Technical Report # 29

Use of Focus Groups to Inform the Construction of a Universally Designed Mathematics Test

Leanne Ketterlin-Geller

Julie Alonzo

Gerald Tindal

University of Oregon

Published by

Behavioral Research and Teaching
University of Oregon • 175 Education
5262 University of Oregon • Eugene, OR 97403-5262
Phone: 541-346-3535 • Fax: 541-346-5689
http://brt.uoregon.edu

Abstract

Focus groups were used to examine the needs of students taking a computer adaptive math assessment. The assessment designed using universal design and item response theory. Focus groups included teachers, administrators, child advocacy groups, and 3rd grade students and parents from the following groups: students with disabilities, students in general education, and students learning English as a second language. Focus groups identified using the mouse, understanding directions and accessing the format of the items as sources of confusion. Additionally, a survey suggested that English Language Learners have differential access to computers. These findings will be used to revise the assessment.

Introduction

Focus groups have been used for a variety of purposes and are a well-established method for gathering information from different groups of people (Morgan & Krueger, 1998). In marketing, focus groups serve as an efficient method for gathering a great deal of information about a particular product in a short amount of time. In social science research, they have informed such diverse fields as psychology, sociology, and education (Patton, 2002). Focus groups are particularly preferred when there is a large gap between researchers and the people they are studying, when researchers want to understand issues related to diversity, and when a friendly, respectful research method is desired (Vaughn, Shumm, & Sinagub, 1996). As a result, they are well suited to research efforts involving the development or refinement of products founded on the principles of Universal Design (with an emphasis on applicability for all). The purpose of this report is to describe a research effort designed to gain information about a universally designed mathematics test from a broad range of educational stakeholders using focus groups.

Universal Design

Originating in architecture, universal design is a practical solution for designing and creating equally accessible products and environments that can be used by people in a particular community (Center for Universal Design, 1997). In this context, *community* refers expansively to all members of the population that might conceivably interact with the environment. Central to the principles of universal design is the premise that access should not only be available, it should be integrated smoothly into the actual design such that all individuals benefit.

Additionally, by considering the needs of all members of the community during the conceptualization phase, environments are naturally created to be accessible to all, and therefore, the features that make a particular environment accessible are seamlessly integrated. In architecture, for instance, universal design might result in a gently sloping ramped entryway leading to wide front doors that are opened by remote sensors. In such a design, people entering the building—whether using wheelchairs, walkers, crutches, or moving without assistance—use the same entryway, without the need for some members of the community to seek an alternate route for entering the building.

Application of the principles of universal design is not limited to architecture; the principles extend to any situation where an individual must interact with aspects of a given environment. In testing, for example, individuals with different access skills or special needs encounter the testing environment differently. Tests built without an awareness of the ways in which access skills interfere with accurate measurement of the construct being assessed are biased against test takers with deficiencies in these areas in much the same way as buildings without ramps or elevators are biased against people in wheelchairs. In contrast, Universal Design for Assessment (UDA) promotes the design and use of assessment environments that allow all members of a given population (the testing community) to access the material without their various special needs affecting either their access to or their performance in the testing situation.

Students enter the testing environment with a multitude of skills and characteristics. Some of the skills are targeted on the assessment and some are not. The characteristics that are ancillary to the tested construct often impede the demonstration of the targeted skills and knowledge. The extent to which these personal changes impact the expression of targeted skills and knowledge is dependent upon features of the environment. Specific features of the environment can negatively or positively interact with specific personal characteristics. For example, a person with a physical disability that is compensated for by use of a crutch may have difficulty interacting with an environment in which the doors are narrow. Changing the environment to include a wider door may improve the interaction between the user and the environment. Universal design addresses this need by designing environments that account for the interaction between personal characteristics and environmental features.

A promising line of research in the area of UDA involves the use of computer-adapted tests (CATs) that use Item Response Theory (IRT) to deliver each test taker a set of test items custom designed to pinpoint with a high degree of accuracy the individual's skill in the content being assessed (Ketterlin-Geller, 2003). Aspects of UDA incorporated into the CAT being developed in our research include the use of basic skills tests to determine the linguistic complexity with which test items should be presented to individuals taking the test (read aloud, simplified text, or standard administration) and the language format (English only, bilingual Spanish and English) as well as more subtle design features such as the use of large font size to accommodate students with minor vision impairment, presentation of only one item at a time to accommodate students with attention deficits, and simple one-click mouse response requirements to reduce the impact of varying levels of computer proficiency.

Methods

This project is part of a larger, multi-year project designed to investigate UDA and its application in the classroom and large-scale tests. During the first year, a universally designed mathematics test for 3rd-grade students was developed. The test items were written, reviewed by content experts for appropriateness, and pilot tested with over 450 students. IRT analyses following the pilot testing revealed several items that were not functioning appropriately or were redundant with other items, and these items were removed or re-written. During the current year, additional items are being added to the item bank at both the low and high end of the range of item difficulty to reduce the standard error of measurement that results when too few items exist to allow sufficient sampling at a particular ability level. The first year also saw the development of the computer delivery system, incorporating design elements following the principles of universal design that were gleaned from the literature. In the current year, we are refining the UDA, and focus groups are one of the primary methods by which we are gathering information about the design.

This paper discusses the use of focus groups in year two of the study. Following Morgan and Krueger's (1998) suggestions, we began by defining the purpose of our study and the particular groups we were interested in including. We divided our target population into two groups: students and adults. Within the student group, we hoped to include students in general education; students with disabilities or access issues; students with physical disabilities (gross motor or mild vision impairments); and Spanish-speaking English language learners. Within the adult group, we hoped to include parents or guardians of general education, special education, and Spanish-speaking English language learner students; teachers who work with general and special education students and those served by Title I; administrators; technical specialists from schools; and representatives from child advocacy groups.

Having identified our target population, we then recruited participants and trained our facilitators, observers, and note-takers. We held our focus groups and analyzed the data that resulted. Finally, we used the information gathered in the focus group process to revise our test design in preparation for the second year of pilot testing.

Setting and Subjects

Our study took place in December of 2003 in a mid-sized city in the Pacific Northwest. We conducted a total of seven focus groups on four separate nights, coordinating the parent and student groups so they took place on the same evening to reduce the disruption to the families' lives. In all, seven general education students, two students with identified disabilities, and five Spanish-speaking English language learners participated. Eight parents of general education students, three parents of students with identified disabilities, five parents or relatives of Spanish-speaking English language learners, seven teachers and administrators, and five members of child advocacy groups participated.

Thirteen people with varied levels of experience and education participated in the research efforts (See Figure 1). The research team worked with three translators to ensure that all materials were appropriately written in Spanish and that transcripts were translated accurately. A paid school district liaison was hired to recruit participants for each of the focus groups, following structured guidelines developed by the lead researcher (See Appendix A).

Figure 1
Members of the Research Team

Education	Role(s)	Special Skills
Ph.D. in education	Note-taker, Observer	Dissertation on building a validity argument for the UDA CAT mathematics test developed in year one of this study; assessment background
Ph.D. in education	Note-taker, Observer	Technology and assessment background
Doctoral student	Moderator, Note-taker	Experience facilitating diverse groups, background in qualitative research methodologies
Doctoral student	Note-taker, Observer	Assessment background

Doctoral student	Moderator	Experience facilitating diverse groups and working with
Doctoral student	Moderator	Experience with young children and with facilitating diverse groups
Master's in Early Intervention	Observer	Experience with young children
Master's in education	Moderator	Native Spanish speaker
Bachelor's	Note-taker	Technology background
Master's student	Moderator	Background working with adults and students with disabilities
Master's student	Observer	Background working with adults and students with disabilities
Master's student	Note-taker	Background working with adults and students with disabilities
Bachelor's student	Moderator	Native Spanish speaker

Design and Operational Procedures

Conducting the Focus Groups

All the focus groups were held in the evening in an easily accessible location on a local university campus. Refreshments and childcare were provided. Focus group sessions lasted an hour and a half. After a short introduction, participants spent fifteen minutes working individually on the computer-delivered math tests so they were familiar with the various aspects of the design about which they would be asked questions. In that time, they were able to complete the three basic skills placement tests (one on using a computer mouse, one on silent reading rate, and one on reading comprehension) as well as work through a variety of math items. The computer interface gave participants the opportunity to click on a speaker icon to have the computer read items aloud to them. In the Spanish-speaking English-language learner groups, they also were shown different possible ways to display the math items: English only, bilingual side-by-side, and a technological solution to dual language presentation with collapsible language alternatives.

Three members of the research team participated in each focus group, serving as (a) moderator, (b) note-taker, and (c) field observer. Each of the roles had clearly outlined responsibilities. Moderators were instructed to greet participants as they arrived, make introductions, and establish a comfortable environment. During the focus groups, they were to ensure that everyone had the opportunity to speak and to prevent anyone in the group from monopolizing the conversation. Moderator training included guidelines on the types of responses appropriate for guiding the conversation without influencing people’s contributions and suggestions for prompting participants to explain their responses to the prepared questions in greater depth. Moderators also kept track of the amount of time being spent on each question to ensure that all questions were adequately sampled. A copy of the Moderator Handbook used in the training can be found in Appendix B.

Note takers were on hand to capture the discussion as well as to record nonverbal data such as group reactions to particular questions and information about context and setting. Note takers were instructed to remain uninvolved in the discussion. Appendix C shows the note takers’ instruction sheet. To allow for additional triangulation of data sources, field observers also gathered information about group interactions, context, and setting (See Appendix D). In addition, it was the field observer’s responsibility to ensure that the audio equipment was in place and functioning well before each session and to assist participants as needed while they were interacting with the computer delivered math test.

Data Preparation and Analysis

We triangulated data sources as well as data analysis. Triangulation of data sources took place in two ways. First, we structured the focus groups to allow for multiple observers and ensured that we captured each of their perceptions. Three members of the research team participated in each focus group. While a trained moderator facilitated the group interaction, an observer took field notes, and a note-taker captured participants' comments. Immediately after each focus group session, the three researchers met to discuss their observations. Information from these discussions was included in the observer's field notes when the discussions produced novel insights. In addition, each focus group was audio taped, and the tapes were transcribed. Observers' field notes, note-takers' typed comments, and transcriptions of the audio recordings were all considered during analysis, a further source of data triangulation.

We also triangulated the analysis of the data. Analysis was conducted independently by two researchers who later compared their results to provide a measure of reliability. They first met to agree upon a strategy and decided to use the constant comparative approach (Boeije, 2002). Using this approach, their analysis proceeded inductively. They first read through all of the transcripts and field notes and then began a systematic re-reading with analysis.

Analysis began with data from the focus group comprised of special needs students and then moved to the group comprised of special needs students' parents. Analysis continued with data from the general education student group, followed by the parents of general education students, child advocacy group members, teachers, and administrators, and finally Spanish-speaking English language learners and their parents. There was no particular significance for the order in which groups were analyzed. As themes began to emerge, the researchers went back to the data to look for additional instances of these themes as well as for contradictory data.

Because the purpose of these focus groups was to solicit opinions from a variety of people, the researchers took great care to note both generally supported themes and opinions offered by only one or a few participants. Contradictions between opinions shared by different people or groups of people also were noted. These contradictions provided researchers with insights into questions to pursue further. In addition, researchers pinpointed several areas of confusion that would benefit from additional research.

Participants' responses on short written surveys completed at two points during the focus group activity served as a further source of data. Compiled survey responses were used by one of the researchers to check the conclusions reached through the earlier analysis of other data sources. There was a good and consistent match between the different data sources and the two researchers' findings.

Results

Analysis of the field notes, note taker's comments, and transcripts yielded two kinds of information: suggestions that could immediately be implemented to improve the design and delivery of the UDA mathematics test, and contradictory results indicating a need for continued research. The suggestions that emerged as clear direction for our research team could be grouped into three broad categories. The first category included suggestions to make it easier for people who had difficulty using the mouse during their computer interaction. The second category included suggestions to reduce areas of uncertainty or confusion in the directions provided during the computer delivered test. The final category focused on the format of the test, on how items looked on the screen, and ways to make the assistive devices more accessible (such as the speaker icon that would prompt the computer to read a particular item aloud). Triangulated results from the discussions that took place during the focus groups are provided in Appendix E.

Survey Results

In addition to participating in a group discussion about the UDA math test, focus group participants filled out surveys, both before and after interacting with the computer delivered math test. These surveys provided the research team with additional information about the participants' familiarity with computers as well as corroborating evidence to support the opinions shared during the focus group discussions. The results of the student surveys are presented in Tables 1 – 2; Tables 3 -5 display the results of the parent surveys; and Tables 6-9 present the results of the child advocacy group members and teacher surveys.

Table 1
Student Responses on the Survey Filled Out Before Focus Group Sessions

Question	Group	Frequency of Response in Each Possible Category				
		never	once a month	2-3 times a week	once a week	every day
How often do you use computers anywhere?	English	0	0	5	3	1
	ELL	3	1	1	0	0
How often do you use a computer at home?	English	0	3	1	3	2
	ELL	5	0	0	0	0
How often do you use a computer at school?	English	3	1	2	3	0
	ELL	0	0	2	4	0

Table 2
Student Responses on the Survey Filled Out After Focus Group Sessions

Question	Group	Frequency of Response in Each Possible Category		
		easy	just right	hard
How easy was it to understand the person talking on the computer?	English	6	3	0
	ELL	3	2	0
How easy was it to use the speaker button?	English	5	4	0
	ELL	3	2	0
Was the story too easy, just right, or too hard?	English	2	6	1
	ELL	1	4	0
Was the fill-in-the-blank too easy, just right, or too hard?	English	3	5	1
	ELL	0	5	0
		never	sometimes	every time
How often did you click on the speaker button?	English	1	3	5
	ELL	0	4	1
		too small	just right	too big
Was the 'next' button too small, just right, or too big?	English	0	9	0
	ELL	0	5	0
Were the numbers and words on the screen big enough to read easily?	English	3	6	0
	ELL	0	5	0
		☹	😐	☺
How did you feel about taking the computer test?	English	0	1	7
	ELL	0	0	5

Table 3

Parent Responses on the Survey Filled Out Before Focus Group Sessions

Question	Group	Frequency of Response in Each Possible Category				
		never	once a month	2-3 times a week	once a week	every day
How often does your child use the computer at home?	English	0	1	2	5	2
	ELL	0	0	2	0	0
How often does your child use a computer at school?	English	0	0	5	2	1
	ELL	0	0	1	0	0
		never	sometimes	often	usually	always
When your child uses a computer, how often do you usually help him or her?	English	1	6	2	1	0
	ELL	2	0	0	0	0
		no experience	little experience	experienced	quite experienced	very experienced
How experienced is your child with computers?	English	0	2	3	5	0
	ELL	0	3	0	0	0
		Yes		No		
Do you have a computer at home?	English	10		0		
	ELL	2		2		

Table 4

Parent Responses on the Survey Filled Out During Focus Group Sessions

Question	Group	Frequency of Response in Each Possible Category				
		not clear		just right		clear
How clear was the computer voice?	English	1		2		7
	ELL	0		0		4
When selecting an answer, was it clear that a change in color meant that you had selected an answer?	English	4		3		3
	ELL	2		0		0
		never		sometimes		every time
How often did you use the speaker icon?	English	8		1		0
	ELL	0		0		3
		very clear	clear	just right	confus-ing	very confusing
Were the directions [on the computer test] clearly written?	English	4	3	0	1	0
	ELL	1	3	0	0	0
		very easy	easy	just right	challeng-ing	very challeng-ing
Were the reading test items too easy, just right, or too challenging?	English	2	0	3	4	1
	ELL	0	3	0	0	0
		very short	short	just right	long	very long
Was the timing of the Fill-in-the-blank test appropriate for your student?	English	0	1	6	1	0
	ELL	0	1	2	0	0
		much less	less	no change	more anxiety	much more
Do you think your child would experience less or more test anxiety while using the computer to take a test instead of paper and pencil?	English	0	6	1	3	0
	ELL	0	1	0	0	0
		much worse	worse	no difference	better	much better
Do you think your child would perform better or worse on a computer test versus a standard paper-type test?	English	0	2	1	6	1
	ELL	0	0	0	1	0

Table 5

Child Advocacy Group Members and Teacher Responses on the Survey Filled Out After Focus Group Sessions

Question	Frequency of Response in Each Possible Category				
	very clear	clear	just right	confusing	very confusing
Were the directions clearly written?	7	1	1	2	0
	very easy	easy	just right	difficult	very difficult
How easy would it be for students to understand what the computer voice was saying?	3	1	4	1	1
Were the reading tests too easy, just right, or too challenging for 3rd grade students?	0	0	5	3	2
	very short	short	just right	long	very long
Was the timing of the Fill-in-the-blank test appropriate for students?	1	4	4	2	0
	much less anxiety	less anxiety	no change in anxiety	more anxiety	much more anxiety
Do you think students would experience less or more test anxiety while using the computer to take a test instead of paper and pencil?	0	3	4	1	1

Table 7

Teacher Responses on the Survey Filled Out During Focus Group Sessions

Question	Frequency of Response in Each Possible Category			
	much different access	different access	no difference in access	don't know
In your opinion, to what extent do the different student groups have differing levels of access to computers while at home?	5	4	1	1

Question	Frequency of Response in Each Possible Category		
	much different access	different access	no difference in access
To what extent do the different student groups have different levels of access to computers while at school?	1	2	8
	never	sometimes	every time
How often did you use the speaker icon?	7	0	2
	yes	no	not sure
Would 3rd grade students understand how and why to use the speaker icon?	3	2	3

Question	Frequency of Response in Each Possible Category	
	yes	no
In your opinion, should the computer give immediate feedback to students regarding the correctness of their responses?	3	6

Discussion

Focus groups are critical elements in the design of universal environments. Through the use of focus groups it is possible to determine the range of personal characteristics of the proposed users. By identifying these characteristics, it is possible to design an environment that accounts for unique features of the users to provide maximum opportunities for successful interactions. In the current study, we used focus groups to explore the needs of the target population for a universally designed mathematics test. We solicited input from teachers, administrators, and members of child advocacy groups, as well as students and parents of 3rd grade students from 3 groups: students with disabilities, students in general education, and students learning English as a second language.

The purpose of this study was to identify the features of our universally designed mathematics test that could be improved upon to enhance the usability of the program. Several themes emerged from the results of these focus groups. First, many students and adults reported having difficulty using the mouse due to the confusion with the right- and left- click buttons. Second, the directions were noted as being confusing or lacking critical information. Suggestions included clarifying the language and providing additional information about resources available to the student such as scratch paper and the read-aloud option. Finally, the format of the items was distracting (e.g., read-aloud voice talked too fast), did not clearly convey the information (e.g., font size and “next” button were too small), and did not provide enough external support (e.g., no “help” button was available). See Appendix E for a detailed summary of the findings. These environmental features interfere with the assessment of the targeted skills and thus cause construct irrelevant variance. Subsequent designs of the mathematics test must mediate these effects by changing the design or providing additional support. Supplemental reviews are necessary by additional members of the educational community to verify these findings.

An additional purpose of the focus group was to better understand the personal characteristics of the users. As such, we asked the participants about the quality and quantity of students' interaction with computers. Each group was asked independently of the others to determine if differences in access to computers existed across the subpopulations. One finding to note is that the ELL and English only students seem to have differential access to computers at home. This may have implications for the interpretation of their scores obtained on tests delivered by computer due to the influence of differential computer skills.

By using focus groups we are able to identify features of the universally designed mathematics test that inhibited student performance. This research methodology provides a unique avenue for gathering valuable information to improve our practices.

References

Boeije, H. (2002). A purposeful approach to the constant comparative method in the analysis of qualitative interviews. Quality and Quantity. 36, pp. 391-409.

Center for Universal Design. (1997). What is universal design? NC State University: The Center for Universal Design.

Ketterlin-Geller, L. R. (2003). Establishing a validity argument for universally designed assessments. Unpublished Doctoral Dissertation, University of Oregon, Eugene, OR.

Krueger, R. A. (1998). Analyzing and reporting focus group results. Focus Group Kit (Vol. 6) Thousand Oaks, CA: Sage.

Morgan, D. L. & Krueger, R. A. (1998). The Focus Group Kit (Vols. 1-6). Thousand Oaks, CA: Sage.

Patton, M. Q. (2002). Qualitative Research and Evaluation Methods (3rd ed.). Thousand Oaks, CA: Sage.

Vaughn, S., Schumm, J. S., Sinagub, J. (1996). Focus group interviews in education and psychology. Thousand Oaks, CA: Sage.