8-1-2006

"None of the Above" as an Answer Option in Observational Based Multiple-Choice Questions

Stephen King
Western Kentucky University

Follow this and additional works at: http://digitalcommons.wku.edu/theses

Part of the Education Commons, and the Psychology Commons

Recommended Citation
http://digitalcommons.wku.edu/theses/288

This Thesis is brought to you for free and open access by TopSCHOLAR®. It has been accepted for inclusion in Masters Theses & Specialist Projects by an authorized administrator of TopSCHOLAR®. For more information, please contact topscholar@wku.edu.
NONE OF THE ABOVE AS AN ANSWER OPTION IN OBSERVATION BASED MULTIPLE-CHOICE QUESTIONS

A Thesis
Presented to
The Faculty of the Department of Psychology
Western Kentucky University
Bowling Green, Kentucky

In Partial Fulfillment
Of the Requirements for the Degree
Master of Arts

By
Stephen M. King

August 2006
"NONE OF THE ABOVE" AS AN ANSWER OPTION IN OBSERVATION BASED MULTIPLE CHOICE QUESTIONS

Date Recommended       June 1, 2006

__________________________
Director of Thesis

__________________________
Elizabeth L. Shoemaker

__________________________
Papa Jarree

__________________________
Dean, Graduate Studies and Research       Date

__________________________
Emmett Grav

6/21/06
Table of Contents

Abstract............................................................................................................iv
Introduction......................................................................................................1
Method.............................................................................................................6
Results............................................................................................................8
Discussion.....................................................................................................10
References...................................................................................................13
Appendix A – Photographs............................................................................16
Appendix B - Meeting Scene Questions.......................................................19
Appendix C – Party Scene Questions.............................................................22
Appendix D – Instructions............................................................................25
This study examined the characteristics of items using *none of the above* (NOTA) as an answer option in observation based multiple-choice questions. Previous research has examined only the use of a NOTA option in academic knowledge based testing, not in visual recognition testing. Item difficulty and discrimination were examined for three different item formats: (a) items without a NOTA option, (b) items with NOTA as a distracter, and (c) items with NOTA as the correct answer. The questions were based on two photographs with similar content. A total of 98 participants from a large southeastern university completed a visual recognition test containing all three item types. Results revealed no difference in item discrimination between items without a NOTA option and items with a NOTA option, but did indicate that items with a NOTA option were more difficulty. A discussion of the results, limitations, and suggestions for future research is provided.
Introduction

In 1926, B.R. Buckingham suggested that teachers use a new testing technology to accurately and efficiently gather desired information from their students (Cizek & O’Day, 1994). This new technology would later come to be known as *multiple-choice* testing and would become one of the most popular forms of measurement in the United States (Smith, 1982). Yet after nearly 80 years, the optimal design of the item format has yet to be determined.

One issue that arises when using multiple-choice items is whether to use *none of the above* (NOTA) as a distracter item. Many textbooks that discuss NOTA suggest that it not be used as an item distracter because of increased item difficulty and reduced discrimination (Frays, 1991). In support of these suggestions, there has been some empirical research that demonstrates NOTA items are more difficult than items without NOTA as an answer option (Tollefson, 1987).

Tollefson (1987) compared the difficulty and discrimination (corrected item-total correlation) of items using NOTA as the correct answer, items with NOTA as a distracter, and items without NOTA as a distracter or correct answer. The results suggested that items using NOTA as the correct answer were more difficult than those without a NOTA option. Also, items using NOTA as a distracter tended to be easier than those with NOTA as the correct answer, but only slightly more difficult than items with one correct answer. There were no differences across the three forms in terms of item discrimination. In a similar study, Oosterhof and Coats (1984) found the same results when NOTA was used as the correct answer.
Despite the fact that a NOTA option makes items more difficult, there is evidence that suggests NOTA is a viable answer option. Frary’s (1991) study of the NOTA option spanned across eight academic disciplines at the college level; virtually all other studies comparing NOTA and non-NOTA options used tests covering only a single topic area. Frary (1991) found that, on average, NOTA items were minimally more difficult than non-NOTA items, and discrimination indexes for the two types of items were equal to the second decimal place. This research suggests that use of a NOTA distracter is unlikely to affect the psychometric properties of the test.

Further research on NOTA includes a meta-analysis conducted by Knowles and Welch (1992). The meta-analysis included 20 studies of item difficulty and 11 studies of item discrimination, with some studies dating to the 1940s. Results showed that using a NOTA option resulted in nonsignificant effect sizes for both item discrimination and item difficulty, indicating that there is no difference in item difficulty or item discrimination when NOTA items were used versus items without NOTA.

Since the Knowles and Welch meta-analysis, there have only been two published empirical studies about the use of NOTA. Crehan, Haladyna, and Brewer (1993) found that items using a NOTA answer option are more difficult, but are not more or less discriminating than items without a NOTA option. Gross (1994) found that items with a NOTA option had lower item discrimination coefficients. Gross (1994) suggested that NOTA should not be used because it reduces an item’s ability to distinguish between test takers with full knowledge and test takers with incomplete knowledge.

All of the research on the use of NOTA to date has been on knowledge based achievement tests, usually in an academic setting. There are, however, applications of
multiple-choice tests to areas outside the realm of knowledge based measurement. One such area is recognition testing.

A recognition test is a test that requires the test taker to recognize information by matching information presented with information previously gathered. One distinction between an achievement test and a recognition test is that during the recognition test, pieces of information might interfere with each other (Anderson, 1995). This interference may be both retroactive and proactive. Retroactive interference occurs when information currently being learned interferes with previously stored information. In proactive interference, previously stored information interferes with new information (Baron, 1999). During a recognition test, both retroactive and proactive interference interfere with test taker performance. Chandler (1994) found that subjects who studied related pictures had a reduced ability to discriminate between the two pictures, even though their confidence increased. His results suggested that when similar information is retained, the information may become distorted during recall.

The idea that similar information may be distorted during recall is an important issue within the world of work. Consider the job of a corrections officer. While filling out incident reports, it is important for a corrections officer to recall exactly what occurred. Some of the information from different incidents, however, may be very similar. The information retained for a given incident, therefore, may become distorted during recall, resulting in an inaccurate incident report.

Distorted recall has been given much attention, especially in the area of eyewitness memory. Researchers in this area have studied what is known as the misinformation effect (e.g., Bekerian & Powers, 1983; Belli, 1989; Loftus, Miller, &
Burns, 1978; McCloskey & Zarragoza, 1985). The misinformation effect occurs when a person observes an event, is exposed to misleading information related to the event, then during recall reports the misinformation as part of the observed event (Pickel, 2004). One explanation for the misinformation effect is source monitoring error (Lindsay, 1990; Lindsay & Johnson, 1989; Zaragoza & Lane, 1994, 1998). With source monitoring error, a memory formed from one source is incorrectly recalled as coming from another source or is even attributed to both sources (Pickel, 2004). Pickel also noted that errors are more likely to occur when source cues are not easily remembered, if the cues related to the two different memory sources are similar, or if the source attribution is made in a hurried fashion. Source monitoring error relates to Chandler’s (1994) finding that studying related pictures distorts recall. It is possible that Pickel’s three sources of increased error may be affecting the subjects in Chandler’s study.

The Present Study

Using a NOTA option on personnel selection test items that assess the ability to attribute information to the correct source may help discriminate between those who can correctly recall information and those who cannot correctly recall information. Thus, using NOTA on selection tests may help bolster the selection of the most qualified applicants. It is the purpose of this study to determine the role of a NOTA answer option in multiple-choice questions used in a visual recognition test. A visual recognition test is a test that requires the test taker to accurately identify information presented in a visual format.

The role of NOTA will be assessed in the present study by comparing item difficulty and item discrimination of three forms of multiple-choice items: (a) items
without a NOTA option, (b) items with NOTA as a distracter, and (c) items with NOTA as the correct answer. It is hypothesized that item difficulty and item discrimination will be greater for items with a NOTA option than for items without a NOTA option.

Hypothesis 1: Items with a NOTA option than will be more difficult than items without a NOTA option.

Hypothesis 2: Item discrimination will be greater for items with a NOTA option than for items without a NOTA option.
Method

Subjects

A total of 98 students at a public southeastern university completed the test for extra credit in various courses. The sample was comprised of both undergraduate and graduate students.

Instrument

Ten questions were developed from two different pictures (see Appendix A for black and white versions), for a total of twenty questions (Appendix B and Appendix C). Each question had three variations: (a) items without NOTA, (b) NOTA as a distracter option, and (c) NOTA as the correct answer. Each test version had 20 items in total, with six items having NOTA as a distracter, six items with NOTA as the correct answer, and eight items without a NOTA option. No two versions of the test had the same items within each variation. A total of 19 versions of the test were formed. A twentieth version of the test was formed and used but was later found to be incorrectly designed (not enough non-NOTA items). All data from that form were deleted.

For those items for which the NOTA option is a distracter, one distracter was randomly be selected to be replaced by NOTA. This replacement method will be employed instead of replacing the least functioning item technique used in previous research (e.g., Crehan, et al., 1993). Replacing the least functioning distracter (i.e., distracters with the lowest discrimination indices) could have the unintentional effect of raising the item’s discrimination index, dependent on how rarely the distracter in question is endorsed.
Procedure

In groups of up to twenty people, participants were given standardized instructions (located in Appendix D) regarding what they should examine in the photographs. Participants then individually looked at the two pictures during a one-minute study period. After the one-minute study period, the pictures were returned to an envelope. Participants then answered twenty questions (ten from each picture) about the two pictures. Participants completed the questions in less than 5 minutes.
Results

The first hypothesis stated that items with a NOTA option would be more difficult than items without a NOTA option. Item difficulties (percent of test takers with a correct answer for given item) were calculated for each of the three item types. Table 1 lists item difficulty means and standard deviations for all three item types. Paired samples t-tests were executed in order to test the hypothesis. Overall, items using NOTA as a distracter or the correct answer were significantly more difficult than items without a NOTA option, \( t(97) = 2.60, p < .05 \). The magnitude of this difference was moderate \( (d = .39) \). A comparison of items without a NOTA option to each of the possible uses of a NOTA option yielded conflicting results. Items having no NOTA option were not significantly easier than items with NOTA as a distracter, \( t(97) = 1.21, p > .05 \). Items without the NOTA option, however, were significantly easier than items with NOTA as the correct answer, \( t(97) = 3.19, p < .05 \). There was a modest effect size for this difference \( (d = .44) \).

Table 1

*Item Difficulty Means and Standard deviations*

<table>
<thead>
<tr>
<th>Item Type</th>
<th>( k )</th>
<th>( M )</th>
<th>( SD )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Items without NOTA</td>
<td>8</td>
<td>.52</td>
<td>.19</td>
</tr>
<tr>
<td>Items with NOTA as distracter or correct answer</td>
<td>12</td>
<td>.45</td>
<td>.15</td>
</tr>
<tr>
<td>Items with NOTA as a distracter only</td>
<td>6</td>
<td>.48</td>
<td>.22</td>
</tr>
<tr>
<td>Items with NOTA as the correct answer only</td>
<td>6</td>
<td>.43</td>
<td>.24</td>
</tr>
</tbody>
</table>

*Note.* \( k \) = number of items.
The second hypothesis stated that item discrimination would be greater for NOTA items than for items with a NOTA option. To analyze item discrimination, corrected item-total correlations were calculated and compared with an independent samples $t$-tests. Table 2 lists mean corrected item-total correlations and standard deviations for each item type. Overall, items using a NOTA option were not significantly more discriminatory than items without a NOTA option, $t(18) = 1.38, p > .05$. A significant difference was found between items with NOTA as the correct answer and items without a NOTA option, $t(12) = 1.87, p < .05$. The magnitude of this difference was large ($d = 1.01$). Items with NOTA as a distracter were not more discriminatory than items without a NOTA option, $t(12) = .43, p > .05$. Finally, given that many analyses of correlations employ Fisher’s $r$ to $z$ transformation of correlation coefficients (Hays, 1988), all 20 corrected item-total correlations were transformed to Fisher’s $z$ coefficients, and the analysis was repeated. The results were nearly identical in terms of $t$ statistics and identical in terms of conclusions. As such, the results will not be reprinted.

Table 2

*Item Discrimination Means and Standard deviations*

<table>
<thead>
<tr>
<th>Item Type</th>
<th>$k$</th>
<th>$M$</th>
<th>$SD$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Items without NOTA</td>
<td>8</td>
<td>.07</td>
<td>.09</td>
</tr>
<tr>
<td>Items with NOTA as distracter or correct answer</td>
<td>12</td>
<td>.12</td>
<td>.09</td>
</tr>
<tr>
<td>Items with NOTA as a distracter only</td>
<td>6</td>
<td>.08</td>
<td>.07</td>
</tr>
<tr>
<td>Items with NOTA as the correct answer only</td>
<td>6</td>
<td>.16</td>
<td>.10</td>
</tr>
</tbody>
</table>

*Note.* $k =$ number of items.
Discussion

Hypothesis 1 stated that items with a NOTA option would be more difficult than items not using NOTA as an answer option. The results of the item difficulty analysis supported this hypothesis. The percent of people answering the typical item with a NOTA option correctly was lower than the percent of people answering the typical non-NOTA item correctly. This finding is inconsistent with Knowles and Welch’s (1992) meta-analytic findings on item difficulty. The Knowles and Welch meta-analysis, however, was based on the use of NOTA in academic testing, not on visual recognition testing as was used in this study. Therefore, the finding that items with a NOTA option are more difficult could be beneficial to test developers looking to increase the difficulty of multiple-choice visual recognition tests. Some researchers (Crehan & Haladyna, 1991; Tollefson, 1987) suggest that increased item difficulty is reason not to use a NOTA option; however, item difficulty levels for NOTA items in this study were still close to the .50 level of item difficulty. In the absence of other information about the test taking population, the .50 level of item difficulty is the optimal level when the purpose of the test is to discriminate between different levels of achievement.

The results indicated that items using NOTA did not have greater corrected item-total correlations (i.e., item discrimination) than items without NOTA. This finding is contrary to what was hypothesized with our second hypothesis. These results could be, in part, due to issues with power. In an analysis of item discrimination the $N$ for significance tests is determined by the number of items, not the number of people. Naturally, having more people will lead to more stable results, but the traditional method of testing for significant differences among item discrimination coefficients (e.g., Tollefson, 1987) is
executed at the item level. Given that the total number of items was 20, power was limited. Analysis of discrimination should be repeated with a longer test in order to have greater statistical power.

A second problem with the analysis of Hypothesis 2 is that the traditional method for testing for differences among item discrimination parameters may violate a key assumption of independent samples t-tests. Specifically, independent samples t-tests assume that each case is independent of the other cases (Hays, 1988). Independence means that the data in one case is different from the data in the other cases. Typically, each case represents a person and each person counts as only one case. In the analyses done above, each case is not a person (as is usual), but rather an item (specifically, an item-total correlation coefficient). Each item-total correlation coefficient is computed on the same 98 people. Although each item is different from the others, they all share the same people. Thus, it is arguable that the cases in analysis of item discrimination are not independent and the results are not correct.

Further research might include examining the reliability of NOTA items compared to non-NOTA items. Because corrected item-total correlations are one but way to estimate reliability, it is desirable to estimate reliability with methods apart from the internal consistency family (e.g., test-retest). Furthermore, future research could look at the use of video instead of using still pictures. Using moving video with multiple events happening at the same time would create a more realistic situation than photographs.

There were several limitations in the present study. First, college students were used as test takers rather than job applicants. Clearly, there is a difference in motivation between the two types of test takers, potentially affecting results. Second, the pictures
were not work related. The pictures did not portray work-related events, but instead day-to-day situations and interactions. Different picture subject matter may alter responses.

The third limitation is the time limit used to study the pictures. The determination to use a one-minute study period was not based on research, but instead was based on the authors’ applied experience with visual recognition testing in personnel selection.

In summary, items with a NOTA option used in observation based multiple-choice questions were more difficult than non-NOTA items. Further research is needed, however, to make a determination about the usefulness of NOTA items in personnel selection decisions.
References


Appendix A

Photographs
Appendix B

Meeting Scene Questions
1. How many adults are visible in the meeting scene picture?
   A) 4
   B) 5
   C) 6
   D) 7 *

2. In the meeting scene, what color is the man’s coat in the middle of the picture?
   A) White
   B) Blue
   C) Black
   D) Green *

3. How many people in the meeting scene are wearing glasses?
   A) 1
   B) 2
   C) 3 *
   D) 4

4. In the meeting scene, what is the child seated at the table doing with his hand?
   A) Holding a can.
   B) Touching his mouth. *
   C) Holding his mom’s arm.
   D) Reaching for something.

5. In the meeting scene, what is the man in the middle of the picture doing with his right hand?
   A) Holding a cup.
   B) Resting it on his hip.
   C) Placing it in his pocket. *
   D) Letting it hang by his side.

6. What is on the wall behind the woman standing on the far right in the meeting scene?
   A) A light switch. *
   B) A picture.
   C) A window.
   D) An exit sign.
7. In the meeting scene, what is the center of attention of the people standing in the middle of the picture?
   A) The camera.
   B) The woman on the left side of the group.
   C) The man in the middle of the group.
   D) The woman on the right side of the group. *

8. What does the woman standing on the far right side of the meeting scene have on her upper sleeve?
   A) A patch
   B) A pocket *
   C) Embroidery
   D) Fringe

9. In the meeting scene, what color is the hat on the man seated closest to the camera?
   A) Green
   B) Blue
   C) Tan *
   D) Black

10. How many people in the meeting scene are holding cups?
   A) 3
   B) 4 *
   C) 5
   D) 6

(* indicates Correct Answer)
Appendix C

Party Scene Questions
1. How many people are visible in the party scene picture?
   A) 3
   B) 4
   C) 5 *
   D) 6

2. Who is the center of attention of the people in the party scene?
   A) The woman on the far right.
   B) The woman in the middle.
   C) The man in the middle.
   D) The women on the far left. *

3. In the party scene, what is the older woman on the far right doing with her left hand?
   A) Resting it on her hip.
   B) Holding an object. *
   C) Pointing at something.
   D) Holding her other arm.

4. What color is the man’s tie in the party scene?
   A) White *
   B) Blue
   C) Black
   D) Green

5. In the party scene, how many rings does the woman in the middle have on her right hand?
   A) 1 *
   B) 2
   C) 3
   D) 4

6. What is the man in the party scene holding in his right hand?
   A) A scarf.
   B) Sunglasses.
   C) A book.
   D) A cup. *

7. In the party scene, what color are the far right woman’s earrings?
   A) White
   B) Blue
   C) Green *
   D) Black
8. What is the woman in the middle of the group in the party scene doing with her left hand?
   A) Resting it on her hip
   B) Holding a cup.
   C) Pointing at something.
   D) Holding her other arm. *

9. What is the color of the table cloth on the table behind the people in the party scene?
   A) White *
   B) Blue
   C) Green
   D) Black

10. In the party scene, what does the woman on the far right have pinned to her dress?
    A) A ribbon.
    B) Flowers. *
    C) A pin.
    D) A nametag.

(* indicates Correct Answer)
Appendix D

Instructions
During this one minute study period you will view 2 photos. You must try to memorize as much as you can about the people, objects and events shown in each of the photographs. You will not be allowed to take notes during this study period. Look closely at the people in the scene: who they are, what they are wearing, what they are doing, etc. Note the setting of the scene: what objects are present, how objects are positioned, what people are doing with objects, etc. After the study period, the photographs will be placed back in the envelope and you will then have to answer questions about them.

Now open envelope A and study the two pictures for one minute.

[After 30 seconds]

You have 30 seconds remaining.

[After 30 more seconds, 60 seconds total]

Now put the pictures back into the envelope, and remove the question sheet.

Please answer these questions now.