Verifying survey items for Construct Validity: A two-stage Sorting Procedure for Questionnaire Design in Information Behavior Research

Naresh Kumar Agarwal
Graduate School of Library and Information Science, Simmons College
300 The Fenway, Boston MA 02115
T +1 617 521 2836 F +1 617 521 3192 agarwal@simmons.edu

ABSTRACT
Construct validity refers to the degree to which inferences can legitimately be made from the operationalizations in a study to the theoretical constructs on which those operationalizations were based. When both convergent and discriminant validities are satisfied, construct validity is said to be satisfied. Verifying survey items for construct validity becomes important, especially when these items are self-developed and not based on questionnaires used in past studies. Yet, there is lack of clarity on how to go about doing so. While researchers in other fields have come up with simple techniques, these have not been readily applied to the field of information behavior. Using the case of questionnaire design for source choice in information behavior, this methodological paper describes the use of a 2-stage sorting procedure based on Moore and Benbasat's 1991 work. The procedure can serve as a guide to researchers using questionnaire design for studies in information seeking behavior and should help lend greater rigor to such studies.

Keywords
Construct validity, sorting procedure, unstructured sorting, structured sorting, convergent validity, discriminant validity, source choice, information behavior, context.

INTRODUCTION AND MOTIVATION

Construct validity

'…since the measuring device has been constructed by the observer, …we have to remember that what we observe is not nature in itself but nature exposed to our method of questioning' (Heisenberg, 1958). Constructs are higher level concepts which are not directly observable or measurable (nature) while variables (sometimes used interchangeably with indicators or measures) seek to measure the underlying construct (nature exposed to our method of reasoning). For example, *hard work* can be seen as a construct (not directly measurable), while *number of hours spent working on a research paper* can be seen as a way of measuring hard work. There can be more than one measure or indicator for the same construct. Thus, when we expose nature to our method of questioning, we come up with operational definitions and measures for our constructs. This process is generally understood as the process of operationalization. 'Construct validity refers to the degree to which inferences can legitimately be made from the operationalizations in your study to the theoretical constructs on which those operationalizations were based' (Trochim, 2006). Using our example, it refers to the extent to which number of hours spent working on a research paper adequately measures or represents hard work. 'Construct validity involves generalizing from your program or measures to the concept of your program or measures' (Trochim, 2006). Trochim writes that we might think of construct validity as a 'labeling' issue. 'When you measure what you term "self esteem" (or *hard work*, in our example), is that what you were really measuring?' (Trochim, 2006).

Convergent and Discriminant validities

Surveys are appropriate for research questions about self-reported beliefs or behaviors (Neuman, 2003). When working with multiple constructs in a survey study, it is important to satisfy convergent and discriminant validities in order to satisfy construct validity. 'If you can demonstrate that you have evidence for both convergent and discriminant validity, then you've by definition demonstrated that you have evidence for construct validity. But, neither one alone is sufficient for establishing construct validity.' (Trochim, 2006). To satisfy convergent validity, 'measures of constructs that theoretically should be related to each other are, in fact, observed to be related to each other (that is, you should be able to show a correspondence or convergence between similar constructs).' (Trochim, 2006). In explaining discriminant validity, Trochim writes that 'measures of constructs that

ASIST 2011, October 9-13, 2011, New Orleans, LA, USA.

theoretically should not be related to each other are, in fact, observed to not be related to each other (that is, you should be able to discriminate between dissimilar constructs).' Let us consider two constructs – happiness and love. For happiness, let us come up with three indicators: HAPP1 the number of times one smiles in a day, HAPP2 the degree to which one is content with one's life and HAPP3 the degree to which one helps other people. Similarly, we could come up with a number of measures or indicators for love e.g. LOVE1 and LOVE2. To satisfy construct validity, once data has been gathered from survey respondents, HAPP1, HAPP2 and HAPP3 should have a high degree of correlation with each other (convergent validity), but low degree of correlation with LOVE1 and LOVE2 (discriminant validity). LOVE1 and LOVE2 should also have a high degree of convergence with each other.

Motivation for the paper

The basic steps in survey research involve formulating the study objectives, selecting data collection techniques (interviews or questionnaires), developing the survey instrument, pretesting the questionnaire with expert judges, pilot testing the questionnaire with a sample from the study population, analyzing the data for convergent and discriminant validities, collecting the main data from the study population and entering it in an excel sheet, analyzing the gathered data and writing and disseminating results. This study relates to the *developing the survey instrument* stage of survey research. In designing a questionnaire during survey research, it is recommended that survey questions (referred to as items) are based on what was tested in past studies (see e.g. Stone, 1978). This helps enhance validity. For each item or question relating to construct or variable, it is good practice to write the source or past study from which the items were taken or adapted, or to write *self-developed* if done so (see Figure 1).

Construct / Variable 1	Survey item 1	Source
	Survey item 2	Self-developed
	Survey item 3	Source; source
	Survey item 4	Source
	Survey item 5	Source
Construct / Variable 2	Survey item 1	Source
	Survey item 2	Self-developed
….	….	….

Figure 1 Constructs/variables, related items and where they're taken from

Verifying survey items for construct validity becomes important, especially when these items are self-developed and not based on questionnaires used in past studies. Yet, there is lack of clarity on how to go about doing so. While researchers in other fields have used techniques for this (e.g. Davis (1986, 1989), Moore and Benbasat (1991) and Kankanhalli, Tan and Wei (2005) have used a sorting procedure in their studies in the field of information systems), these have not been readily applied to the field of information behavior.

The Case

When faced with a need for information to fulfill a task at work, a person often asks a friend or a colleague or goes to a website, a search engine, or refers to a book, manual, etc. Each of these may be understood as a source of information. 'An important question in information-seeking behavior is where people go for information and why information seekers prefer to use one source type rather than another when faced with an information-seeking task or need for information' (Agarwal, Xu and Poo, 2011, p.1087). Agarwal, Xu and Poo (2011) designed a questionnaire and conducted a survey of 352 working professionals in Singapore to study the contextual factors affecting the use of different types of sources by information seekers in a work environment. The variables (or constructs) they studied included:

- Variables of the task or environment – task importance, task urgency, task complexity
- Variables pertaining to the source – source quality, access difficulty, communication difficulty
- Variables pertaining to the seeker-source relationship – inherent lack of comfort
- Control variables pertaining to the seeker – seeker's learning orientation, task self efficacy, tenure in the role, tenure in the organization, gender, age, education
- Control variables pertaining to the environment – favorable learning environment, size of team

They examined the effect of these variables on source use. See Agarwal, Xu and Poo (2011, p.1089) for their research model.

Objective of the paper

Using the case of questionnaire design for this study on source use in information behavior by Agarwal, Xu and Poo (2011), this methodological paper describes the use of a 2-stage sorting procedure to verify construct (convergent and discriminant validities) in the survey instrument design stage of the study. This procedure is based on the one used by Moore and Benbasat (1991), who had extended it from an earlier work by Davis (1986, 1989). Kankanhalli, Tan and Wei (2005), among others, have used the procedure in their study. The procedure is described briefly in Agarwal, Xu and Poo (2011, p.1092) and is discussed in detail in this paper.

The procedure can serve as a guide to researchers using questionnaire design for studies in information seeking behavior and should help lend greater rigor to such studies.

INSTRUMENT DEVELOPMENT AND OPERATIONALIZATION OF CONSTRUCTS

To operationalize their constructs, Agarwal, Xu and Poo (2011) came up with definitions of constructs in their research model based on past literature (see Table 1).

Attribute of	Construct	Abbreviation	Definition
Source	Use	USE	The degree to which the person actually utilized the information source. They examined 3 aspects of source use – frequency of use of each source, percentage of time each source was utilized and the order of utilizing each source as compared to other sources (of other source types).
	Quality	QUA	The novelty, reliability, breadth and depth of information content the source carries that has applicability and relevance to the task at hand
	Access Difficulty	ACC	The time and effort required, and the difficulty encountered in reaching a particular information source i.e. establishing the channel of communication with the source (*before* the person actually starts using it)
	Communication Difficulty	CMM	The difficulty in interacting with, conversing with, and understanding the information source (once the person has reached the source)
Problem situation / Task	Importance	IMP	Importance of the outcome of the task with the seeker's well being
	Urgency	URG	Need to accomplish the task sooner than later
	Complexity	CMP	The degree to which a task is challenging, difficult to understand, requires considering many aspects and takes a long time to learn
Environment	Learning Environment	ENV*	The degree to which the seeker's work environment is favourable for learning and information seeking.
Seeker / Actor	Learning Orientation	ORT*	The seeker's attitude towards learning – degree to which the seeker believes that his/her competence can be improved
	Task Self Efficacy	EFF*	The degree to the information seeker considers himself/herself an expert in doing the task at hand
Seeker-Source relationship	Inherent Lack of Comfort	CFT	The degree to which the seeker feels uncomfortable, nervous, embarrassed or concerned about his/her image/self-image before using an information source. Here, image refers to the way another person(s) sees the seeker. Self-image refers to the way the seeker sees himself/herself.

* Control Variables

Table 1 Definition of Constructs

As suggested by Stone (1978), wherever possible, survey items (questions) were adapted from prior studies to enhance validity. When items were not available, new questions were developed based on survey of literature. See Agarwal, Xu, and Poo (2011, p.1095-1096) for the items and their sources from past literature.

A TWO-STAGE SORTING PROCEDURE FOR VERIFYING SURVEY ITEMS FOR CONSTRUCT VALIDITY

Since some of the survey items were self-developed (while the rest were adapted from prior studies), Agarwal, Xu and Poo, (2011) followed three steps to verify the construct validity. 'First, experienced researchers in the field were consulted to discuss the wording of each item. The instrument was revised based on the feedback collected' (p.1092). Then, they subjected all their survey items to a two-stage conceptual validation exercise using the procedure recommended by Moore and Benbasat (1991) (see Kankanhalli, Tan and Wei 2005 for an example of prior usage of this procedure). The first stage consisted of 'unstructured sorting', and the second stage 'structured sorting' of all items in the survey.

Unstructured Sorting

In the first stage, 4 graduate students (let us call them *judges*), 'who were not informed of the research model and constructs were invited to sort all items into an unrestricted number of categories. They were also asked to name each category.' (p.1092). First, they were given an envelope each. Each envelope consisted of strips of paper, with a survey item printed on each strip. There were a total of 52 strips in each envelope (one for each survey item) to cover the 52 initial survey items belonging to 11 constructs in total (including 3 constructs for control variables). The strips in the envelope were randomly mixed. Each judge was asked to take out the strips and put them in different piles such that each pile had related questions (those that measured a common construct). If any question appeared to belong to more than one pile, the judges were asked to place them in a separate pile. After completing the grouping exercise, the judges were also supposed to give a name or

label (as well as definition) to each pile (to indicate the construct that all the items in that pile measured), and fill a form shown in Figure 2. A mock example using hypothetical items from a hypothetical research model was also shown to each judge to demonstrate how to conduct the sorting exercise.

This process was very useful in identifying ambiguously-worded survey items/questions. The names/labels given by the judges for the different piles were very close to the names of the actual constructs. As shown in Table 2, the four judges correctly placed close to 87% of the survey items into their rightful construct piles.

Your Name_____
Contact No. _____

Label	Definition	Pile No. (staple each pile and assign a number to it)	No. of items in pile
:			

Total No. of items in all piles (fill in the end) _____

Figure 2 Form each judge filled for unstructured sorting

Target Category	Actual Category												Total Qs	Hit Rate (%)
	USE	QUA	CFT	CMM	ACC	ENV	EFF	CMP	IMP	URG	ORT	Other		
USE	19											1	20	95
QUA		19										1	20	95
CFT			13		5							2	20	65
CMM		1		5	6								12	41.67
ACC		1			19								20	95
ENV						17						3	20	85
EFF							20						20	100
CMP								19				1	20	95
IMP									16				16	100
URG							1			19			20	95
ORT											18	2	20	90
AVERAGE														86.97

Table 2 Results of Unstructured Sorting Exercise

After the first round of sorting, a number of items from different constructs were altered because they were either ambiguous or did not fit well with the other items in the pile. Table 3 shows the items changed after the first round of sorting i.e. unstructured sorting. Portions of items changed (to remove ambiguity) are highlighted in **bold**. On the suggestion of judges, 5 items were added after the first round of sorting. These were IMP4 (task importance), CFT6 (inherent lack of comfort), CMM1 / CMM3 (communication difficulty with the source) and USE1 (source use). Kim and Mueller (1981) posit that it is desirable to have 3 or more items per construct to ensure better measurement properties for each construct. These additions took the total number of items/questions from 52 to 57.

Construct	ItemCode	Item before unstructured sorting	Item after unstructured sorting
Access Difficulty	ACC2	It would be very **difficult** to get to [].	It would be very hard to get to [].
Learning Environment	ENV2	Everyone around me **is asking** for information s/he needs.	In my organization, everyone around me feels free to ask for information s/he needs.
Task Self Efficacy	EFF5	I have good knowledge about the task.	I have good knowledge about **this** task.
Task Complexity	CMP4	I need to consider so many factors to do this task.	This **task requires me** to consider so many factors.
Task Importance	IMP4*		I give a lot of **weightage** to this task.
Task Urgency	URG5	I have enough time to accomplish this task.	I **do not have** enough time to accomplish this task.
Source Quality	QUA3	[] has novel (new) knowledge which **can be used to solve** the problem.	[] has novel (new) knowledge related to the problem.
	QUA4	[] has reliable knowledge which	[] has reliable knowledge relevant to the

		can be used to solve the problem.	problem.
Inherent Lack of Comfort	CFT2	**It is embarrassing** to use [] for information.	I would be embarrassed to use [] for information.
	CFT5	I do not have adequate knowledge about [].	Using [] will hurt my image (the way another person(s) sees me).
	CFT6*	-	Using [] will hurt **my self-image** (the way I see myself).
Communication Difficulty	CMM1*	-	It is difficult to reach a **common understanding** of the problem with [].
	CMM2	It is not easy to get immediate feedback from [].	It is difficult to communicate with [].
	CMM3*	-	It is difficult to **explain** my problem to [].
	CMM4	It is difficult to **reach a common understanding** of the problem with [].	It is difficult to make [] understand my problem most of the time.
	CMM5	It is **not easy to** extract information from [].	It is difficult to extract useful information from [].
Source Use	USE1*	-	Among all the sources of information available to me, **I used** [] for problem-solving information.
	USE2	For information to solve this problem, I used [] very frequently.	I used [] very frequently for problem-solving information.

*Item added after unstructured sorting
[] indicates an information source of a particular type e.g. face-to-face source, online source, etc. (see Agarwal, Xu and Poo, 2011).

Table 3 Changes to survey items after unstructured sorting

Target Category	Actual Category												Total Qs	Hit Rate (%)
	USE	**QUA**	**CFT**	**CMM**	**ACC**	**ENV**	**EFF**	**CMP**	**IMP**	**URG**	**ORT**	**Other**		
USE	16	1										7	24	66.667
QUA		17										3	20	85
CFT			24										24	100
CMM				17								3	20	85
ACC					20								20	100
ENV						19					1		20	95
EFF							20						20	100
CMP		1					1	16				2	20	80
IMP							1	1	18				20	90
URG										19		1	20	95
ORT											20		20	100
AVERAGE														90.606

Table 4 Results of Structured Sorting Exercise

Structured Sorting

Four more students participated as judges in the second round of sorting. Here, the job of the judge was made easier than that of the judge in the first round. This is because the number of categories was specified beforehand, and the labels and definitions provided for each construct (see Figure 3). All the judge had to do was to determine which label and definition each item best conformed to. A 'does not fit' category was also provided for putting the items that the judge thought did not fit in any of the specified categories.

Your Name_____
Contact No. _____

Pile No. (staple each pile and assign a number to it)	Label	Definition	No. of items in pile
1.	Source	The degree	

	Use	to …	
2.	Source Quality	The novelty, rel…	

:

11.	Task Complexity	The degree to …	
00.	Does not fit	Item(s) that do..	

Total No. of items in all piles (fill in the end) _____

Figure 3 Form each judge filled for structured sorting

Each judge was given the 57 reworded items (printed in individual strips of paper, mixed and put inside an envelope). This time, all judges correctly placed close to 91% of items in the correct construct piles (see Table 4).

For items placed in the 'does not fit' category or in wrong categories, further changes were made to remove ambiguity. +Item added after structured sorting

Table 5 shows the items changed after the second round i.e. structured sorting. Portions of items changed are highlighted in **bold**. One item QUA6 was added for source quality to cover the depth of knowledge that an information source has. This took the total number of items to 58 for the final questionnaire.

Construct	ItemCode	Item before structured sorting	Item after structured sorting
Task Complexity	CMP4	This task requires me to consider so many **factors**.	This task requires me to consider so many aspects.
Task Urgency	URG5	I **do not have enough time** to accomplish this task.	There is a pressing need to get this task done soon.
Learning Orientation	ORT5	I **believe** in life-long learning.	I continuously work towards upgrading my knowledge and skills.
Learning Environment	ENV1	My **organization encourages** me to seek knowledge.	In my organization, we always ask each other for work-related knowledge.
	ENV4	My **organization encourages** me to share knowledge.	Most colleagues in my organization are ready to share their knowledge.
Source Quality	QUA5	[] has **broad** knowledge related to the problem.	[] has broad/wide knowledge related to the problem.
	QUA6^{+}	-	[] has **deep** knowledge related to the problem.
Communication Difficulty	CMM1	It is difficult to **reach a common understanding** of the problem with [].	While using [] for my problem, the "conversation" with [] is painful.
	CMM2	It is difficult to **communicate** with [].	While using [] for my problem, it is difficult to "converse" with [].
	CMM3	It is difficult to explain my problem to [].	**While using [] for my problem**, it is difficult to explain to [].
	CMM4	It is difficult to make [] understand my problem most of the time.	**While using [] for my problem**, it is difficult to make [] understand most of the time.
	CMM5	It is difficult to extract useful information from [].	**While using [] for my problem**, it is difficult to extract useful information from [].
Inherent Lack of Comfort	CFT5	Using [] **will hurt** my image (the way another person(s) sees me).	Using [] will not be nice for my image (the way another person(s) sees me).
	CFT6	Using [] **will hurt** my self-image (the way I see myself).	Using [] will not be nice for my self-image (the way I see myself).
Source Use	USE1	Among all the sources of information available to me, I used [] for problem-	Among all the sources of information available to me, I

		solving information.	used [] **a lot** for problem-solving information.
	USE2	I used [] very **frequently** for problem-solving information.	I used [] very often for problem-solving information.
	USE3	How **often** did you use [each of] the following sources for this problem? (very infrequently.. …very frequently)	How frequently did you use the following sources for this **specific problem/part of the task**? (very infrequently… …very frequently)
	USE4	How **frequently** did you use the following sources for this problem? Tick the appropriate choice (only one) for each source (didn't use at all; used less than once each week; used about once each week; used several times a week; used about once each day; used several times a day).	I used [] (several times a day; about once a day; several times a week; about once a week; **about once in 2-3 weeks**; **less than (once in 2-3 weeks)**; didn't use at all)

[+]Item added after structured sorting

Table 5 Changes to survey items after structured sorting

'Finally, a pretest was conducted to finetune the instrument, whereby the survey was administered on 12 graduate students. They were asked to comment on their understanding of each item after filling out the survey. Minor revisions were made based on their feedback.'(Agarwal, Xu and Poo, 2011, p.1092). Agarwal, Xu and Poo then go on to describe the steps of pilot data collection and main data collection in their survey research.

CONCLUSIONS AND IMPLICATIONS

As seen in Table 2 and Table 4, while there is no explicit cut-off, higher the percentage of items placed in the correct category (hit rate), the better it is. An average hit rate of above 90% after the second step of the sorting process should enable one to proceed to the next step. An average hit rate below 80% in either step may prompt one to reword the questions and repeat the sorting step until one gets a better hit rate. Those constructs which show a hit rate of below 80% (e.g. CFT and CMM in Table 2 and USE in Table 4) might need further examination. It is, eventually, for the researcher to make a call when to proceed to the next step (ideally when the overall hit rate is above 90%, at least after the second step).

Something that is to be kept in mind is that the procedure explained in this paper is one in a series of iterative processes when designing a questionnaire. Such a process may include: 1) pretesting a questionnaire with the help of experts to point out any problems in the wording or sequence of questions; 2) two step sorting procedure as described in this paper; 3) doing a pilot data collection and carrying out an exploratory factor analysis (see Anderson and Gerbing, 1988) to check for factor loadings and degree of adherence to convergent and discriminant validities; 4) doing the main data collection and carrying out confirmatory factor analysis (see, for example, Agarwal, Xu and Poo, 2011). After each of these processes (including each step of sorting), the questionnaire may be revised and the preceding process repeated if needed. Proceeding to the next step is recommended only when the researcher is satisfied with the current process. Thus, each process is iterative.

One might wonder why not go on to the pilot testing stage right away without doing the additional sorting stage described here. The difference is that the pilot testing stage is expensive. One has to reach out to a number of participants from the sample under study and a minimum sample size will need to be obtained in order to factors to load correctly in exploratory factor analysis. Then, if changes in the questionnaire are needed, one has to discard all the data gathered and might not be able to reach out to the same set of participants as well (to minimize the maturation threat to internal validity). On the other hand, the sorting procedure can be done with very few judges who need not be from the study sample.

Thus, the two-step sorting procedure is very helpful in verifying construct validity especially when new survey items have been developed. It helps simplify the future steps of exploratory factor analysis when analyzing pilot data and confirmatory factor analysis when analyzing the data from the main study in a survey. It can easily be applied to future survey studies in information behavior and can serve as a guide to researchers. It should help lend greater rigor to future studies. Further work can be done in coming up with an online interface to simulate this sorting procedure.

REFERENCES

Agarwal, N.K., Xu, Y. (C.), Poo, D.C.C. (2011). A Context-based Investigation into Source Use by Information Seekers. *Journal of the American Society for Information Science and Technology*, 62(6), 1087-1104.

Anderson, J.C. and Gerbing, D.W. (1988). Structure equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, 103(3), 411-423.

Davis, F.D. (1986). A Technology Acceptance Model for Empirically Testing New End User Information Systems: Theory and Results. Unpublished Doctoral Dissertation, Massachusetts Institute of Technology.

Davis, F.D. (September, 1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology, *MIS Quarterly*, 13(3), 319-340.

Heisenberg, W. (1958). The Copenhagen Interpretation of Quantum Theory (Chapter 3). In *Physics and Philosophy*. George Allen and Unwin Edition, 1959. Retrieved May 31, 2011 from http://www.marxists.org/reference/subject/philosophy/works/ge/heisenb3.htm

Kankanhalli, A., Tan, B.C.Y. and Wei, K.-K. (March, 2005). Contributing knowledge to electronic knowledge repositories: An empirical investigation. *MIS Quarterly*, 29(1), 113-143.

Kim, J. and Mueller, C.W. (1981). *Factor Analysis: Statistical Methods and Practical Issues*, Newbury Park, CA: Sage Publications.

Moore, G.C. and Benbasat, I. (1991). Development of an instrument to measure the perceptions of adopting an Information Technology Innovation. *Information Systems Research*, 2(3), 173-191.

Neuman, W.L. (2003). Social Research Methods: Qualitative and Quantitative Approaches, 5th Edition, Pearson: Boston, MA.

Stone, E.F. (1978). *Research Methods in Organizational Behavior*. Goodyear: Santa Monica, CA.

Trochim, W.M. (2006). The Research Methods Knowledge Base, 2nd Edition. Retrieved May 31, 2011 from http://www.socialresearchmethods.net/kb/