



# HHS Public Access

Author manuscript

*Med Decis Making*. Author manuscript; available in PMC 2019 August 01.

Published in final edited form as:

*Med Decis Making*. 2018 August ; 38(6): 673–682. doi:10.1177/0272989X18785159.

## Reducing “I Don’t Know” Responses and Missing Survey Data: Implications for Measurement

Deanna C. Denman<sup>1</sup>, Austin S. Baldwin<sup>1</sup>, Andrea C. Betts<sup>2,3,4</sup>, Amy McQueen<sup>5</sup>, and Jasmin A. Tiro<sup>2,3</sup>

<sup>1</sup>Department of Psychology, Southern Methodist University, Dallas, Texas

<sup>2</sup>Department of Clinical Sciences, University of Texas Southwestern Medical Center, Dallas, Texas

<sup>3</sup>Harold C. Simmons Comprehensive Cancer Center, Dallas, Texas

<sup>4</sup>Department of Health Promotion and Behavioral Sciences, UTHealth School of Public Health in Dallas, Texas

<sup>5</sup>Division of General Medical Sciences, School of Medicine, Institute for Public Health, Washington University in St. Louis

### Abstract

**Background**—“I don’t know” (DK) responses are common in health behavior research. Yet analytic approaches to managing DK responses may undermine survey validity and researchers’ ability to interpret findings.

**Objective**—Compare usefulness of a methodological strategy for reducing DK responses to three analytic approaches: (1) excluding DKs as missing data, (2) recoding them to the neutral point of the response scale, and (3) recoding DKs with the mean.

**Methods**—We used a four-group design to compare a methodological strategy, which encourages use of the response scale after an initial DK response, to three methods of analytically treating DK responses. We examined 1) whether this methodological strategy reduced the frequency of DK responses, and 2) how the methodological strategy compared to common analytic treatments in terms of factor structure and strength of correlations between measures of constructs.

**Results**—The prompt reduced DK response frequency (55.7% of 164 unprompted participants vs. 19.6% of 102 prompted participants). Factorial invariance analyses suggested equivalence in factor loadings for all constructs across groups. Compared to excluding DKs, recoding strategies and use of the prompt improved the strength of correlations between constructs, with the prompt resulting in the strongest correlations (.589 for benefits and intentions, .446 for perceived susceptibility and intentions .329 for benefits and perceived susceptibility).

---

Correspondence concerning this article should be addressed to Jasmin A. Tiro, Department of Clinical Sciences, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd. Dallas, TX 75390-9066; Phone: (214) 648-0263; [jasmin.tiro@utsouthwestern.edu](mailto:jasmin.tiro@utsouthwestern.edu).

Findings from this study were presented at the 38<sup>th</sup> Annual Meeting of the Society of Behavioral Medicine in San Diego, CA in 2017.

**Limitations**—This study was not designed *a priori* to test methods for addressing DK responses. Our analysis was limited to an interviewer-administered survey, and interviewers did not probe about reasons for DK responses.

**Conclusion**—Findings suggest use of a prompt to reduce DK responses is preferable to analytic approaches to treating DK responses. Use of such prompts may improve validity of health behavior survey research.

---

When people are asked to report their attitudes and beliefs in surveys about health issues (e.g., knowledge, risk), “I don’t know” (DK) is a common response that may undermine survey validity (1). Missing data due to survey nonresponse is an issue that is pervasive across survey research (2,3). Managing DK responses analytically can be done through different approaches (2–5). For example, DK responses can be excluded as missing data, presumed to mean “neither” and recoded as a neutral midpoint in a Likert response scale (6), or treated as a meaningful categorical response (5). However, there is no clear best practice, and each approach has limitations.

Most commonly, DK responses are treated as missing data and excluded from analysis (1,7,8). Extensive DK reporting therefore results in large amounts of missing data and threatens the validity of study findings. For example, excluding DK responses from analyses reduces statistical power and could increase the chance of making a Type II error (i.e., failure to detect an effect that is present) (4). Also, exclusion of DK responses could bias the study sample. Literature reviewed here mainly focuses on health behavior research and risk perception. In these studies, individuals who provide DK responses tend to have lower health literacy and are at greater risk of health disparities (1,6,7,9). DK responses are also associated with lower education, lower numeracy, lower income, and minority status (1,6). Systematically excluding DK responses therefore hinders generalizability of research findings and ability to understand whether psychosocial constructs differentially influence behavior for populations at greatest risk for disease.

Although alleviating some of these issues, recoding DK responses through single imputation (i.e., into the neutral or mean-level) or treating them as a meaningful categorical variable presents conceptual and interpretive challenges. Although DK responses have traditionally been thought to reflect a lack of opinion, research suggests that a number of factors may prompt an individual to provide a DK response (5,10). These include insufficient knowledge of the disease, difficulty processing probabilities, fear of providing an inaccurate answer, concern their response will not be confidential, or lack of motivation to cognitively process the question and formulate a response (5,6,11–13). Some literature suggests that when trying to express uncertainty (such as when asked to estimate probabilities), participants use “50–50” to suggest they do not know (6). Other work has indicated that participants interpret and use the midpoint category of scales to suggest “neutral” or “uncertain” (14). Previous work has suggested that some constructs from popular health behavior theories (e.g. Theory of Planned Behavior [TPB] and Health Belief Model [HBM] (15,16)), such as behavioral intentions, may be less susceptible to DK responses than perceived susceptibility (or perceived risk) (6). In the context of perceived risk for disease, DK responses may represent meaningful uncertainty or a legitimate lack of knowledge regarding their risk (9,17). In

contrast, selection of DK for behavioral intentions may indicate a lack of thinking about this topic or an unwillingness to provide an opinion. Thus, use of DK may be qualitatively different when responding to perceived risk items, and uniform interpretation of DK responses is likely not justifiable as some responses are meaningful but others are not. Implementing a prompt in response to DK responding could be beneficial in reduction of non-meaningful DK responses. However, use of prompts has at times been questioned due to the possibility of influencing participants responses or forcing them to guess (18). Although methodological strategies to minimize non-meaningful DK responses (e.g. probing) are generally recommended over analytical approaches, few have been tested for their impact on the validity of findings (5,19).

We had the opportunity to compare a methodological strategy to prevent DK responses to three common analytic practices for treating DK responses in an ongoing project testing an intervention promoting the human papillomavirus (HPV) vaccine with parents of vaccine-eligible children attending a safety-net health system (20). Infection with certain types of HPV, the leading cause of cervical cancer, anogenital cancer, and genital warts, is preventable with vaccination. Routine vaccination for boys and girls aged 11 to 12 (and catch up vaccination up to age 21 for men and 26 for women) is recommended, but rates are suboptimal (21). Initially, in the baseline telephone survey we observed many DK responses, although this was not offered as a valid response option. Partway through data collection, we implemented a methodological strategy to prompt participants when they initially provided DK responses. Thus, participants were not randomized to receive, or not receive, the protocol with prompt. No study procedures or recruitment methods were changed in any other way. Though the study was not *a priori* designed to test different methods for handling DK responses, we conducted this secondary analysis of survey data to compare the effect of one methodological strategy (addition of the prompt) to the three most commonly used analytic simple imputation strategies for treating missing data--

*Analytic Treatment #1:* Coding DK responses as missing and excluding them from analyses.

*Analytic Treatment #2:* Recoding DK responses into the neutral midpoint of the Likert scale (i.e., neither agree nor disagree).

*Analytic Treatment #3:* Recoding DK responses with the computed item-level mean.

*Methodological Strategy #4:* Responding to participants' initial DK responses with a prompt that reassures them, reiterates the question, and encourages use of the response scale.

Given the widespread use, but limitations of these analytic approaches (3,22), we sought to establish whether the prompting strategy was a feasible method for addressing DK responses without influencing the content of participants' responses. Specifically, we examined the following research questions:

1. To what extent does use of the methodological strategy (i.e. giving participants who respond with DK a prompt) vs. not reduce the frequency of DK responses?

2. Does the prompt generate a similar factor structure and correlations among perceived susceptibility, perceived benefits, and behavioral intentions when compared to analytic approaches for managing missing data (i.e., test of factorial invariance or equivalence)?

We hypothesized that use of a prompt would result in less DK responding. We also hypothesized that the underlying factor structure of constructs would be similar (i.e. equivalent factor loadings), but the correlations between factors would be stronger (i.e. larger magnitude because of fewer missing data).

## Methods

Data were collected as part of a larger multi-stage study to develop a tablet-based self-persuasion intervention to promote HPV vaccination among parents undecided about the vaccine. The analyses reported here used survey data from the first stage of the larger project (20).

### Participants

Study participants (n=266) were parents or guardians of unvaccinated adolescents attending a safety-net health system serving a diverse population of low-income, under- and uninsured individuals living in Dallas County, Texas (20). Trained research assistants called parents and asked them to report their vaccine decision stage (e.g., never thought about the vaccine, undecided about it, do *not* want it, or do want it). Parents who had never thought about the vaccine or were undecided were invited and consented to participate. Research assistants administered the baseline survey over the telephone (20).

### Survey Measures and Procedure

Participants responded to several items assessing their HPV knowledge, psychosocial predictors of vaccination, and demographics. For the present investigation, analyses focused on three HBM constructs commonly assessed in the HPV vaccine literature: intentions to vaccinate, perceived benefits of vaccination, and perceived susceptibility to infection (16). Intentions were assessed with three items asking about the likelihood of parents vaccinating their children (e.g., “In the next year, how likely is it that you will get the HPV vaccine for your child?”). Perceived benefits were assessed with six items measuring parents’ degree of agreement regarding the positive aspects of HPV vaccination (e.g., “Getting my child the HPV vaccine will be good for my child’s health.”). Perceived risk was assessed with three items measuring the extent to which parents believed their children were vulnerable to HPV infection, cervical cancer, and genital warts (e.g. “How likely is it that your child will become infected with HPV at some point in his/her life?”). All items were adapted from prior measures (23–25). Survey responses were presented on a five-point Likert scale with responses ranging from *strongly agree/very likely* to *strongly disagree/very unlikely*.

We implemented the prompt (methodologic strategy) one year after study initiation when we noted the large number of DK responses given by parents.

**Initial protocol**—DK was not offered as a response option, but was recorded when expressed by the parent; 164 parents completed the baseline survey under this protocol. Among these 164 parents, we analytically treated DK responses in 3 different ways: coding as missing (Group 1), coding into the neutral point of the Likert-scale (Group 2), and coding into the item-level mean (Group 3).

**Prompting protocol**—When participants provided a DK response, research assistants were instructed to reassure the participant and given the opportunity to select a valid response option, using the following prompt: “It’s ok that you don’t know; there is no right or wrong answer. We just want to know your OPINION. Does your opinion match with any of the following choices?” Research assistants then repeated the Likert scale options from *strongly agree/very likely* to *strongly disagree/very unlikely*. Participants’ final responses were recorded as either a response on the scale or a persistent DK; 102 parents completed the survey under the prompting protocol, and constituted Group 4.

## Analyses

### Don’t Know Response Frequency

To test our hypothesis that the prompt would decrease DK responses, we compared the number of participants responding DK for each item in the initial protocol and prompting protocol (Group 1 vs. Group 4).

### Factorial Invariance

To determine whether the factor structures of the constructs were equivalent across the four approaches, we created four groups. Group 1 included the 164 survey respondents pre-protocol change in which DKs were coded as missing and excluded. Group 2 consisted of the same 164 survey respondents for which the DKs were recoded into the neutral point of the Likert scale (i.e. “neither”). Group 3 consisted of those 164 survey respondents but DKs were recoded into the item-level mean. Group 4 included the 102 participants who received the methodological prompt protocol. Any persistent DK responses in this group were coded as missing as is consistent with common practice. Through confirmatory factor analysis, we tested the factorial invariance (equivalence) of the three HBM constructs across the four groups (Figure 1). Unconstrained models allowed factor loadings to freely vary (Model 1), whereas the constrained model (Model 2) set factor loadings equivalent across the four groups (26). We calculated overall model fit indices and the global test of equality of covariance structures, which follows a chi-square distribution. A non-significant chi-square result indicates no significant decrement in model fit when factor loadings are constrained and suggests equivalence (i.e., participants’ interpretation of the constructs are similar across these approaches for dealing with DK responses).

### Correlations

To determine how the different approaches for managing DK responses affected the strength of associations between constructs, we examined the magnitude and direction of correlations for each of the four groups. We expected that a reduction in the number of DK responses (i.e., missing data) would result in stronger correlations between constructs. Consistent with

common practice, we created a mean score for each latent construct to compare zero-order correlations. We also computed Cronbach's alpha for each construct per group.

## Results

### Participant Characteristics

Participant demographic information is displayed in Table 1. Participants were compared across groups (initial protocol vs. prompting protocol) with no significant differences found. Additionally, participants who used and did not use DK were compared across demographics (reported in Supplemental Table 1). No significant differences in DK use were found across adolescent age, gender, parental education, or parental age ( $p$ s > .05).

### Don't Know Response Frequency

The number of participants responding DK one or more times decreased in the methodological strategy group when participants were reassured and prompted with the response scale a second time (55.7% of 164 participants supplied 1 or more DK responses when unprompted vs. 19.6% of 102 participants in the methodological strategy group). Table 1 shows the frequency of DK responses for each of the 10 survey items. Use of DK responses decreased for perceived susceptibility (3 items) from 31.1% to 10.8%, perceived benefits (4 items) from 39.5% to 13.7%, and behavioral intentions (3 items) from 14.4% to 5.9%. Of the 10 items we examined, the mean number of total DK responses per survey was 1.5 before the protocol change and .56 after the protocol change.

### Factorial Invariance

The  $\chi^2$  difference between Model 1 (unconstrained allowing all factors, variances, and covariances to vary freely between the three groups) and Model 2 (constrained factor loadings to be equal across groups) were not significant. This suggested measurement equivalence in factor loadings for all constructs (see Tables 2 and 3) and that all indicators loaded in a similar magnitude across approaches. We also ran models with the missing values in Group 4 recoded into the neutral point and into the item-level mean. The results did not differ (reported in Supplemental Tables 2 and 3); the models with these other two groups are not discussed further.

### Correlations

*Group 1:* Relative to the other approaches, the magnitudes of correlations were generally weakest when DK responses were coded as missing data and dropped from analyses (Table 4). Correlations when treating DK responses as missing data ranged from .235 (benefits and perceived susceptibility) to .429 (intentions and perceived susceptibility).

*Group 2:* The magnitudes of correlations were generally stronger for the group that recoded DK as the neutral than those seen when DK responses were treated as missing data (Table 4). Correlations when DK responses were recoded as the neutral midpoint ranged from .237 (benefits and perceived susceptibility) to .491 (intentions and benefits).

*Group 3:* The magnitudes of correlations were generally stronger for the group that recoded DK as the mean than those seen when DK responses were treated as missing data (Table 4). Correlations when DK responses were recoded as the item-level means ranged from .208 (benefits and perceived susceptibility) to .403 (intentions and benefits).

*Group 4:* Relative to the other approaches, the magnitudes of correlations were strongest when participants providing initial DK responses were given the prompt (Table 4). Correlations when participants received the prompt ranged from .329 (benefits and perceived susceptibility) to .589 (intentions and benefits). We tested the differences between correlations in Group 4 vs. Groups 1–3. Despite the magnitude of the correlations being consistently stronger, there were no significant differences when comparing the correlations ( $p > .05$ ).

Cronbach's alpha for each unidimensional construct per group are also reported in Table 4. Results were similar across groups and consistently demonstrated good internal consistency reliability ( $\alpha > .75$ ).

## Discussion

Our results suggest that a simple prompt reassuring participants that questions are intended to solicit their opinions, and giving participants a second opportunity to consider their response, may be a superior method for managing DK responses than either excluding them as missing data or imputing them into the means or the neutral midpoint. Notably, "I don't know" was not offered as a valid response option in our survey about HPV vaccination, yet prior to our method change, nearly 56% of participants responded "I don't know" at least once when asked about perceived susceptibility, benefits of vaccination, and intentions to vaccinate. This proportion was reduced to approximately 20% after instituting the prompt.

The structure of the latent variables was not impacted by any of the methods of addressing missing data. However, the majority of factor loadings were strengthened (7 out of 10 items) by both simple imputation strategies and with the implementation of the prompt. In addition, the direction of association and statistical significance of associations were unchanged. Notably, correlations were the strongest in the methodological strategy group when interviewers prompted participants and suggest that mitigation of DK responding both decreases noise in the measures and provides greater precision in detecting the associations. However, it is important to note differences in the magnitude of the correlations were not statistically significant and these findings should be replicated in larger samples before we have conclusive evidence that the change in magnitude is reliable. Considering that these data are cross-sectional and involve intentions rather than behavior, future work should also confirm whether use of a prompt increases the magnitude of predictive relations between attitudes and behavior.

Given that the larger study included only participants who had not yet decided whether or not to pursue HPV vaccination, DK responses may have reflected uncertainty around the related constructs. However, the substantial reduction (but not elimination) in the proportion of participants responding DK after the methodology change, combined with strengthened

reliability and factor loadings for most constructs, suggests that many participants held opinions on these survey items that were consistent with their other stated beliefs. Thus, in this case, initial DK responses may reflect hesitancy in responding due to lack of knowledge or expertise, whereas persistent DK responses may reflect meaningful uncertainty or indecision. With respect to perceived susceptibility, a DK response may also reflect avoidance of risk information, and may be more likely to be used in the context of HPV vaccination when asking parents about their child's risk of infection. Our survey items did not involve numeracy or quantification of risk as in related work (1,7), but rather asked participants to state the degree to which they agreed with an assertion or to the likelihood of a particular outcome; thus, it is important to note that issues of health literacy and health behavior performance may underlie the observed DK responses, particularly persistent use of DK. These DK responses may be meaningful and important in informing both interventions to promote health behavior performance and assessment of participants' health literacy or understanding of risk.

The main study was not designed *a priori* to test methods for addressing DK responses and the impact of missing data; thus, it is subject to several limitations. First, interviewers did not probe about the reason for DK responses, so we do not know the various factors triggering their responses. However, the larger number of participants selecting a valid response after receiving the prompt suggests the reduction in DK responses was due to uncertainty about which option to choose or fear of providing an 'incorrect' response. Although there was an overall reduction in DK, it is important to note that some DK responses persisted (e.g., 11% of those who received prompts continued to give a DK response to the perceived risk item). Future research on DK responses should examine the participants' degree of certainty and whether prompts encourage more thoughtful consideration of survey items. Within the health behavior literature, provoking increased deliberative thought could be beneficial and increase motivation for positive health behaviors, or could backfire. Second, our sample consisted of low-income, minority parents; thus, results may generalize best to these populations and our protocol change may not improve DK responses in other populations (1). Third, our analysis was limited to an interviewer-administered survey. Self-administered, web-based surveys could make use of similar prompts or require responses before advancing. Paper-based surveys are limited in their ability to offer reassurance regarding participants' opinions and the lack of a right or wrong answer, and thus may be subject to more missing data. Also, it may not be appropriate to clarify DK response on constructs for which DK might be a valid or reasonable response option (i.e. knowledge). Thus, it is important for researchers to consider the various factors influencing DK responding and be mindful of these when designing surveys. Finally, we did not test the analytic strategy of multiple imputation because it assumes data is missing at random (27). Although multiple imputation offers advantages over simple imputation strategies, multiple imputation models must be carefully specified based on characteristics of each study and dataset (28–30), and may introduce bias when missingness is extensive (29,30), as it was for some items in our study. Despite these limitations, introducing use of a prompt partway through data collection provided a unique opportunity to examine patterns in the prevalence of DK responses among parents undecided about the HPV vaccine, and how these patterns impacted measurement.



Future research should examine the meaning and function of DK responses in greater depth. As previously noted, DKs may stem from lack of motivation or ability to form a valid response (5), from uncertainty (9), or low knowledge of disease risk (1). Given that reasons for providing a DK response may vary within a sample, recording participants' thoughts behind their DK responses may enable the development of tailored prompts. In addition, qualitative research and cognitive testing of items before survey implementation may illuminate motivations for supplying a DK response in different populations, as well as items that elicit responses inconsistent with other beliefs. Cognitive testing may help improve the salience of survey items and identify discordant underlying beliefs.

## Conclusions

Responses of "I don't know" are common in health behavior research and threaten the validity of study findings. Our study suggests using a methodological prompt is a superior approach to capture more accurate or precise responses than analytic approaches for managing missing data (excluding them or recoding them as the scale midpoint). The simple prompt, reassuring participants and reminding them of response options, improved the strength of correlations between constructs and reduced the frequency of DK responses. Use of prompts may increase the validity of survey-based research among populations experiencing health disparities and facilitate more accurate data collection to elucidate behavioral determinants and design more effective interventions.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

**Funding:** All phases of this study were supported by a National Institutes of Health grant (PIs: Tiro, Baldwin; 1R01CA178414). Additional support provided by the UTSW Center for Translational Medicine, through the NIH/National Center for Advancing Translational Sciences (UL1TR001105) and the Harold C. Simmons Cancer Center (1P30 CA142543). Ms. Betts is supported by a pre-doctoral fellowship, University of Texas School of Public Health Cancer Education and Career Development Program (National Cancer Institute/NIH R25 CA57712).

## References

1. Hay JL, Orom H, Kiviniemi MT, Waters EA. "I Don't Know" My Cancer Risk: Exploring Deficits in Cancer Knowledge and Information-Seeking Skills to Explain an Often-Overlooked Participant Response. *Med Decis Making*. 2015 May; 35(4):436–45. [PubMed: 25810268]
2. Dillman DA, Eltinge JL, Groves RM, Little RJA. *Survey nonresponse*. New York: Wiley; 2002. *Survey nonresponse in design, data collection, and analysis*.
3. Little Roderick JA, Rubin Donald B. *Statistical analysis with missing data*. 2. Wiley-Interscience; 2002.
4. Graham JW. *Missing Data Analysis: Making It Work in the Real World*. *Annu Rev Psychol*. 2009 Jan; 60(1):549–76. [PubMed: 18652544]
5. Krosnick JA, Holbrook AL, Berent MK, Carson RT, Michael Hanemann W, Kopp RJ, et al. The impact of "no opinion" response options on data quality: Non-attitude reduction or an invitation to satisfice? *Public Opin Q*. 2002; 66(3):371–403.
6. Bruine de Bruin W, Fishbeck, Paul S, Stiber Neil A, Fischhoff Baruch. What Number is "Fifty-Fifty"? Redistributing Excessive 50% Responses in Elicited Probabilities. *Risk Anal*. 2002; 22(4): 713–23. [PubMed: 12224745]

7. Waters EA, Hay JL, Orom H, Kiviniemi MT, Drake BF. “Don’t Know” Responses to Risk Perception Measures: Implications for Underserved Populations. *Med Decis Making*. 2013 Feb; 33(2):271–81. [PubMed: 23468476]
8. Waters EA, Weinstein ND, Colditz GA, Emmons K. Explanations for side effect aversion in preventive medical treatment decisions. *Health Psychol*. 2009; 28(2):201–9. [PubMed: 19290712]
9. Waters EA, Kiviniemi MT, Orom H, Hay JL. “I don’t know” My Cancer Risk: Implications for Health Behavior Engagement. *Ann Behav Med*. 2016 Oct; 50(5):784–8. [PubMed: 26935308]
10. Krosnick JA. Survey research. *Annu Rev Psychol*. 1999; 50:537–67. [PubMed: 15012463]
11. Tourangeau Roger. Attitude measurement: a cognitive perspective. In: Hippler HJ, Schwarz N, Sudman S, editors *Social Information Processing and Survey Methodology*. New York, NY: Springer-Verlag; 1987. 149–62.
12. Warnecke RB, Johnson TP, Chávez N, Sudman S, O’Rourke DP, Lacey L, et al. Improving question wording in surveys of culturally diverse populations. *Ann Epidemiol*. 1997; 7(5):334–42. [PubMed: 9250628]
13. Nápoles-Springer AM, Santoyo-Olsson J, O’Brien H, Stewart AL. Using cognitive interviews to develop surveys in diverse populations. *Med Care*. 2006; 44(11):S21–30. [PubMed: 17060830]
14. Nadler JT, Weston R, Voyles EC. Stuck in the Middle: The Use and Interpretation of MidPoints in Items on Questionnaires. *J Gen Psychol*. 2015 Apr 3; 142(2):71–89. [PubMed: 25832738]
15. Ajzen I. The theory of planned behavior. *Organ Behav Hum Decis Process*. 1991; 50(2):179–211.
16. Rosenstock IM. Historical origins of the health belief model 2. 1974; 4:328–35.
17. Savage I. Demographic influence on risk perceptions. *Risk Anal*. 1993; 13(4):413–20. [PubMed: 8234949]
18. Foddy W. Probing: A dangerous practice in social surveys? *Qual Quant*. 1995; 29(1):73–86.
19. De Leeuw ED. Reducing missing data in surveys: An overview of methods. *Qual Quant*. 2001; 35(2):147–160.
20. Tiro JA, Lee SC, Marks EG, Persaud D, Skinner CS, Street RL, et al. Developing a Tablet-Based Self-Persuasion Intervention Promoting Adolescent HPV Vaccination: Protocol for a Three-Stage Mixed-Methods Study. *JMIR Res Protoc*. 2016; 5(1)
21. Walker TY, Elam-Evans LD, Singleton JA, Yankey DR, Markowitz LE, Fredua B, et al. National, Regional, State, and Selected Local Area Vaccination Coverage Among Adolescents Aged 13–17 Years — United States, 2016. *Morbidity and Mortality Weekly Report*. 2017:874–82. Report No.: 66. [PubMed: 28837546]
22. Schafer JL, Olsen MK. Multiple Imputation for Multivariate Missing-Data Problems: A Data Analyst’s Perspective. *Multivar Behav Res*. 1998 Oct; 33(4):545–71.
23. Bynum SA, Wigfall LT, Brandt HM, Richter DL, Glover SH, Hébert JR. Assessing the influence of health literacy on HIV-positive women’s cervical cancer prevention knowledge and behaviors. *J Cancer Educ*. 2013; 28(2):352–356. [PubMed: 23564430]
24. Gerend MA, Barley J. Human papillomavirus vaccine acceptability among young adult men. *Sex Transm Dis*. 2009; 36(1):58–62. [PubMed: 18830138]
25. Gerend MA, Shepherd JE. Predicting human papillomavirus vaccine uptake in young adult women: Comparing the health belief model and theory of planned behavior. *Ann Behav Med*. 2012; 44(2): 171–180. [PubMed: 22547155]
26. Byrne BM. Testing for multigroup invariance using AMOS graphics: A road less traveled. *Struct Equ Model*. 2004; 11(2):272–300.
27. Schafer JL. Multiple imputation: a primer. *Stat Methods Med Res*. 1999; 8(1):3–15. [PubMed: 10347857]
28. Hayati Rezvan P, Lee KJ, Simpson JA. The rise of multiple imputation: a review of the reporting and implementation of the method in medical research. *BMC Med Res Methodol* [Internet]. 2015 Dec.15(1) [cited 2018 Mar 6]. Available from: <http://bmcmredresmethodol.biomedcentral.com/articles/10.1186/s12874-015-0022-1>.
29. Lee KJ, Carlin JB. Recovery of information from multiple imputation: a simulation study. *Emerg Themes Epidemiol*. 2012; 9(1):3. [PubMed: 22695083]

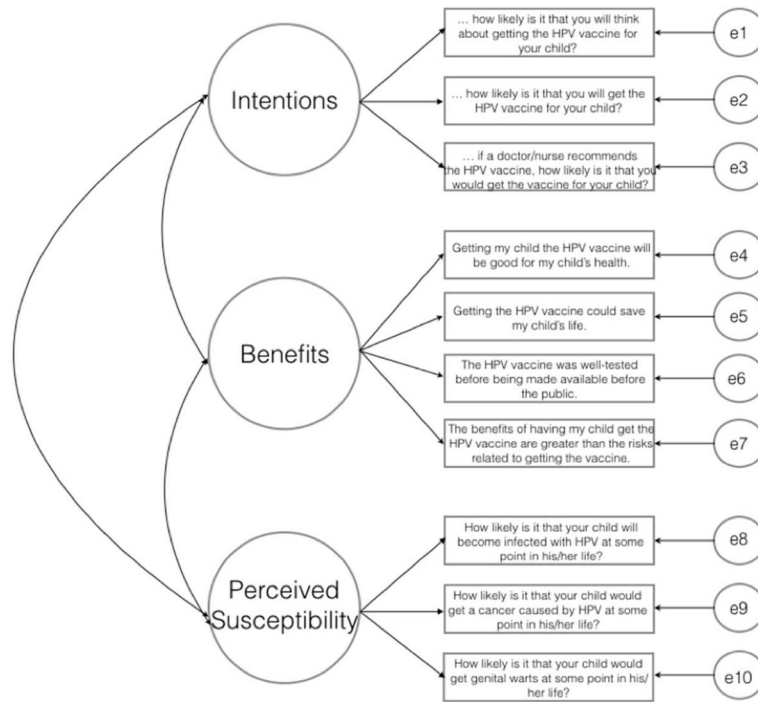
30. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ*. 338:b2393.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript



**Figure 1.**  
Three-factor model of latent constructs

Table 1

## Participant Demographics

Variable	Initial Protocol			Revised Protocol			Difference testing*
	Mean	Range	Range	Mean	Range	Range	
Parent Age	39.4	27-72	41.15	22-75			$p=.09$
Child Age	13.6	11-17	13.4	11-17			$p=.23$
Parent Sex	N	%	N	%			$p=.12$
Female	163	97.6	92	92.9			
Male	4	2.4	7	7.1			
Parent Language							$p=.91$
English	67	40.1	39	39.4			
Spanish	100	59.9	60	60.6			
Parent Race/Ethnicity							$p=.89$
Hispanic	116	69.5	70	70.7			
Non-Hispanic White	3	1.8	1	1.0			
Non-Hispanic Black	47	28.1	28	28.3			
Unknown race/ethnicity	1	.6	0	0			
Parent Education							$p=.18$
Less than high school	44	26.3	19	19.2			
Some high school	35	21.0	25	25.3			
High school/Vocational degree	56	33.5	27	27.3			
Some college/College graduate	32	19.2	28	28.2			
Adolescent Sex							$p=.61$
Male	85	50.9	54	54.5			
Female	82	49.1	45	45.5			

\* Participants were compared across groups (initial protocol vs. prompting protocol) using independent samples t-tests and Fisher's exact chi-square analyses

Table 2  
 Frequency of Don't Know responses and factor loadings for each group (analytic and methodologic approaches)

Construct/Item	Number DK Responses				Factor Loadings			
	Initial Protocol N=164	Revised Protocol N=102	Group #1 (Analytic code as missing) N=164	Group #2 (Analytic code to neutral) N=164	Group #3 (Analytic code to mean) N=164	Group #4 (Promoting Intervention) N=102		
<b>Intentions</b>								
...how likely is it that you will think about getting the HPV vaccine for your child?	11	3	.84	.83	.82	.83		
... how likely is it that you will get the HPV vaccine for your child?	16	2	.87	.87	.89	.85		
...if a doctor or nurse at the clinic recommends the HPV vaccine, how likely is it that you would get the vaccine for your child?	10	2	.82	.84	.78	.88		
<b>Perceived benefits</b>								
Getting my child the HPV vaccine will be good for my child's health.	17	17	.79	.79	.75	.86		
Getting the HPV vaccine could save my child's life.	22	3	.81	.71	.76	.63		
The HPV vaccine was well tested before being made available to the public.	50	13	.79	.71	.73	.61		
The benefits of having my child get the HPV vaccine are greater than the risks related to getting the vaccine.	27	6	.61	.58	.58	.72		
<b>Perceived Susceptibility</b>								
How likely is it that your child would become infected with HPV at some point in their life?	3	8	.82	.80	.80	.80		
How likely is it that your child would get a cancer caused by HPV at some point in their life?	34	5	.81	.77	.76	.83		
How likely is it that your child would get genital warts at some point in their life?	35	10	.78	.74	.77	.85		

**Table 3**

Testing the invariance of the three-factor model across the four groups (Analytic #1, Analytic #2, Analytic #3 and Methodological Strategy #4)

<b>Model</b>	$\chi^2$	<i>df</i>	<b>CFI</b>	<b>RMSEA (90% CI)</b>
1. Unconstrained	172.45	128	.982	.024 (.014, .033)
2. Equality of Factor Loadings	197.48	167	.987	.017 (.000, .026)
<i>Model Comparisons</i>				
	$\chi^2$ difference	Difference <i>df</i>	<i>p</i> -value	
Model 1 vs. Model 2	25.03	39	.959	

**Table 4**

Correlations among Health Belief Model constructs for each of four groups

<b>Group #1 (Analytic code as missing)</b>	<b>Intentions</b>	<b>Benefits</b>	<b>Perceived Susceptibility</b>
Intentions	1		
$\alpha = .89$			
Benefits	.404**	1	
$\alpha = .86$	(N=159)		
Perceived Susceptibility	.429**	.235**	1
$\alpha = .83$	(N=142)	(N=141)	
<b>Group #2 (Analytic code to neutral)</b>	<b>Intentions</b>	<b>Benefits</b>	<b>Perceived Susceptibility</b>
Intentions	1		
$\alpha = .88$			
Benefits	.491**	1	
$\alpha = .79$	(N=164)		
Perceived Susceptibility	.412**	.237**	1
$\alpha = .84$	(N=164)	(N=164)	
<b>Group #3 (Analytic code to mean)</b>	<b>Intentions</b>	<b>Benefits</b>	<b>Perceived Susceptibility</b>
Intentions	1		
$\alpha = .87$			
Benefits	.403**	1	
$\alpha = .80$	(N=164)		
Perceived Susceptibility	.383**	.208**	1
$\alpha = .82$	(N=164)	(N=164)	
<b>Group #4 (Methodological prompt)</b>	<b>Intentions</b>	<b>Benefits</b>	<b>Perceived Susceptibility</b>
Intentions	1		
$\alpha = .89$			
Benefits	.589**	1	
$\alpha = .80$	(N=95)		
Perceived Susceptibility	.446**	.329**	1
$\alpha = .88$	(N=94)	(N=91)	

\*\*  
p < .01 level