Are Do Not Include Statements Helpful? October 2009

Monica Dashen, ¹ U.S. Bureau of Labor Statistics, 2 Massachusetts Ave., NE Washington DC 20212

¹ The opinions expressed in this paper are those of the author and do not reflect official policy of the Bureau of Labor Statistics. The author gratefully acknowledges the contributions of Stella Godbolt in data collection and entry. This work also profited greatly from discussions with Karen Goldenberg, Jake Kane, Bill Mockovak, and Mark Sauer.

Abstract

The present work investigated how helpful "do not include" statements (DNI), instructions to exclude certain items from categories, are to respondents' decision making. Study participants were asked to consult an industry description and determine whether the descriptions accurately reflected a business described in an accompanying scenario. Some industry descriptions contained "do not include" statements while others did not. Accuracy measures were tabulated. The observed findings indicate that when DNI statements were present, respondents made more correct rejections. By contrast, respondents made more correct acceptances *without* DNI statements.

Key Words: Questionnaire Design or Survey Methods

1. Introduction

The goal of this work is to determine how helpful "do not include" statements (DNI), instructions to exclude certain items from categories, are to respondents' decision making. DNI statements, which are also referred to as "exclude" statements, can occur anywhere in a survey and are intended to guide respondents so that their answers are as intended by survey designers. DNI statements are prevalent in both household surveys (e.g., the decennial Census) and establishment surveys (e.g., Annual Refiling Survey).

Survey designers intend DNI statements to help respondents better understand category definitions, thereby eliminating or at least reducing misclassification errors. When they work as intended, DNI statements help define categories by indicating what the category is *not*. However, there is the possibility that DNI statements may confuse respondents, leading them to be more likely to reject items that do belong to a category. Of particular interest are DNI statements that rely on percentages. For example, a survey that seeks to determine business categories might instruct a proprietor to exclude her business from the "hair salon" industry category if manicures and other nail services account for more than 50% of the business. The goal here is to find out whether DNI statements help or hurt—or both—performance on verification tasks.

At this point, it is important to understand that there are two types of DNI statements:

- (1) All or None: A DNI statement that excludes the item from the category based on *any* presence of some quality, service, or factor. For example, suppose the question is whether a particular business should be classified in the industry category "beauty salon." An all-or-none DNI statement might instruct a respondent *not* to include the business in the category if *any* part of the business consists of giving manicures.
- (2) Partial: A DNI statement that excludes the item from the industry category based on *a certain percentage or amount* of some quality, service, or factor. For example, a partial DNI statement might instruct the respondent to include the business above in the "beauty salon" category if less than 49% of the business consists of manicures. In contrast, a partial DNI statement might instruct the respondent to exclude the business above in the "beauty salon" category if more than 50% of the business consists of manicures.

This work seeks to determine the degree to which the Partial DNI statements help—or hinder—respondents' accuracy rates in yes/no decisions. The present work will not focus on All or None DNI statements. In the remainder of this paper, I will only refer to partial DNI statements.

A verification task is a vehicle for the examination of the role of Partial DNI statements in decision making. In such a task, respondents are asked to determine whether something, (e.g., a business) is still a member of a category. There are four response categories: (1) correct acceptances (yes when it really was yes), (2) correct rejections (no when it really was no), (3) incorrect acceptances (yes when it really was no) and (4) incorrect rejections (no when it really was yes).

These four possible response categories can be divided into "similar" and "different." The former means that DNI statements either help people make correct acceptances and correct rejections or tend to lead respondents to make incorrect acceptances and incorrect rejections. "Different" means that DNI statements help with one category (correct acceptances or correct rejections) and promote incorrect responses in the other category.

The present work uses the Annual Refiling Survey (ARS), an establishment survey, to examine the effects of Partial DNI statements in a verification task involving decision accuracy. In this survey, industry descriptions include DNI statements. Respondents

see written descriptions of the industries that correspond to the North American Industry Classification System (NAICS) code currently assigned to their businesses. These descriptions consist of general statements characterizing economic activity followed by lists of included examples and excluded items. The include statements provide examples; the exclude statements list items. Respondents are asked to indicate whether their business correspond to the relevant statements. The data collected from the ARS are used to update the industry classifications based on the NAICS code. The ARS is designed to catch any changes in business activities. For example, a beauty shop may report an increase of 60 % more nail services than the previous reporting period (4-years ago) where only 10% of the business activities were devoted to nail services. This increased percent in nail services requires a reclassification of the business from a beauty salon to a nail salon industry.

This research approach differs from other establishment studies where researchers generally visit the respondents at their place of business and make inquiries about their business activities (e.g., Phipps et al., 1993; Eldridge et al., 2000). The more familiar methodology would not meet the needs of our study because respondents are expected to provide an answer for more than one industry. It would be unreasonable to expect individual ARS respondents to answer any questions about a business or industry other than their own.

In addition, the present work looks at the NAICS code assigned. One of the most common statistical tasks of businesses and government research agencies is classification: the organization of units into hierarchies of categories. For instance, the U.S. Bureau of Labor Statistics classifies business activities by NAICS code for the purposes of calculating standardized business statistics. Industrial data are widely used by researchers in government, business, and academia who need a comprehensive, up-to-date, and flexible system. U.S. federal statistical agencies utilize the NAICS in surveys of businesses to assess which areas of the economy are growing and which are stagnating.

Many researchers have looked at classification systems in establishment surveys and decision making. For example, Conrad and Couper (2001) and Malhotra and Krosnick (2007) looked at how respondents classified occupations in the Standard Occupational Classification scheme. These researchers found that length and ambiguity of the definition affected time taken to complete the question. The presence of the DNI statements was not directly manipulated in these studies. The present work seeks to extend these studies by measuring the effect of DNI statements.

2. Overview

The present work seeks to determine whether "do not include" statements help respondents make correct classification decisions. Participants from the general public were asked to consult the industry description on the ARS forms and determine whether the descriptions accurately reflected a business described in an accompanying scenario. Some industry descriptions contained "do not include" statements while others did not, a difference that allowed me to see whether the presence of such statements helped respondents make more correct decisions. In this study, I did not manipulate the Include statements, as these statements represent an integral part of the definition.

3. Methodology

3.1 Participants.

Thirty-one participants (13 males and 18 females) responded to an advertisement in a local area newspaper (Washington DC, USA) and received \$40.00 each in compensation for their participation. The participants' mean age was 41.3, and their average educational level was 15 years of schooling (or three years of college).

3.2 Experimental Design.

This is a two-by-two experimental design wherein Industry Description and Scenario are categories for which the presence of DNI statements was manipulated. The first factor, Industry Description, refers to those descriptions currently printed on the ARS form and has two values: original (the current industry descriptions as they appear on the ARS form) and test (the same after DNI statements have been removed). Respondents saw both original and test conditions but for different industries. That is, they either always saw the original or always saw the test condition for a specific industry. Since nothing else in the industry description changed, this structure allows me to evaluate the effect of the "Does not include" statement.

The second factor, Scenario, refers to descriptions of businesses and has two values: straightforward and complex. The "straightforward" scenario value reflects information from the "include" side of the NAICS industry definition (description or examples). The other scenario value is "complex," in that it incorporates information from both the include and "Does not include" side of the NAICS industry definition. We used the complex scenario because in the real world many businesses encompass a variety of products, services, or activities, some of which could overlap with other NAICS classifications. ²

² For example, using the *Shoe* stores, a straightforward scenario that is correctly coded in this industry might describe a store that sells athletic shoes, boots, sandals, slippers. A straightforward scenario that is not correctly coded for this industry could describe a business that sells only golf supplies or hosiery—that is, the business does something different from what is described, but is not specifically addressed by the exclusions. A complex scenario that is correctly coded might be a business that sells sandals (30% of total products sold), athletic shoes (45%), and boots (25%), while a complex scenario that does not fit the description could be a shoe store or a store that sells golf shoes (60%) and other golf supplies (40%). What makes these situations complex is that they refer to one of the

Each industry description has a corresponding straightforward and complex scenario, as seen in Appendix A. Not all scenarios fit the Industry Description. For this reason, study participants should reject some of the industry descriptions based on the scenario. The Y or N at the end of each scenario, in Appendix A, indicates whether the description should be accepted (Y) or rejected (N), given the scenario.

3.3 Materials.

3.3.1 *Industries*.

24 industries were chosen. The selection of industries to test was a key component of the research. (See Appendix A, for a sample form.) My goal in industry selection was to ensure that the descriptions were relatively homogeneous thereby increasing the likelihood that the results would be attributed to the experimental condition and not to an uncontrolled attribute of the industry description. The industries chosen were based on the following criteria:

- 20,000 or more establishments in an industry (fairly large sized industries), as denoted by the *Employment and Wages: Annual Averages* (BLS, 2003). Two exceptions to this criterion were the Facilities Support Services and Continuing Care Retirement Communities industries.
- 60-100 total number of words in the description (as counted by Microsoft Word) were selected.
- 2 to 16 include statements in the industry description. (We chose not to use the number of "excludes" as a criterion, because the number varied widely and nonlinearly relative to the "includes.")

Scenarios. Forty-eight scenarios, two per industry, accompany industry descriptions. (See Appendix B, for a sample scenario.) The scenarios help create a "reality" for respondents. Without scenarios, respondents familiar with only one or a few industries might be unable to answer questions about those industries with which they are not familiar. Moreover, scenarios are particularly useful because researchers know the correct answers for each scenario. We control the situations described, the extent to which they do or do not match the industry definitions, and the amount of ambiguity in the situations. Because respondents are members of the general public, though, it is important to ensure that the scenarios deal with familiar businesses and everyday situations. The present work's use of scenarios fits the recent trend of using scenarios in establishment surveys (Goldenberg et al., 2002; Goldenberg, 1998; Morrison et al., 2004; Stettler et al, 2000). Until recently, scenarios have been used primarily in household surveys.

3.4 Procedure.

There are two phases in this study, as follows:

Phase 1. Reviewing Instructions. Each participant received instructions and an informed consent form. The instructions pertained to the Mock ARS forms corresponding to each scenario. (See Appendix A, for more details about the actual Mock ARS survey, including a Shoe industry description.) All respondents read instructions about the forms and scenarios. (See Appendix B, for a corresponding Shoe scenario.) To ensure that the respondents understood the task at hand, the experimenter summarized the instructions after the respondents had read them and encouraged the respondents to ask any questions about the study before getting started.

Phase 2. Completing Mock ARS Forms and Scenario Packets. Participants received 48 packets containing Mock ARS forms and corresponding scenarios and were expected to complete each packet on their own one at a time (self-administered). Participants were discouraged from returning to their previously answered forms. The Mock ARS forms show a fictitious business at a specific physical location in the fictitious state of Utana. The industry questions contain the industry description corresponding to the scenario, in either the original or the test condition. While participants received the entire form, the questions following the industry description were greyed out so as not to distract the respondent. Appendix B shows sample mock ARS forms for the original and test conditions. The packet also contains the worksheets with a scenario, as well as a self-administered difficulty rating. The scenarios pertained to corresponding industry description. The forms also contained a self-administered difficulty rating task. We asked participants to rate each industry verification task from 1 to 5, where "1" is very easy and "5" is very hard, before going on to the next scenario. A randomization procedure was used to allocate industries and scenarios so that participants saw either the original or the test version of each description. This procedure ensured that participants did not see both scenarios for an industry together, scattered correct and incorrect descriptions, and helped to provide a good mix of complex and straightforward scenarios.

4. Results

Preliminary analysis was performed to assess any difficulties that the respondents may have. One analysis involved the mean proportion correct ranged from the lowest of .15 (Furniture Store) to the highest.87 (Shoes). Further analyses indicate that the low performance of 15% on the *Furniture Store* was atypical, probably the result of respondents' failure to understand the industry description. For this reason, the *Furniture Store* was eliminated from future analyses. Another analysis involved, 21 out of 24 industries had a mean proportion correct above 50%. These two findings indicate that this task was neither too difficult nor too easy. The most likely explanation for the low performing categories is the unclear definitions.

Recall that not all of the scenarios were created to match the ARS forms descriptions. I separated the scenarios according to whether the correct response for the ARS form was "Yes" or "No." Table 1 summarizes the mean proportion correct by verification task, scenario, and industry description.

	Scenario					
	Straightforward		Complex			
	Original	Test	Original	Test		
Verification Task						Mean Verif. Task total
Correct Acceptance	0.70	0.65	0.61		0.77	.68
Correct Rejection	0.73	0.69	0.71		0.57	.68
Mean Straightforward	.69		Mean Complex		.67	

There are two points about the role of DNI statements in Table 1 above. These two points were based on a three-way repeated measures ANOVA where scenario, industry description and verification task were the factors. This repeated measure ANOVA yielded a significant interaction of $\underline{F}(1,30)=8.05$, $\underline{p}<.01$. All t-test comparisons (including a Bonferroni correction) are reported below:

- 1. Correct Acceptances. I found that when a respondent verified the existing information (i.e., the correct acceptance) in industry description, there was a tendency for the test version of the industry description to yield a greater number of correct acceptances for the complex scenario (.77) than the original industry description (.61); <u>t</u> (30)=2.53 p<.05. That is, for the complex scenario, the absence of the DNI statement seemed to produce better data, a result consistent with my expectations. Recall that the complex scenario consists of both include and exclude statements while the straightforward scenario consists of include only. The fact that the difference was observed only in the complex scenario, where DNI statements are referred to, suggests that the exclude statements confused respondents when making a correct verification.
- 2. <u>Correct Rejections</u>. I found that when a respondent correctly rejected an industry description, there was a tendency for the original version of the industry description (.71) to yield a greater number of a correct rejections than the test version (.57), t(30)=3.22, p<.05 (This finding however was not the case for the straightforward scenario.) That is, for the complex scenario, the presence of the Does Not Include seemed to produce a large number of correct rejections than the complex scenarios with out DNI statements. No other meaningful and significant comparisons were found for the correct rejections portion of the table.

One possible hypothesis for these two findings is that respondents simply do not read carefully enough or fail to read DNI statements. I can reject this hypothesis, however, because, were it true, there should be no difference between the original and test scenarios. By omitting the DNI statements, the test provides a control and further indicates that, because of the significant differences in the mean accuracies, DNI statements did have an effect. This effect is inconsistent with the possibility that respondents did not read the statements.

Future research on this topic should focus on reaction times to measure how long people are spending on DNI statements and determine the ambiguity of the statements. Previous work by Conrad and Couper (2001) and Malhotra and Krosnik (2007) has shown that the more ambiguous a DNI statement, the longer the reaction time.

These findings have practical implications for survey designers. Because of the mixed results, care should be exercised when DNI statements are used. DNI statements are likely to be helpful when there is a strong possibility of an incorrect industry code. For example, DNI statements are likely to help for similar industries whose activities fluctuate considerably as, for example, might be the case for a hair and nail salon in which the relative proportions of the activities change. In contrast, DNI statements are generally not helpful when a business has been given a correct NAICS code and performs both Include and Do Not Include activities. For example, if a beauty salon performs nail services, it should be classified as a beauty salon provided the nail services do not constitute the majority of the business's activity. However, a respondent whose beauty salon has been correctly classified may erroneously change the classification in reliance on a DNI statement about nail services—even though nail services make up less than half of the business. To remedy this situation, I recommend underscoring the importance of the percentage rule (e.g., 51%) where an establishment that performs a DNI activity less than a certain percentage (e.g., 51%) of the time can still be a member of the industry.

The findings from the present work may apply to other countries that have a similar sort of survey involving industry classification updates. In the United Kingdom, the Office of National Statistics uses the Business Register Employment Survey (BRES), in which respondents write in their establishments' activities and these activities are coded later to update the European NACE system and Standard Industrial Classification system. In addition, Canada uses the Quality Assurance Survey (QAS) to verify the corresponding NAICS code. Of course, the QAS is not the only Canadian survey involved in updating industrial activity tabulations.

5. Acknowledgements

The opinions expressed in this paper are those of the author and do not reflect official policy of the Bureau of Labor Statistics. The author gratefully acknowledges the contributions of Stella Godbolt in recruiting respondents, data entry and other research assistance. This work also profited greatly from discussions with Karen Goldenberg, Bill Mockovak, and Mark Sauer.

6. References

- Conrad, F.G., and Couper, M.P., (2001) Classifying Open-Ended Reports: Coding Occupation in the Current Population Survey. Paper presented at the *Federal Committee on Statistical Methodology (FCSM) Research Conference*, Arlington, VA, November.
- Eldridge, Jack, Jean Martin, and Amanda White. 2000. "The Use of Cognitive Methods to Improve Establishment Surveys in Britain." *ICES II: The Second International Conference on Establishment Surveys. Invited papers.* Alexandria, VA: American Statistical Association, pp. 307-316.
- Goldenberg, Karen. 1998. "Results of NCA Form Test." Unpublished report, U.S. Bureau of Labor Statistics.
- Goldenberg, Karen L., Diane K. Willimack, Sylvia K. Fisher, and Amy E. Anderson. 2002. "Measuring Key Economic Indicators in U.S. Government Establishment Surveys." Paper presented at the International Conference on Improving Surveys, Copenhagen, Denmark.
- Malhotra, N. and Krosnick, J. Procedures for Updating Classification Systems: A Study of Biotechnology and the Standard Occupational Classification System. *Journal of Official Statistics*, 23 (3): 409-432.
- Morrison, Rebecca L., Kristin Stettler, and Amy E. Anderson. "Using Vignettes in Cognitive Research on Establishment Surveys." *Journal of Official Statistics* 20 (2): 319-340.
- Murphy, John B. 1998. "Introducing the North American Industry Classification System." Monthly Labor Review 121 (7): pp. 43-47.
- Phipps, Polly A., Shail Butani, and Young I. Chun. 1993. "Designing Establishment Survey Questionnaires." *BLS Statistical Note Number 35*. Washington, D.C.: U.S. Bureau of Labor Statistics.
- Stettler, Kristen, Rebecca L. Morrison, and Amy E. Anderson. 2000. "Results of Cognitive Interviews Studying Alternative Formats for Economic Census Forms." Paper presented at the Second International Conference on Establishment Surveys, Buffalo, NY.
- Walker, James A. and John B. Murphy. 2001. "Implementing the North American Industry Classification System at BLS." Monthly Labor Review 124 (12):15-21.
- U.S. Office of Management and Budget (OMB). 1997. North American Industry Classification System. United States, 1997. Washington, DC: Executive Office of the President.

- U.S. Office of Management and Budget (OMB). 2002. North American Industry Classification System. United States, 2002. Washington, DC: Executive office of the President.
- U.S. Bureau of Labor Statistics. 2003. Employment and Wages. Annual Averages, 2002. Washington, DC.

Appendix A

Sample Mock ARS Form

	HANNAH BANANA 3 FLORIDA WAY KEYS, UA 3265-9876
	KE13, GA 3203 3070
3	We need the name and direct mailing address for the business using this Unemployment Insurance account, regardless of who prepares the form. This information does not affect mailings for tax purposes. Are the name and mailing address shown in Item 2 correct for the business using this Unemployment Insurance account?
	☐ YES☐ NO Please print corrections or additions to the right of the printed address in Item 2. ☐
	this form
	In addition to your mailing address, please tell us where your business is physically located (street and number). The physical location address is the place where you conduct your business and receive deliveries, so it cannot be a Post Office Box or a rural route number.
	Our records show that this business in Utana is physically located at:
	3 FLORIDA WAY KEYS, UA 3265-9876
	Is this address correct for the location in Utana?
5	Is the following information correct for the address in Item 4? UTANA COUNTY: WATERCRESS YESContinue with Item 6
	NOPlease print corrections in this space and then continue with Item 6
6	According to our records, the business operating under Unemployment Insurance account 1234567890 in Utana mainly provides goods and services to the general public. Is this correct? ("The general public" includes individual consumers, other businesses, and organizations.)
[[] YES, we MAINLY provide goods and services to the general public [] NO, we are part of a larger company and we MAINLY support other locations of OUR company
7	Does this business have a website?
	YESPlease enter your website address hereContinue with Item 8
8	Does the business using Unemployment Insurance account 1234567890 IN UTANA have only one physical location in this state? (Do not count client sites or offsite projects that will last less than a year.)
	 YES (One physical location)Continue with Item 9 on the back NO (More than one physical location) Please attach a separate sheet. For each site, (1) list physical location address, (2) show number of employees, and (3) answer Items 6 and 9 - 11. Continue with Item 9
	PLEASE CONTINUE WITH ITEM 9 ON THE BACK OF THIS PAGE.
	OFFICE USE FY05 02/08/2005

EMPI

This report is mandatory under Section 320.5 of the Utana Unemployment Insurance Code and Section 320-1

29 U.S.C. 2. Your cooperation is needed to make the results of this survey complete, accurate, and

Title 22 of the Utana Code of Stergunlant Somery Restainth Methodsizes Mby009aw,

1234567890 IN UTANA.

The questions on this form concern the work location(s) using Unemployment Insurance account number

9		s show that the <i>main</i> activity of the business using U.I. number 0 in UTANA is:
:	Examples ir * Athletic * Boots * Galoshes * Sandals * Tennis sh * Ballet sl * Children' * Men's dre * Slippers	noes .ippers s shoes
		ICLUDE retailing hosiery. ICLUDE retailing specialty sports shoes such as bowling shoes, golf shoes, or spiked shoes.
10	during the activity.)	may not do everything listed above, does the information in Item 9 accurately describe the <i>main</i> business in Utana past 12 months? (If the business has been closed, sold, or moved out of this state, please answer in terms of its forme <i>Please SKIP to Item 12</i>
	NO	Continue with Item 11
	main busing Then give utotal 100%.	etailed information to assign the correct industry code to this business. In the space provided below, describe your ess activities, goods, products, or services in this state, as though you were telling a prospective employee what you do us the approximate percentage of sales or revenues resulting from each item. See examples below. Percentages should be used third party agent for the business named in Item 2, such as a payroll service or accountant, please review with your client.
	distri EXAI	products: What are they, and what do you do with them? Do you design, manufacture, sell directly to consumers, bute to wholesalers, install, repair, or do something else with them? What are these goods or products made of? MPLE 1: <i>Major appliances: Sell to public 40%; Sell to retailers 30%; Repair 30%</i> EXAMPLE 2: <i>Install fiber optic cable</i>
		urers: What are your main products? What are your most important materials? What are the main production methods MPLE: Weaving cotton broadwoven fabrics 80%; Spinning cotton threads 20%
	broker: mana	agement, or similar services, what are your major activities?
		MPLE 1: Hair cutting & styling 65%; Manicures 25%; Facials 10% EXAMPLE 2: Long distance trucking, less than truckl
	100% EXAI 100%	MPLE 3: Marketing consulting: Planning strategy 60%, Sales forecasting 40% EXAMPLE 4: Cleaning private homes
	Construct	tion or Building Trades: Is the work mostly residential or nonresidential? Single- or multi-family? New or remodeling? MPLE: Electrical contractor: Wiring new homes 51%; Electrical refurbishing of office buildings 49%
	List most	
	%	
	important	
	%	
	activities	
	%	
	100%	PLEASE PRINT CLEARLY

Appendix B: Sample Shoe Store Scenario

NAICS	Industry	Scenario Type:	Scenario Type: Complex
448210	In-store retail sales of all types of new footwear, except hosiery and specialty sports shoes. Examples include, but are not limited to: * Athletic shoes * Boots * Galoshes * Sandals * Tennis shoes * Ballet slippers * Children's shoes * Men's dress shoes * Slippers * Women's casual shoes DOES NOT INCLUDE retailing hosiery. DOES NOT INCLUDE retailing specialty sports shoes such as bowling shoes, golf shoes, or spiked shoes.	Hannah Banana is a children's clothing store that features bananarelated apparel. Her two biggest sellers are bananashaped purses and banana-printed sweatshirts. Grandmothers and young girls are amongst Hannah's biggest fans. N	Shoe Train Inc. is a children's shoe store. While selling shoes, Mr. Small, the owner wears an engineer hat, stripped overalls, and whistle. The kids call him "Engineer Small." The floor of which "Engineer Small" presides over looks like a train engine where the customer chairs and salespeople stools are located near the front. "Engineer Small" is considering selling soccer, baseball shoes or athletic shoes, in children's sizes due to the increasing demand of his customers.

^{*} Based on the scenarios, study participants should reject the corresponding industry description. The Y or N at the end of the scenario indicates whether the description should be accepted (Y) or rejected (N), given the scenario.