear whether such factors may or may not interfere with language influences.

Fourth, the conditions of the cited experiments and the survey setting may further contribute to the fragile language effects we observed. In bilingual experiments, participants usually are aware or can guess that their languages are at issue (Grosjean, 1998). Thus, whatever self-representation is activated is attended by awareness of the language spoken (Schrauf, 2000). In contrast, in surveys, there are often many distractions to which respondents are subjected – in addition to the survey sponsor or multiple question topics surrounding the question on language skills, various other factors at the respondent's house (e.g., a stranger in the house, ringing telephone, loud TV, crying children, displayed cultural icons/primes) may interfere with the linguistic conditioning. Furthermore, interviewer's observable characteristics, such as ethnic belonging or accent (in addition to age in the Japanese culture, or gender in the Arabic world) may interact with language influences. We suspect that the effect of language priming culture in interviewer administered surveys is dependent on interviewer characteristics – when interviewers use the language of the dominant culture, but belong to an ethnic minority and/or have an accent, the language of interview will be unlikely to prime the dominant culture. Thus, in interviewer administered surveys, language effects may be masked or mitigated by the presence and interaction with an interviewer. We were unable to investigate this with the available data.

Finally, despite the deliberate efforts in both surveys to ensure cultural relevance and equivalence across language versions, the translation of the questionnaire remains an unexplored possibility. It is possible that the particular words used (and their concreteness) in the Spanish language versions of the surveys have diminished the difference in perceived affective characteristics of the questions under study. This may be more of a threat to some questions than others; arguably, translation of marital status is less suspect than measures of mental health to differences in translation. Alternatively, despite the evidence in the existing literature, the examined topics themselves may not be strong enough contrasts to allow us to detect language influences. We based our assertions about the existing cultural differences in norms related to mental and physical health on existing literature, but did not have direct measures of the stigma associated with them in the Hispanic and American cultures.

2. Implications for survey practitioners

The empirical results suggest language influences may pose less of a threat to surveys of immigrants and ethnic minorities than the psycholinguistic and cross-cultural psychological research suggests. However, it is premature at this point to comfortably conclude that language of survey administration does not contribute to measurement differences in the responses of bilingual-bicultural individuals. The existence of some effects requires further

investigation of the conditions and mechanisms that produce them. The presented results suggest that language can cue the interpretive frame bicultural bilingual respondents adopt. In turn, this implies that leaving the choice of language to the bilingual respondent (or interviewer) may not always be a good practice.

Further, the reported language effects in this investigation, even though weak, suggest that the same questions can be perceived to have different affective characteristics depending on language and the cultural norms it activates; thus, more or less socially desirable opinions are expressed depending on language. Knowing in advance how cultures differ in terms of question affective characteristics may better inform questionnaire design and various techniques may be used to reduce social desirability or sensitivity across language versions.

Finally, we were unable to test the effect of language-dependent recall in this investigation of language influences, but the presented literature from other fields suggests that language may affect both quality and quantity of recall. For the survey practitioner this implies that language of administration in bilingual-bicultural respondents may be switched throughout the survey, depending on life periods and/or life domains of interest. However, in order to do this, we need to invent better measures of bilingualism than self-reports.

3. Next Steps

To truly understand the role of language effects in surveys involving bilingual respondents, an examination in an experimental survey setting should be conducted. The theoretical framework presented in this dissertation demonstrates that language effects can occur in any two cultures, but the items where such effects will be exhibited are dependent on the differences between cultures and/or the match between the language of information encoding and language of interview. We expect to detect the strongest influence of language when two cultures take strongly diverging positions on an issue. Cultural differences are often related to religion, societal structure, customs, and habits. For example, the Arabic culture differs from the American culture in many of these aspects and there is a rich pool of topics on which the two cultures take opposite standpoints (e.g., the role of women in society, the approval of premarital sex). Further, bilingual immigrants acquire and use their languages for different purposes and in different life-domains (Grosjean, 1997) – for example, at home and at work. We can expect strong language influences when the recall of autobiographical events of interest is requested in a language different from the language of initial encoding.

In fall 2008 we are launching a small-scale web survey of adult Arab-American immigrants from the Detroit area. The proposed experiment has a within-subject design with a two week period between two administrations of the same instrument. In addition to measuring attitudes on topics where the

American and Arab cultures take opposing standpoints and questions on fictitious issues¹⁶, we would attempt to replicate language-dependent recall in a survey setting. We focus on events that take place in the respondent's home and at the respondent's work, hypothesizing that the recall of work related information will be facilitated by being asked in English, whereas the recall of home related information will be facilitated by being asked in Arabic, reflecting respondents' differential language use in these two settings. Overall, the design will take into consideration possible issues that prevented us from detecting strong language effects in the NIS and NLAAS data. Specifically, respondents' bilingual competence will be tested in the screener and various measures of acculturation will be employed to ensure the sample has sufficient exposure and practice with the new language and culture. To prevent possible confounding interviewer effects (and for budget reasons), the survey will be self-administered. Finally, the questionnaire will be pretested to ensure the questions possess different affective characteristics across the two cultures and will allow us to capture language influences as intended.

If, under such controlled conditions, we fail to detect strong language influences, the effect of language of administration in surveys of bicultural bilingual respondents may be considered negligible. However, further investigation into what causes some of the predicted effects to appear (as in the NIS and NLAAS data) will be conducted. Special attention will be focused

¹⁶ Admitting ignorance is not acceptable in the Arabic culture, so it is expected that respondents will be more likely to express an opinion when asked in Arabic, but opt for "Don't Know" when asked in English.

on levels of acculturation, levels of language fluency and interviewer effects. Any discovered differences will motivate additional examination to disentangle the possible mechanisms that play a role at the various stages of the response formation process.

References

- Allegría, M., D. Vila, M. Woo, G. Canino, D. Takeuchi, M. Vera, V. Febo, P. Guarnaccia, S. Aguilar-Gaxiola and P. Shrout (2004). "Cultural Relevance and Equivalence in the Nlaas Instrument: Intergrating Etic and Emic in the Development of Cross-Cultural Measures for a Psychiatric Epidemiology and Services Study of Latinos." International Journal of Methods in Psychiatric Research **13**(4): 270-288.
- Allport, G. (1954). The Nature of Prejudice. Reading, MA, Addison-Wesley.
- Alva, S. A. (1991). "Academic Invulnerability among Mexican American Students: The Importance of Protective Resources and Appraisals." Hispanic Journal of Behavioral Sciences **13**: 18–34.
- Alva, S. A. (1995). "Psychological Distress and Alcohol Use in Hispanic Adolescents." Journal of Youth and Adolescence **24**.
- Bazzoui, W. and I. Al-Issa (1966). "Psychiatry in Iraq." British Journal of Psychiatry **112**: 827-832.
- Berry, J. (1990). Psychology of Acculturation. Proceedings of Nebraska Symposium on Motivation: Cross-cultural Perspectives, Lincoln, NE.
- Berry, J. and D. Sam (1996). Acculturation and Adaptation. Handbook of Cross-Cultural Psychology: Social Behavior and Applications. J. Berry, M. Segall and C. Kagitcibasi. Boston, Allyn&Bacon. **3**: 291-325.
- Bickel, B. (1997). Spacial Operations in Deixis, Cognition and Culture: Where to Orient Oneself in Belhare. Language and Conceptualization. J. Nuyts and E. Pederson. Cambridge, University Press.
- Bodenhausen, G. V. and R. S. Wyer (1987). Social Cognition and Social Reality: Information Acquisition and the Use of Laboratory and the Real World. Social Information Processing and Survey Methodology. H. J. Hippler, N. Schwarz and S. Sudman. New York, Springer: 6-41.
- Bond, M. H. and K. S. Yang (1982). "Ethnic Affirmation Versus Cross-Cultural Accomodation: The Variable Impact of Questionnaire Language on Chinese Bilinguals from Hong Kong." Journal of Cross-Cultural Psychology **13**: 169-185.
- Boroditsky, L. (2001). "Does Language Shape Thought? Mandarin and English Speakers' Conception of Time." Cognitive Psychology **43**: 1-22.
- Boroditsky, L., L. Schmidt and W. Phillips (2003). Sex, Syntax and Semantics. Language in Mind. D. Gentner and S. Goldin-Meadow. Cambridge, MA, MIT Press: 61-79.
- Botha, E. (1968). "Verbally Expressed Values of Bilinguals." Journal of Social Psychology **75**: 159-164.
- Bower, G. H. (1981). "Mood and Memory." American Psychologist **36**: 129-148.

- Bower, G. H., K. P. Monteiro and S. G. Gilligan (1978). "Emotional Mood as a Context for Learning and Recall." Journal of Verbal Learning and Verbal Behavior **17**: 573-585.
- Bowerman, M. and S. Choi (2003). Space under Construction: Language-Specific Spatial Categorization in First Language Acquisition. Language in Mind: Advances in Studies of Language and Thought. D. Gentner and S. Goldin-Meadow. Cambridge, MA, MIT Press: 387-427.
- Brach, C., I. Fraser and K. Paez (2005). "Crossing the Language Chasm." Health Affairs **24**: 424-434.
- Bradburn, N. M., L. J. Rips and S. K. Shevell (1987). "Answering Autobiographical Questions: The Impact of Memory and Inference on Surveys." Science **236**: 157-161.
- Breslau, J., S. Aguilar-Gaxiola, G. Borges, R. C. Castilla-Puentes, K. S. Kendler, M.-E. Medina-Mora, M. Su and R. C. Kessler (2007). "Mental Disorders among English-Speaking Mexican Immigrants to the Us Compared to a National Sample of Mexicans." Psychiatry Research **151**: 115-122.
- Brewer, W. F. (1988). Memory for Randomly Sampled Autobiographical Events. Remembering Reconsidered: Ecological and Traditional Approaches to the Study of Memory. U. Neisser and E. Winograd. Cambridge, Cambridge University Press: 21-90.
- Briere, J. and C. Lanktree (1983). "Sex-Role Related Effects of Sex Bias in Language." Sex Roles **9**(5): 625-631.
- Bugelski, B. R. (1977). "Imagery and Verbal Behavior." Journal of Mental Imagery **1**: 39-52.
- Caetano, R. (1987). "Alcohol Use and Depression among U.S. Hispanics." Addiction **82**: 1245-1251.
- Caetano, R. and C. L. Clark (1999). "Trends in Situational Norms and Attitudes toward Drinking among Whites, Blacks, and Hispanics: 1984-1995." Drug and Alcohol Dependence **54**: 45-56.
- Canino, G. and M. Bravo (1994). "The Adaptation and Testing of Diagnostic and Outcome Measures for Cross-Cultural Research." International Review of Psychiatry **6**: 281-286.
- Canino, G., R. Lewis-Fernandez and M. Bravo (1997). "Methodological Challenges in Cross-Cultural Mental Health Research." Transcultural Psychiatric Research Review **34**(2): 163-184.
- Cannell, C., P. Miller and L. Oksenberg (1981). Research on Interviewing Techniques. Sociological Methodology 1981. S. Leinhardt. San Francisco, Jossey-Bass: 389-437.
- Cateano, R. and C. L. Clark (1999). "Trends in Situational Norms and Attitudes toward Drinking among Whites, Blacks, and Hispanics: 1984-1995." Drug and Alcohol Dependence **54**: 45-56.
- Cervantes, R. C., A. M. Padilla and W. N. Salgado de Snyder (1991). "The Hispanic Stress Inventory: A Culturally Relevant Approach to Psychological Assessment." Psychological Assessment **3**: 438-47.
- Chaleby, K. (1987). "Women in Polygamous Marriages in out-Patient Psychiatric Services in Kuwait." International Journal of Family Psychiatry **8**(1): 25-34.

- Chang, I., S. C. Lapham and K. J. Barton (1996). "Drinking Environment and Sociodemographic Factors among Dwi Offenders." Journal of Studies on Alcohol **57**: 659–669.
- Chee, M. W., D. Caplan, C. S. Soon, N. Sriram, E. W. L. Tan, T. Thiel and B. Weeks (1999). "Processing of Visually Presented Sentences in Mandarin and English Studies with Fmri." Neuron **23**: 127-137.
- Chen, C., S. Lee and H. W. Stevenson (1995). "Response Style and Cross-Cultural Comparisons of Rating Scales among East Asian and North American Students." Psychological Science **6**: 170-175.
- Chiu, L. H. (1972). "A Cross-Cultural Comparison of Cognitive Styles in Chinese and American Children." International Journal of Psychology **7**: 235-242.
- Chomsky, N. (1965). Aspects of the Theory of Syntax. Cambridge, MA, MIT Press.
- Chuang, Y. (1999). Fusion: The Primary Model of Bicultural Competence and Bicultural Identity Development in a Taiwanese-American Family Lineage, Stanford University, Palo Alto, CA.
- Chun, K. T., J. B. Campbell and J. H. Yoo (1974). "Extreme Response Styles in Cross-Cultural Research: A Reminder." Journal of Cross-cultural Psychology **5**: 465-480.
- Clark, W. B. and M. Hill, Eds. (1991). Alcohol in America: Drinking Practices and Problems. Albany, NY, State University of New York Press.
- Collet, C. (2007 of Conference). Asian-Americans, Item Nonresponse and Neutrality in Telephone Surveys: Some Hypotheses and Tests. Paper Presented at AAPOR, Anaheim, CA.
- Collins, K., D. Hughes, M. Doty, B. Ives, J. Edwards and K. Tenney (2002). Diverse Communities, Common Concerns: Assessing Health Care Quality for Minority Americans: Findings from the Commonwealth Fund 2001, Health Care Quality Survey. New York, NY, Commonwealth Fund.
- Converse, J. and S. Presser (1986). Survey Questions: Handcrafting the Standardized Questionnaire. Thousand Oaks, CA, Sage Publications.
- Couper, M. P., E. Singer and R. A. Kulka (1998). "Participation in the 1990 Decennial Census: Politics, Privacy and Pressure." American Politics Quarterly **26**: 59-80.
- Dawson, D. A., B. F. Grant, P. S. Chou and R. P. Pickering (1995). "Subgroup Variation in U.S. Drinking Patterns: Results of the 1992 National Longitudinal Alcohol Epidemiologic Study." Journal of Substance Abuse **7**: 331-344.
- De Las Fuentes, C., A. J. Baron and M. J. T. Vasquez (2003). Teaching Latino Psychology. Teaching Gender and Multi-Cultural Awareness: Resources for the Psychology Classroom. P. Bronstein and K. Quina. Washington, DC, American Psychological Association.
- DeMaio, T. J. (1984). Social Desirability and Survey Measurement: A Review. Surveying Subjective Phenomena. C. F. Turner and E. Martin. New York, Russell Sage Foundation. **2**.

- Dufour, R. and J. F. Kroll (1995). "Matching Words to Concepts in Two Languages: A Test Concept of the Concept Mediation Model of Bilingual Memory." Memory and Cognition **23**(2): 166-180.
- Earl, M. (1969). "A Cross-Cultural and Cross-Language Comparison of Dogmatism Scores." Journal of Social Psychology **4**: 1-20.
- Edgerton, R. B. and M. Karno (1971). "Mexican-American Bilingualism and the Perception of Mental Illness." Archives of General Psychiatry **24**: 286-290.
- Eich, J. E., H. Weingartner, R. C. Stillman and J. C. Gillin (1975). "State-Dependent Accessibility of Retrieval Cues in the Retention of a Categorized List." Journal of Verbal Learning and Verbal Behavior **14**: 408-417.
- Einstein, A. (1954). Ideas and Opinions. New York, Crown.
- Ervin, S. M. and C. E. Osgood (1954). "Second Language Learning and Bilingualism." Journal of Abnormal and Social Psychology **49**: 139-146.
- Ewing, K. P. (1990). "The Illusion of Wholeness: Culture, Self and the Experience of Inconsistency." Ethos **18**(3): 251-278.
- Feldman, R. H. L. (1975). "The Effect of Administrator and Language on Traditional-Modern Attitudes among Gusii Students in Kenya." Journal of Social Psychology **96**: 141-142.
- Fordman, S. and J. Ogbu (1986). "Black Students' School Success: Coping with the "Burden of 'Acting White'"." The Urban Review **18**: 176-206.
- Fowler, F. J. J. (1995). Improving Survey Questions: Design and Evaluation. Thousand Oaks, CA, Sage Publications.
- Fu, G., K. Lee, C. A. Cameron and F. Xu (2001). "Chinese and Canadian Adults' Categorization and Evaluation of Lie and Truth Telling About Prosocial and Antisocial Behaviors." Journal of Cross-cultural Psychology **32**(720-727).
- Geertz, C. (1993). The Interpretation of Cultures. New York, Basic Books.
- Gentner, D. (2003). Why We're So Smart. Language in Mind. D. Gentner and S. Goldin-Meadow. Cambridge, MA, MIT Press: 195-235.
- Gil, A. G. and W. A. Vega (1996). "Two Different Worlds: Acculturation Stress and Adaptation among Cuban and Nicaraguan Families." Journal of Social and Personal Relationships **13**: 435-56.
- Gil, A. G., W. A. Vega and J. M. Dimas (1994). "Acculturative Stress and Personal Adjustment among Hispanic Adolescent Boys." Journal of Community Psychology **22**: 43-54.
- Godden, D. R. and A. D. Baddeley (1975). "Context-Dependent Memory in Two Natural Environments: On Land and under Water." British Journal of Psychology **66**: 325-331.
- Goldstone, R. L. (1998). "Perceptual Learning." Annual Review of Psychology **49**: 585-612.
- Grant, B. F., F. S. Stinson, D. S. Hasin, D. A. Dawson, S. P. Chou and K. Anderson (2004). "Immigration and Lifetime Prevalence of Dsm-Iv Psychiatric Disorders among Mexican Americans and Non-Hispanic Whites in the United States: Results from the National Epidemiologic

- Survey on Alcohol and Related Conditions." Archives of General Psychiatry **61**: 1226–1233.
- Greenberg, J. H. (1978). Generalization About Numeral Systems. Universals of Human Language. Word Structure. J. H. Greenberg. Stanford, CA, Academic Press. **3**.
- Greenfield, T. K. and L. A. Kaskutas (1998). "Five Years' Exposure to Alcohol Warning Label Messages and Their Impacts: Evidence from Diffusion Analysis." Applied Behavioral Science Review **6**: 39-68.
- Grice, P. (1989). Studies in the Way of Words. Cambridge, MA, Harvard University Press.
- Grosjean, F. (1997). "The Bilingual Individual." Interpreting: International Journal of Research and Practice in Interpreting **2**: 163-187.
- Grosjean, F. (1998). "Studying Bilinguals: Methodological and Conceptual Issues." Bilingualism: Language and Cognition **1**: 131-149.
- Gurin, P., A. Hurtado and T. Peng (1994). "Group Contacts and Ethnicity in the Social Identities of Mexicanos and Chicanos." Personality and Social Psychology Bulletin **20**: 512-532.
- Haberstroh, S., D. Oyserman, N. Schwarz, U. Kuhnen and L. Ji (2002). "Is the Interdependent Self More Sensitive to Question Context Than the Independent Self? Self-Construal and the Observation of Conversational Norms." Journal of Experimental Social Psychology **38**: 323-329.
- Harkness, J. (2003). Questionnaire Translation. Cross Cultural Survey Methods. J. A. Harkness, F. R. Van de Vijver and P. P. Mohler. New York, Wiley.
- Harkness, J. A., F. R. Van de Vijver and P. P. Mohler, Eds. (2003). Cross Cultural Survey Methods. New York, Wiley.
- Harris, K., M. Edlund and S. Larson (2005). "Racial and Ethnic Differences in Mental Health Problems and Use of Mental Health Care." Medical Care **43**: 775–784.
- Hayashi, E. (1992). "Belief Systems, the Way of Thinking, and Sentiments of Five Nations." Behaviormetrica **19**: 127-170.
- Health and Human Services Department (2001). Mental Health: Culture, Race, and Ethnicity—a Supplement to Mental Health: A Report of the Surgeon General. M. Rockville, U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Mental Health Services.
- Heeringa, S., J. Wagner, M. Torres, N. Duan, T. Adams and P. Berglund (2004). "Sample Designs and Sampling Methods for the Collaborative Psychiatric Epidemiology Studies (Cpes)." International Journal of Methods in Psychiatric Research **13**: 221-240.
- Hoffman, C., I. Lau and D. R. Johnson (1986). "The Linguistic Relativity of Person Cognition: An English-Chinese Comparison." Journal of Personality and Social Psychology **51**: 1097-1105.
- Hofstede, G. (1980). Culture's Consequences. Beverly Hills, CA, Sage.
- Hong, Y. Y., C. Y. Chiu and T. Kung (1997). Bringing Culture out in Front: Effects of Cultural Meaning System Activation on Social Cognition. Progress in

- Asian Social Psychology. K. Leung, Y. Kashima, U. Kim and S. Yamaguchi. Singapore, Wiley. **1**: 135-146.
- Hui, C. H. and H. C. Triandis (1989). "Effects of Culture and Response Format on Extreme Response Style." Journal of Cross-Cultural Psychology **20**: 296-309.
- Hunt, E. and F. Agnoli (1991). "The Whorfian Hypothesis: A Cognitive Psychology Perspective." Psychological Review **98**: 377-389.
- Jason, B. and C. J. Cherpitel (2004). "Measuring Accuracy of Self-Reported Vehicular Offences and Blood Alcohol Concentration Levels among Whites and Mexican Americans Mandated to Drinking Driver Programs." Addictive Behaviors **29**: 1163-1169.
- Jasso, G., D. S. Massey, M. R. Rosenzweig and J. P. Smith (2005). The U.S. New Immigrant Survey: Overview and Preliminary Results Based on the New-Immigrant Cohorts of 1996 and 2003. Immigration Research and Statistics Service Workshop on Longitudinal Surveys and Cross-Cultural Survey Design: Workshop Proceedings. B. Morgan and B. Nicholson. London, Crown Publishing: 29-46.
- Ji, L., N. Schwarz and R. Nisbett (2000). "Culture, Autobiographical Memory and Behavioral Frequency Reports: Measurement Issues in Cross-Cultural Studies." Personality and Social Psychology Bulletin **26**: 586-594.
- Ji, L., Z. Zhang and R. Nisbett (2002). Culture, Language and Categorization, Queens University, Kingston, Ontario.
- Ji, L., Z. Zhang and R. Nisbett (2004). "Is It Culture or Is It Language? Examination of Language Effects in Cross-Cultural Research on Categorization." Journal of Personality and Social Psychology **87**(1): 57-65.
- Johnson, T. P. and F. J. R. Van de Vijver (2003). Social Desirability in Cross-Cultural Research. Cross-Cultural Survey Methods. J. Harkness, F. J. R. Van de Vijver and P. Mohler. Hoboken, NJ, John Wiley and Sons: 195-204.
- Kashima, Y. (2000). "Conceptions of Culture and Person for Psychology." Journal of Cross-Cultural Psychology **31**: 14-32.
- Kay, P. and W. Kempton (1984). "What Is the Sapir-Whorf Hypothesis?" American Anthropologist **86**: 65-79.
- Kemmelmeier, M. and B. Y. M. Cheng (2004). "Language and Self-Construal Priming: A Replication and Extension in a Hong Kong Sample." Journal of Cross-Cultural Psychology **35**: 705-712.
- Khosroshahi, E. (1989). "Penguins Don't Care, but Women Do: A Social Identity Analysis of a Whorfian Problem." Language and Society **18**: 505-525.
- Kim, K., N. Relkin, K. Lee and J. Hirsch (1997). "Distinct Cortical Areas Associated with Native and Second Languages." Nature(388): 171-174.
- Kroll, J. F. and A. Sholl (1992). Lexical and Conceptual Memory in Fluent and Nonfluent Bilinguals. Cognitive Processing in Bilinguals. R. J. Harris. Amsterdam, Elsevier Science: 191-204.

- Kroll, J. F. and E. Steward (1994). "Category Interference in Translation and Picture Naming: Evidence for Assymetric Connection between Bilingual Memory Representations." Journal of Memory and Language **33**(2): 149-174.
- LaFromboise, T., H. Coleman and J. Gerton (1993). "Psychological Impact of Biculturalism: Evidence and Theory." Psychological Bulletin **114**: 395-412.
- Lambert, W. E., J. Havelka and C. Crosby (1958). "The Influence of Language-Acquisition Contexts on Bilingualism." Journal of Abnormal and Social Psychology **56**: 239-244.
- Laroche, M., C. Kim, M. K. Hui and M. A. Tomiuk (1998). "Test of a Nonlinear Relationship between Linguistic Acculturation and Ethnic Identification." Journal of Cross-Cultural Psychology **29**: 418-433.
- Lee, K., F. Xu, G. Fu, C. A. Cameron and S. Chen (2001). "Taiwan and Mainland Chinese and Canadian Children's Categorization and Evaluation of Lie- and Truth-Telling: A Modesty Effect." British Journal of Developmental Psychology **19**: 525-542.
- Levine, R. V., A. Norenzayan and P. K. (2001). "Cross-Cultural Differences in Helping Strangers." Journal of Cross-Cultural Psychology **32**: 534-560.
- Levinson, S. (1996). Frames of Reference and Molyneux's Question: Cross-Linguistic Evidence. Language and Space. P. Bloom, M. Peterson, L. Nadel and M. Garrett. Cambridge, MIT Press.
- Levinson, S. (2003). Space in Language and Cognition: Explorations of Cognitive Diversity. Cambridge, University Press.
- Lisansky, J. (1981). Interpersonal Relations among Hispanics in the United States: A Content Analysis of the Social Science Literature. Technical Report #3. Urbana-Champaign, University of Illinois, Department of Psychology.
- Loftus, E. F., K. D. Smith, M. R. Klinger and J. Fiedler (1992). Memory and Mismemory for Health Events. Questions About Questions: Inquiries into the Cognitive Basis of Surveys. J. M. Tanur. New York, Russell Sage Foundation: 102-137.
- Longacre, R. E. (1956). "Review of "Language and Reality" by Wilbur M. Urban and "Four Articles on Metalinguistics" by Benjamin Lee Whorf." Language **32**: 298-308.
- Lopez, S. R. and P. J. J. Guarnaccia (2000). "Cultural Psychopathology: Uncovering the Social World of Mental Illness." Annual Review of Psychology **51**: 571-598.
- Lucy, J. A. (1992). Language Diversity and Thought: A Reformulation of the Linguistic Relativity Hypothesis. Cambridge, Cambridge University Press.
- Lucy, J. A. and S. Gaskins (2001). Grammatical Categories and the Development of Classification Preferences: A Comparative Approach. Language Acquisition and Conceptual Development. M. Bowerman and S. Levinson. Cambridge, MA, Cambridge University Press: 257-283.
- Lucy, J. A. and R. A. Shweder (1979). "Whorf and His Critics: Linguistic and Nonlinguistic Influences on Color Memory." American Anthropologist **81**: 581-615.

- Lucy, J. A. and J. V. Wertsch, Eds. (1987). Vygotsky and Whorf: A Comparative Analysis. Social and Functional Approaches to Language and Thought. Orlando, FL, Academic Press.
- Maass, A., D. Salvi, L. Arcuri and G. R. Semin (1989). "Language Use in Intergroup Contexts: The Linguistic Intergroup Bias." Journal of Personality and Social Psychology **57**: 981-993.
- Maddahian, E., M. D. Newcomb and P. M. Bentler (1988). "Adolescent Drug Use and Intention to Use Drugs: Concurrent and Longitudinal Analyses of Four Ethnic Groups." Addictive Behaviors **13**: 191-195.
- Majid, A., M. Bowerman, S. Kita, D. B. M. Haun and S. Levinson (2004). "Can Language Restructure Cognition? The Case for Space." Trends in Cognitive Sciences **8**(3): 108-114.
- Marian, V. and M. Kaushanskaya (2004). "Self-Constraint and Emotion in Bicultural Bilinguals." Journal of Memory and Language **51**: 190-201.
- Marian, V. and U. Neisser (2000). "Language-Dependent Recall of Autobiographical Memories." Journal of Experimental Psychology: General **129**(3): 361-368.
- Marin, G., R. J. Gamba and B. V. Marin (1992). "Extreme Response Style and Acquiescence among Hispanics." Journal of Cross-Cultural Psychology **23**: 498-509.
- Marin, G. and B. V. Marin (1991). Research with Hispanic Populations. Newbury Park, CA, Sage.
- Marin, G. and H. C. Triandis (1985). Allocentrism as an Important Characteristic of the Behavior of Latin Americans and Hispanics. Cross-Cultural and National Studies. R. Diaz Guerrero. Amsterdam, Elsevier: 85-104.
- Marin, G., H. C. Triandis, H. Betancourt and Y. Kashima (1983). "Ethnic Affirmation Versus Social Desirability: Explaining Disparities in Bilinguals' Responses to a Questionnaire." Journal of Cross-Cultural Psychology **14**(2): 173-186.
- Matsumoto, A. and C. Stanny (2006). "Language-Dependent Access to Autobiographical Memory in Japanese-English Bilinguals and U.S. Monolinguals." Memory **14**(3): 378-390.
- Mazzoni, S., P. Boiko, W. Katon and J. Russo (2007). "Depression and Disability in Seasonal and Migrant Hispanic Agricultural Workers." General Hospital Psychiatry **29**(5): 450 - 453.
- McConnell, A. R. and R. H. Fazio (1996). "Women as Men and People: Effects of Gender-Marked Language." Personality and Social Psychology Bulletin **22**: 1004-1013.
- Means, B., A. Nigam, M. Zarrow, E. F. Loftus and M. Donaldson (1989). Autobiographical Memory for Health-Related Events. Vital and Health Statistics. Washington, D.C., US Government Printing Office.
- Mendoza-Denton, R., Y. Shoda, O. Ayduk and W. Mischel (1999). Applying Cognitive-Affective Processing System (Caps) Theory to Cultural Differences in Social Behavior. Proceedings of Merging past, present, and future in cross-cultural psychology: Selected proceedings of the 14th

- International Congress of the International Association for Cross-Cultural Psychology, Lisse, the Netherlands, Swets and Zeitlinger.
- Menon, G. (1994). Judgements of Behavioral Frequencies: Memory Search and Retrieval Strategies. Autobiographical Memory and the Validity of Retrospective Reports. N. Schwarz and S. Sudman. New York, Springer-Verlag: 161-172.
- Ng, S. H. (1990). "Areocentric Coding of Man and His Memory by Language Users." Journal of Experimental Social Psychology **26**: 455-464.
- Nisbett, R. (2003). The Geography of Thought. How Asians and Westerners Think Differently. And Why. New York, Free Press.
- Okasha, A. and F. Lotalif (1979). "Attempted Suicide: An Egyptian Investigation." Acta Psychiatrica Scandinavica **60**: 69-75.
- Olson, D. H. (1986). "Circumplex Model Vii: Validation Studies and Faces Iii." Family Process **25**: 337-51.
- Olson, D. H. (1989). Circumplex Model of Family Systems Viii: Family Assessment and Intervention. Circumplex Model: Systemic Assessment of Treatment of Families. D. H. Olson, C. S. Russell and S. D. H. New York, Haworth Press.
- Ortega, A. N., R. Rosenheck, M. Alegria and R. A. Desai (2000). "Acculturation and the Lifetime Risk of Psychiatric and Substance Use Disorders among Hispanics." Journal of Nervous and Mental Disease **188**: 728-735.
- Ortiz, V. (1995). The Diversity of Latino Families. Understanding Latino Families: Scholarship, Policy and Practice. R. E. Zambrana. Thousand Oaks, CA, Sage.
- Oyserman, D., H. Coon and M. Kemmelmeier (2002). "Rethinking Individualism and Collectivism: Evaluation of Theoretical Assumptions and Meta-Analyses." Psychological Bulletin **128**(1): 3-72.
- Oyserman, D. and S. W. Lee (2007). Priming "Culture": Culture as a Situated Cognition. Handbook of Cultural Psychology. S. Kitayama and D. Cohen. New York, Guilford Press.
- Padilla, A. M. (1986). Acculturation and Stress among Immigrants and Later Generation Individuals. The Quality of Urban Life: Social, Psychological and Physical Conditions. D. Frick, H. Hoefert, H. Legewie, R. Mackensen and R. K. Silbereisen. Berlin, Germany, de Gruyter.
- Padilla, A. M. (1994). Bicultural Development: A Theoretical and Empirical Examination. Theoretical and Conceptual Issues in Hispanic Mental Health. R. G. Malgady and O. Rodriguez. Melbourne, FL, Robert E. Krieger: 20-51.
- Padilla, A. M., R. C. Cervantes and M. Maldonado (1988). Psychosocial Stress in Mexican Immigrant Adolescents.
- Pederson, E., E. Danziger, D. Wilkins, S. Levinson, S. Kita and G. Senft (1998). "Semantic Typology and Spatial Conceptualization." Language and Society **74**: 557-589.
- Phinney, J. and M. Devich-Navarro (1997). "Variations in Bicultural Identification among African American and Mexican American Adolescents." Journal of Research on Adolescence **7**: 3-32.

- Portes, A. (1992). "Mental Illness and Help Seeking Behavior among Mariel, Cuban and Haitian Refugees in South Florida." Journal of Health and Social Behavior **4**: 283-298.
- Rosenbaum, P. and D. B. Rubin (1983). "The Central Role of Propensity Score in Observational Studies for Causal Effects." Biometrika **70**: 41-55.
- Rosenbaum, P. and D. B. Rubin (1984). "Reducing Bias in Observational Studies Using Subclassification on the Propensity Score." Journal of the American Statistical Association **79**: 516-524.
- Ross, C. E. and J. Mirowsky (1984). "Socially Desirable Response and Acquiescence in a Cross-Cultural Survey of Mental Health." Journal of Health and Social Behavior **25**: 189-197.
- Ross, M., W. Q. E. Xun and A. Wilson (2002). "Language and the Bicultural Self." Personality and Social Psychology Bulletin **28**(8): 1040-1050.
- Rotheram-Borus, M. J. (1993). Biculturalism among Adolescents. Ethnic Identity: Formation and Transmission among Hispanics and Other Minorities. M. E. Bernal and G. P. Knight. Albany, NY, State University of New York Press.
- Sabogal, F., G. Marin and R. Otero-Sabogal (1987). "Hispanic Familism and Acculturation: What Changes and What Does Not?" Hispanic Journal of Behavioral Sciences **9**: 397-412.
- Sabogal, F., G. Marin, R. Otero-Sabogal, B. V. Marin and E. J. Perez-Stable (1987). "Hispanic Familism and Acculturation: What Changes and What Doesn't?" Hispanic Journal of Behavioral Sciences **9**: 397-412.
- Sapir, E., Ed. (1929). The Status of Linguistics as a Science. Culture, Language and Personality. Berkeley, University of California Press.
- Schrauf, R. W. (2000). "Bilingual Autobiographical Memory: Experimental Studies and Clinical Cases." Culture and Psychology **6**(4): 387-417.
- Schrauf, R. W. and D. C. Rubin (1998). "Bilingual Autobiographical Memory in Older Adult Immigrants: A Test of Cognitive Explanations of the Reminiscence Bump and the Linguistic Encoding of Memories." Journal of Memory and Language **39**: 437-457.
- Schrauf, R. W. and D. C. Rubin (2000). "Internal Languages of Retrieval: The Bilingual Encoding of Memories for the Personal Past." Memory and Cognition **28**: 616-623.
- Schwarz, N. (1990). Assessing Frequency Reports of Mundane Behaviors: Contributions of Cognitive Psychology to Questionnaire Construction. Research Methods in Personality and Social Psychology. C. Hendrick and M. S. Clark. Newbury Park, CA, Sage.
- Schwarz, N. (1996). Cognition and Communication: Judgmental Biases, Research Methods and the Logic of Conversation. Hillsdale, NJ, Erlbaum.
- Schwarz, N. (2003). Culture-Sensitive Context Effects: A Challenge for Cross-Cultural Surveys. Cross-Cultural Survey Methods. J. Harkness, F. J. R. Van de Vijver and P. Mohler. Hoboken, NJ, John Wiley and Sons.
- Schwarz, N. and H. J. Hippler (1991). Response Alternatives: The Impact of Their Choice and Ordering. Measurement Error in Surveys. P. Biemer, R. Groves, N. Mathiowetz and S. Sudman. Chichester, Wiley: 41-56.

- Shrout, P., M. Alegria, G. Canino, P. Guarnaccia, W. A. Vega, N. Duan and Z. Cao (2008). "Testing Language Effects in Psychiatric Epidemiology Surveys with Randomized Experiments: Results from the National Latino and Asian American Study." Practice of Epidemiology **168**(2).
- Shweder, R. A. and E. J. Bourne (1982). Does the Concept of the Person Vary Cross-Culturally? Cultural Conceptions of Mental Health and Therapy. A. J. Marsella and G. M. White. Boston, MA, Reidel: 97-137.
- Silva de Crane, R. and C. D. Spielberger (1981). "Attitudes of Hispanic, Black and Caucasian University Students toward Mental Illness." Hispanic Journal of Behavioral Sciences **3**: 241-255.
- Singer, E., N. Mathiowetz and M. P. Couper (1993). "The Impact of Privacy and Confidentiality Concerns on Survey Participation: The Case of the 1990 U.S. Census." Public Opinion Quarterly **57**: 465-482.
- Sloutsky, V. M. (2003). "The Role of Similarity in the Development of Categorization." Trends in Cognitive Sciences **7**: 246-251.
- Smith, A. F. and J. B. Jobe (1994). Validity of Reports of Long-Term Dietary Memories: Data and a Model. Autobiographical Memory and the Validity of Retrospective Reports. N. Schwarz and S. Sudman. Berlin, Springer-Verlag: 121-140.
- Smith, S. M. (1988). Environmental Context-Dependent Memory. Memory in Context: Context in Memory. G. M. Davies and D. M. Thomson. Chichester, England, Wiley.
- Stening, B. W. and J. E. Everett (1984). "Response Styles in Cross-Cultural Managerial Study." Journal of Social Psychology **122**: 151-156.
- Strube, G. (1987). Answering Survey Questions: The Role of Memory. Social Information Processing and Survey Methodology. H. J. Hippler, N. Schwarz and S. Sudman. New York, Springer-Verlag: 86-101.
- Substance Abuse and Mental Health Services Administration (2001). Summary of Findings from the 2000 National Household Survey on Drug Abuse. NHSDA Series H-13. D. P. [SMA]01-3549. Rockville, MD, Office of Applied Studies.
- Sudman, S. and N. Bradburn (1973). "Effects of Time and Memory Factors on Response in Surveys." Journal of the American Statistical Association **68**: 805-815.
- Sudman, S., N. Bradburn and N. Schwarz (1996). Thinking About Answers: The Application of Cognitive Processes to Survey Methodology. San Francisco, Jossey-Bass.
- Thompson, C. P., J. J. Skowronski, S. F. Larsen and A. L. Betz (1996). Autobiographical Memory. Mahwah, NJ, Erlbaum.
- Tourangeau, R., K. Rasinski, J. B. Jobe, T. W. Smith and W. Pratt (1997). "Source of Error in a Survey of Sexual Behavior." Journal of Official Statistics **13**: 341-365.
- Tourangeau, R., L. J. Rips and K. A. Rasinski (2000). The Psychology of Survey Response. Cambridge, U.K.; New York, Cambridge University Press.
- Trafimow, D., E. S. Silverman, R. M. T. Fan and J. S. F. Law (1997). "The Effect of Language and Priming on the Relative Accessibility of the Private Self

- and the Collective Self." Journal of Cross-Cultural Psychology **28**: 107-123.
- Triandis, H. C. (1995). Individualism and Collectivism. Boulder, CO, Westview Press.
- Triandis, H. C., E. E. Davis, V. Vassiliou and M. Nassiakou (1965). Some Methodological Problems Concerning Research on Negotiations between Monolinguals. D. o. Psychology. Urbana, University of Illinois.
- Triandis, H. C., H. Marin, J. Lisansky and H. Betancourt (1984). "Simpatia as a Cultural Script for Hispanics." Journal of Personality and Social Psychology **47**: 1363-1375.
- Tsai, J., Y. Ying and P. Lee (2000). "The Meaning of "being Chinese" and "being American" : Variation among Chinese-American Youth Adults." Journal of Cross-Cultural Psychology **31**: 302-332.
- Tulving, E. and D. Thompson (1973). "Encoding Specificity and Retrieval Processes in Episodic Memory." Psychological Review **80**: 352-373.
- Turner, C. F., L. Ku, S. M. Rogers, L. D. Lindberg, J. H. Pleck and F. L. Sonenstein (1998). "Adolescent Sexual Behavior, Drug Use, and Violence: Increased Reporting with Computer Survey Technology." Science **280**: 867-873.
- Vasquez, M. J. T. (1994). Latinas. Women of Color: Integrating Ethnic and Gender Identities in Psychotherapy. L. Comas-Diaz and B. Greene. New York, Guilford.
- Vega, W. A. (1990). "Hispanic Families in the 1980s: A Decade of Research." Journal of Marriage and the Family **52**: 1015-24.
- Vega, W. A., A. G. Gil, G. J. Warheit, R. S. Zimmermann and E. Apospori (1993). "Acculturation and Delinquent Behavior among Cuban-American Adolescents: Toward an Empirical Model." American Journal of Community Psychology **21**: 113-25.
- Vega, W. A., B. Kolody, S. Aguilar-Gaxiola, E. Alderete, R. Catalano and J. Carveo-Anduaga (1998). "Lifetime Prevalence of Dsm-iii-R Psychiatric Disorders among Urban and Rural Mexican Americans in California." Archives of General Psychiatry **55**: 771-778.
- Wagenaar, W. A. (1986). "My Memory: A Study of Autobiographical Memory over Six Years." Cognitive Psychology **18**: 225-252.
- Warnecke, R. B., T. P. Johnson, N. Chavez, S. Sudman, D. O'Rourke, L. Lacey and J. Horn (1997). "Improving Question Wording in Survey of Culturally Diverse Populations." Annals of Epidemiology **7**(5): 334-342.
- Wassmann, J. and P. R. Dasen (1998). "Balinese Spatial Orientation: Some Empirical Evidence of Moderate Linguistic Relativity." Journal of the Royal Anthropological Institute **4**: 689-711.
- Watkins, D. and A. Gerong (1999). "Language of Response and the Spontaneous Self-Concept: A Test of the Cultural Accommodation Hypothesis." Journal of Cross-Cultural Psychology **30**: 115-121.
- Watkins, D. and M. Regmi (2002). "Does the Language of Response Influence Self-Presentation? A Nepalese Test of the Cultural Accommodation

- Hypothesis." Psychologia: An International Journal of Psychology in the Orient **45**: 98-103.
- Watson, J. B. (1925). Behaviorism. New York, W.W. Norton.
- Weingartner, H. (1978). Human State-Dependent Learning. Drug Discrimination and State-Dependent Learning. B. T. Ho, D. W. Richards and D. C. Chute. New York, Academic Press: 361-382.
- Whitney, P. (1998). The Psychology of Language. Boston, MA, Houghton Mifflin.
- Whorf, B. L. (1940). "Science and Linguistics." Technology Review **42**(6): 229-231.
- Whorf, B. L. (1956). Language, Thought, and Reality: Selected Writings of Benjamin Lee Whorf. New York, Wiley.
- Widlock, T. (1997). "Orientation in the Wild: The Shared Cognition of the Haiom Bushpeople." Journal of the Royal Anthropological Institute **3**: 317-332.
- Williams, M. D. and J. D. Hollan (1981). "The Process of Retrieval from Very Long Term Memory." Cognitive Science **5**: 87-119.
- Yang, K. S. and M. H. Bond (1980). "Ethnic Affirmation by Chinese Bilinguals." Journal of Cross-Cultural Psychology **11**: 411-425.
- Zax, M. and S. Takahashi (1967). "Cultural Influences on Response Style: Comparison of Japanese and American College Students." Journal of Social Psychology **71**: 3-10.