

UNDERSTANDING THE PROCESS OF CONSTRUCTING SCALES INVENTORIES IN THE PROCESS MODELLING DOMAIN

Recker, Jan, Queensland University of Technology, 126 Margaret Street, Brisbane QLD
4000, Brisbane, Australia, j.recker@qut.edu.au

Rosemann, Michael, Queensland University of Technology, 126 Margaret Street, Brisbane
QLD 4000, Brisbane, Australia, m.rosemann@qut.edu.au

Abstract

Empirical research strategies are undisputedly of paramount importance to rigorous and relevant IS research. Employed practices of empirical research in Information Systems, however, are still considered to be problematic. One related problem in this context is the lack of rigorously developed and tested empirical scales inventories that could be used in empirical studies of artefacts relevant to Information Systems. This paper reports on the process of developing research tools for usage in empirical studies of Information Systems artefacts. We use the domain of business process modelling as an example of a relevant IS research domain and report on the development of an scales inventory to measure the various perceptions that individuals may have towards acceptance and continuance of process modelling languages. We describe a multi-stage development approach for scales creation that incorporates feedback from both expert and user perspectives. Our research in progress results in a pre-validated scales inventory that can be used to assist in further empirical studies that study phenomena associated with the business process modelling domain.

Keywords: Process modelling, empirical research, perception measurement

1 INTRODUCTION

This paper reports on research in progress towards the development of an instrument designed to measure the perceptions that individuals may have towards the use of process modelling languages (also known as process modelling notations) for conceptual modelling. It specifically discusses the process of developing such a measurement instrument. As such, it seeks to address a widespread challenge in IS research, namely to develop and conduct empirical studies by means of rigorous and theoretically sound instrument development (Straub, 1989).

IS research is concerned with the investigation of the management, development, operation, use and impact of information systems in organizations (Zmud & Boynton, 1991). This discipline is quite uniquely placed interface of technology and organization, *i.e.*, it addresses the interaction in human-machine systems (Lee, 2001). Accordingly, IS research can be, should be, and is being addressed via many paradigms and multiple research strategies, both of theoretical and empirical nature (McGrath, 1981). While theoretical contributions to IS research are not within our scope, we are concerned with troubles and concerns facing IS researchers conducting empirical studies of IS artefacts. In this regard, from a retrospective, empirical research results in IS have often been mixed and inconclusive (Moore & Benbasat, 1991). One of the noted areas of concerns is related to the lack of theoretical rigour in the development of measurement arrays (or scales inventories) that are sought to tap the explanatory concepts of underlying theoretical models. In fact, the lack of theoretical foundations and the wide range of measurements used by IS researchers without adequate rigorous theoretical justification has repeatedly been noted to be a major cause for a rather incomplete state of empirical knowledge in IS (Kwon & Zmud, 1987; Davis et al., 1989). Not surprisingly, measurement issues are receiving increased attention amongst IS researchers (Doll & Torkzadeh, 1991).

Poor theory development and lacking methodological rigor in measurement studies have plagued IS research in a wide variety of topics (Dickson et al., 1984). It is also clear that theoretically well-founded and rigorously developed construct operationalizations are a pre-requisite for the beginning of a cumulative tradition (Moore & Benbasat, 1991). Not surprisingly, in some areas have papers in leading IS journals reported on the development of domain-appropriate scales inventories, *e.g.*, (Doll & Torkzadeh, 1998; Burton-Jones & Straub, 2006).

In spite of the importance of the research phase of measurement development, which is vital to any empirical research study (Froehle & Roth, 2004), several popular and relevant domains of IS research lack rigorous development procedures as well as reliable and valid scales inventories.

This is especially the case in research of phenomena ascribed to the area of conceptual modelling, which is arguably a popular and relevant area of IS research (Wand & Weber, 2002). Especially in the area of IS analysis and design is conceptual modelling widely established as both an inevitable and promising means (Karimi, 1988) for representing requirements and other domain phenomena in the form of intuitive graphical models (Siau, 2004). Yet, this relevant research domain has traditionally been facing the problem of lacking mature theoretical foundations, *e.g.*, (Weber, 2003), upon which empirical research strategies could be based and from which testable propositions could be generated. In fact, most of the existing approaches for conceptual modelling have been developed on the basis of practical wisdom rather than on a scientific theory (Bubenko, 1986). This insight holds even more so in the area of conceptual modelling for business process management, one of the most popular application areas of conceptual modelling overall (Davies et al., 2006). Many studies have shown the relevance of process modelling to BPM initiatives, *e.g.*, (Davenport, 1993). The recent introduction of legislative frameworks such as the Sarbanes-Oxley Act further contributed to the increasing interest in business process modelling as a way of capturing and graphically documenting the processes of an organization.

Given the popularity of process modelling it is not surprising that, in IS research, a wide range of scholarly articles reports on, and discusses, the role of phenomena related to process modelling, see,

for instance, (Curtis et al., 1992; Bandara et al., 2005). However, by far the largest share of research findings in this space is of conceptual or theoretical nature. In fact, the share of empirical papers is well less than twenty per cent (Moody, 2005). In particular, there is a paucity of empirical studies on process modelling that would investigate success measures of process modelling, such as the acceptance of tools, methods and languages, an individual's attitude towards process modelling or the factors critical to the success of process modelling languages.

Accordingly, our overall research objective is to study the acceptance, and ultimately success, of process modelling in IS practice. While this research objective remains a long-term goal, at the moment our research is driven by the insight that a better understanding of process modelling can only be achieved by means of appropriate empirical research strategies. These, in turn, are dependent on the availability of adequate empirical research tools, *i.e.*, valid and reliable scales inventories to assist in empirical studies on process modelling and related phenomena. Given the aforementioned problems in IS research related to devising and using such measurement tools, the *aim of this paper* is to report on, and discuss, a rigorous and theoretically grounded process for developing a scales inventory for usage in IS research. We use the scenario of process modelling language acceptance to showcase and exemplify this process in the development of a scales inventory designed to measure the various perceptions that an individual may have towards process modelling artefacts, in our case, process modelling languages.

We proceed as follows. The next section introduces theoretical foundations relevant to our research context. Then we report in detail on scales inventory creation and measurement development. This paper concludes with a summary of contributions and an outlook to future research.

2 THEORETICAL FOUNDATIONS AND RELEVANT CONCEPTS

As mentioned above, the starting point for any empirical investigation should be a thorough theoretical basis (Doll & Torkzadeh, 1991). Reference to an existing established theoretical foundation should be given preference in order to demonstrate validity and theoretical rigour. In our research, which we use in this paper for exemplification purposes, we seek to identify and investigate the success of phenomena associated with process modelling. In particular we are interested in the question of user acceptance of process modelling languages. Process modelling can be subsumed under the notion of conceptual modelling, *i.e.*, the process of building a representation of selected phenomena in the problem domain for the purpose of understanding and communication among stakeholders (Kung & Sølvsberg, 1986; Mylopoulos, 1992; Siau, 2004). Studies on the success of process modelling are rare (Bandara *et al.*, 2005), which is surprising given that studies on the success of Information Systems (IS) artefacts, under which such research can be subsumed, belong to the core research directions in IS (Lee *et al.*, 2003). In this context it is often noted that it is foremost the question of the acceptance, and not so much potential superior capabilities, of an IS artefact that determines the realization of its benefits (Davis, 1989).

Of all models that have been proposed to explain the acceptance of IS artefacts, the Technology Acceptance Model (TAM) (Davis, 1989; Davis *et al.*, 1989) has been most influential. Many TAM studies have been published over the years, leading to the statement that TAM denotes one of the few theories unique to the IS discipline that have not only obtained wide-spread acceptance in the field but also considerably high levels of maturity and rigor (Lee *et al.*, 2003). The main advantages of TAM are the parsimony and explanatory power of the model (Venkatesh & Davis, 2000) and the well-researched and validated measurement inventory with high levels of reliability and validity of constructs and measurement scales (Davis, 1989; Segars & Grover, 1993). King and He (2006) further found in their meta-analysis of TAM that, despite of its recent extensions, for example, the TAM2 model (Venkatesh & Davis, 2000), and revisions, for example, the UTAUT model (Venkatesh *et al.*, 2003), primarily the classical model is of high reliability and explanatory power and obtains high levels of robustness. As such, we deem TAM in its original form a suitable starting point for our line

of investigation and see potential and first evidence that it could successfully be applied to the study of the acceptance of process modelling languages, an area to which it has so far not at all been applied.

TAM was initially developed by Davis (1989) to explain and predict voluntary usage of computer systems. TAM assumes that an individual's acceptance of an information system, measured by the intention to use (ItU) the information system, is determined by the two major variables Perceived Usefulness (PU) and Perceived Ease of Use (PEOU). Over the past twenty years TAM has been applied to different IS artefacts (*e.g.*, email, GSS), under different situations (*e.g.*, culture, over time), with different moderating factors (*e.g.*, gender, organizational size), and with different subjects (*e.g.*, students, knowledge workers, managers). We do not wish to recapitulate each TAM study here and instead refer to the annotated overview given, for instance, in (Lee *et al.*, 2003). All these studies have, in essence, shown that the general postulates of TAM hold in a variety of settings, which suggests that TAM is also applicable to the domain of process modelling languages. More specifically, TAM studies have found that the constructs of PU and PEOU directly influence ItU (Davis, 1989; Davis *et al.*, 1989; Moore & Benbasat, 1991). Also, PEOU was found to be a causal antecedent of PU (Venkatesh & Davis, 1996; Venkatesh & Davis, 2000). This means that, often, users tend to find an artefact more useful (and thus tend to intent to use it) when they perceive it as easy to use. However, not all studies found these relationships to be always statistically significant. In particular, while PU has consistently been found to impact the formation of intention to use, support for PEOU has been inconsistent and sometimes of less significance. An explanation for this is speculated to reside in the fact that prolonged exposure to an IS artefact remedies potential concerns about the ease of its use (Chau, 1996).

Following TAM we propose an a-priori model of process modelling language acceptance that hypothesizes the following concepts and relationships:

- the acceptance of a process modelling language can be predicted by an individual's intention to use or to continue to use the process modelling language (ItU).
- the intention to use a process modelling language is jointly determined by an individual's perception of its usefulness (PU) and its ease of use (PEOU).
- the perceived usefulness of a process modelling language is causally influenced by its perceived ease of use.

Note that we slightly changed the definition of ItU to also include scenarios in which users have already been confronted with a process modelling language and hence make a decision to *continue* using it. We see a need for altering the construct in the fact that the initial adoption of a process modelling language is often an organizational decision. Ultimately, however, individual modellers are the ones who use a language and evaluate its acceptability (Ambler, 2004).

Fichman (1992) points out that most studies of IS acceptance are restricted to applying a general model that does not take into account the specific characteristics of the research context. The generality of TAM, which allows for wide applicability, induces a lack of focus on the particular artefact under observation. Hence, while TAM provides the advantage of a rich cumulative tradition, researchers seeking to borrow this theory must take care to ensure that its concepts and variables are being tailored to the specific research context (Segars & Grover, 1993). In particular, the original construct definitions of TAM should be adopted to the context of process modelling. Accordingly, Table 1 gives the original and adopted construct definitions as used in this study. Note here that the adapted definition of PU deviates from the original definition. Moody (2003) argues that the original definition of PU should be extended to reflect the objectives of the particular task for which the artefact is being used. Adopting this insight to the context of process modelling, the definition given in Table 1 reflects the notion of rational selection (Rescher, 1973), which states that, generally, those methods or tools (here: languages) will be adopted that outperform others in achieving intended objectives, *viz.*, which are more effective. Thus, PU represents a perceptual judgment of an artefact's effectiveness (Rescher, 1973). This was deemed to be of particular relevance to process modelling

given the wide range of purposes for which process modelling is being used, and hence the definition was slightly modified.

Construct	Original definition	Adopted definition for study
PU	The degree to which a person believes that using a particular system would enhance his or her job performance (Davis, 1989)	The degree to which a person believes that a particular process modelling language will be effective in achieving the intended modelling objective
PEOU	The degree to which a person believes that using a particular system would be free of effort (Davis, 1989)	The degree to which a person believes that using a particular process modelling language would be free of effort
ItU	The extent to which a person intends to use a particular system (Davis, 1989)	The extent to which a person intends to use or continue to use a particular process modelling language for process modelling tasks

Table 1: Construct definitions

Forthcoming from the specification of the construct definitions is the need to pursue appropriate scales inventories for these constructs. The next section discusses the development and pre-test of new multi-item measurement scales for these constructs.

3 MEASUREMENT CONSTRUCTION AND SCALE DEVELOPMENT

So far, no comprehensive scales inventory exists to measure the variety of perceptions that individuals may have towards phenomena associated with process modelling. Such an instrument, if valid and reliable, would however be vital to empirical studies on process modelling adoption, acceptance, diffusion and success.

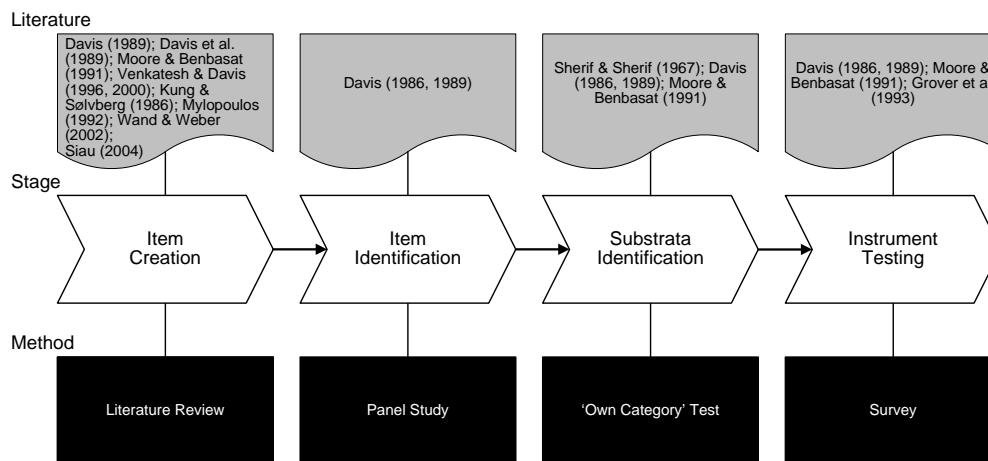


Figure 1. Scales inventory development procedure

In general terms, development of scales inventories should be carried out in multiple stages and should incorporate validation attempts during creation stages rather than ex-post during application stages. In our case, we followed the methodological procedures firstly prescribed by Davis (1989) and later extended and revised by Moore and Benbasat (1991). We selected these methodological procedures over others, e.g., (de Vaus, 2001), as they explicitly pay attention to the objective that the resulting inventories, while being developed for a particular research purpose, are general enough to allow for a wider uptake in other related studies (Moore & Benbasat, 1991, p. 194). The first stage is *item creation*, whose purpose is to create pools of candidate items for each relevant concept of the process

modelling acceptance model. The next stage is *scale development*, whose purpose is to sort the candidate items into meaningful separate concept categories to display construct, convergent and discriminant validity (Moore & Benbasat, 1991). This stage consists of two steps, item identification and substrata identification. The third stage is *instrument testing*, whose purpose is to identify from the pool of candidate items a set of reliable and valid items to be used in later empirical studies. A small empirical experiment was used to get an initial indication of the scales' reliability. Figure 1 gives an overview of the overall procedure and also displays the relevant literature. The following subsections describe each of stages in more detail.

3.1 Item Creation

The objective of the item creation step is to ensure *content validity* of the measurement items, defined as “the degree to which the scope or scale being used represents the concept about which generalizations are to be made” (Bohrnstedt, 1970 p 91). As the validity of scales inventories is built in from the outset (Davis, 1989), it should be established in the construction of the items through appropriate planning and rigorous procedures (Nunnally & Bernstein, 1994) rather than through *ex post* testing. Thus, items should be prepared to fit the content domains of the construct definitions (Anastasi, 1986).

Following these recommendations, candidate items for each of the three introduced constructs (PU, PEOU and ItU) were generated from past literature. It was referred to the Spearman-Brown Prophecy formula used in Davis' (1989) original study as an indication of how many items to create. Davis (1989) reports that at least ten items per construct would be needed to achieve reliability of at least .8.

As per specification of the candidate items, Ajzen and Fishbein's (1980) suggestions were followed to include into the definition of the items the *actual behaviour* (e.g., using a process modelling language), the *target* at which the behaviour is directed (e.g., BPMN as the process modelling language under observation), the *context* in which the behaviour occurs (e.g., for process modelling tasks) and, where possible, a *time frame* (e.g., current and most recent process modelling initiatives). The latter element was not explicitly included in the definitions as the general instructions of the test advised the participants to refer in their responses to the most recent process modelling initiative they have actively been part of. Note here that, in accordance to (Ajzen & Fishbein, 1980), we included in the item specification a specific modelling language, BPMN (BPMP.org & OMG, 2006), which will be the target language in the final field test, in order to make the item more tangible and understandable.

In preparing the candidate items, it was referred to two types of literature. First, previous studies on IS acceptance were reviewed to identify the set of candidate items that previous acceptance studies have shown to obtain highest levels of validity and reliability. Second, we reviewed conceptual modelling literature in order to derive candidate items from relevant concept definitions in the conceptual modelling domain. This was done to appropriately reflect the particularities of our research context and to ensure that all dimensions and domain substrata of the respective construct definition were covered. In Wand and Weber's (2002) research framework on conceptual modelling four primary purposes of conceptual modelling are outlined, *viz.*, supporting communication between developers and users, helping analysts understand a domain, providing input to the design process, and documenting the original requirements for future references. We argue that the multiplicity of purposes for which conceptual modelling can be used must be reflected in the measurement items to ensure appropriate content validity across all potential dimension substrata of the construct. As an example, Table 2 gives the initial item pool for perceived usefulness of a process modelling language. Due to word limitations we omit an in-depth discussion of the scale development procedures and results for all constructs and instead report on illustrative examples.

No	Item definition	
PU1	I find BPMN to provide an effective solution to the problem of representing business processes	(Moody, 2003)

PU2	I find BPMN useful for process modelling	(Davis, 1989; Davis et al., 1989; Venkatesh & Davis, 1996; Venkatesh & Davis, 2000; Moody, 2003)
PU3	I find BPMN useful for the task of designing process models for the purpose of supporting communication between stakeholders	(Kung & Sølvsberg, 1986; Mylopoulos, 1992; Wand & Weber, 2002; Siau, 2004)
PU4	I find BPMN useful for the task of designing process models for the purpose of helping domain understanding	(Kung & Sølvsberg, 1986; Mylopoulos, 1992; Wand & Weber, 2002; Siau, 2004)
PU5	I find BPMN useful for the task of designing process models for the purpose of providing input to systems design	(Kung & Sølvsberg, 1986; Mylopoulos, 1992; Wand & Weber, 2002; Siau, 2004)
PU6	I find BPMN useful for the task of designing process models for the purpose of documenting requirements	(Kung & Sølvsberg, 1986; Mylopoulos, 1992; Wand & Weber, 2002; Siau, 2004)
PU7	I find that using BPMN enables me to accomplish my process modelling task more quickly	(Davis, 1989; Moore & Benbasat, 1991)
PU8	I find that using BPMN for process modelling improves the quality of my process modelling work	(Moore & Benbasat, 1991)
PU9	I find that using BPMN improves my process modelling performance	(Davis, 1989; Davis et al., 1989; Moore & Benbasat, 1991; Venkatesh & Davis, 1996; Venkatesh & Davis, 2000)
PU10	I find that using BPMN increases my process modelling effectiveness	(Davis, 1989; Davis et al., 1989; Moore & Benbasat, 1991; Venkatesh & Davis, 1996; Venkatesh & Davis, 2000)
PU11	I find that using BPMN increases my process modelling productivity	(Davis, 1989; Davis et al., 1989; Venkatesh & Davis, 1996; Venkatesh & Davis, 2000)
PU12	I find that using BPMN makes it easier for me to do process modelling	(Davis, 1989; Moore & Benbasat, 1991)
PU13	I find using BPMN to be advantageous for process modelling	(Moore & Benbasat, 1991)

Table 2: Initial candidate items for perceived usefulness

Item creation for perceived ease of use and intention to use was accomplished in similar fashion.¹

3.2 Scale development

The goals of this stage were two-fold: to formally establish *construct validity* of the candidate items and to assess *construct validity* in terms of *convergent* and *discriminant validity* of these items. In order to achieve the former objective, a panel of process modelling practitioners was asked to assess on a 7-point scale the correspondence between the candidate items and the definitions of the constructs they are intended to measure, following the procedures prescribed by Davis (1989) and extended by Moore and Benbasat (1991). In order to achieve the latter objective, the panel was asked to sort the items into construct categories so that the statements within a category were most similar in meaning to each other and most dissimilar in meaning from those in other categories. The categories were also to be labelled. This step followed the guidelines of the ‘own category’ procedure by Sherif and Sherif (1967). Categorization provides a simple yet powerful indicant of cluster similarity that helps to reflect

¹ Item pools and analysis results can be obtained from the author upon request.

on the domain substrata for each construct and thus to assess coverage and representativeness of the items.

The panel consisted of sixteen members with various yet strong backgrounds in process modelling, including academic staff conducting research in the area of Business Process Management, BPM-affiliated consultants and business analysts, and a number of under- and postgraduate students with both class and practical experience in process modelling-related contexts. By including members with different theoretical and practical expertise it was sought to incorporate adequate proxies for varying types of process modelling practitioners. The scale development procedure was conducted in several steps. First, four panel members were in-face-to-face interviews asked to perform the three tasks ranking, categorization, and labelling. These initial respondents were also to report on given instructions and testing procedures, which were previously pre-tested with a separate panel member to ensure comprehensiveness and comprehensibility. Based on responses received the testing procedure and the instructions were revised before handed out electronically to the remaining eleven panel members. Each test contained an example case of a trial test related to various aspects of an automobile, which were to be ranked, categorized and labelled. This was done to ensure the mechanics of the scale development procedures were fully understood by the participating panel members.

For the ranking task, the responses of the panel members were averaged and then ranked to obtain an order of candidate items with respect to their content validity and to identify potential candidates for elimination. In eliminating items, however, it must be considered whether the remaining item pool contains appropriate representativeness of potential domain substrata of the theoretical construct (Bohrnstedt, 1970). Hence the categorization task was performed in order to identify, aside from content validity, items that do not display sufficient discriminant and convergent validity, *viz.*, to identify domain substrata that the item pool has excessive, or not enough, coverage of. For the categorization task, panel members were asked to place the candidate items in up to five categories so that the statements within a category are similar in meaning to each other and dissimilar to statements in the other categories. Following Davis (1989) the similarity data was cluster analysed by assigning to the same cluster items that at least seven members (equalling 44 %) placed in the same category. By comparing and reflecting on the chosen labels for the associated categories, the resulting clusters were given an appropriate label. In effect, the resulting clusters can be considered to adequately reflect distinct domain substrata for the considered construct and thus serve as a basis for identifying a set of items to comprehensively cover the domain content of the construct. In performing the clustering of the categories obtained from the panels, two coders separately clustered the categories, then met to defend their clusters and created a second, joint draft.

Again, we here exemplarily report on the item pool for perceived usefulness.² The ranking task resulted in an order of construct validity of the candidate items that can be used to eliminate items that demonstrate low validity (*e.g.*, items PU1, PU7, PU9, PU11). The categorization task most notable resulted in a cluster that twelve (75 %) panel members identified from the pool of PU candidate items. This cluster, namely *relevance to modelling purpose*, in turn reflects a substratum for the PU construct that has to be covered by the measurement item pool. Furthermore, a number of categories obtained appear to indicate the existence of two clusters related to the *effectiveness and efficiency* of a language and the *general usefulness* of a language. However, both clusters failed to obtain the required overall support (38 %, respectively). Merging these two clusters to a new cluster, *overall usefulness*, resulted in support of 63 %. This in turn indicated that the notions of usefulness and effectiveness/efficiency are strongly related to each other and do not denote distinct substrata. In summation, the categorization task resulted in two supported substrata for the PU construct, *relevance to modelling purpose* and *overall usefulness*, both of which obtained considerably high support from the panel categorization. A second step was to assess whether panel members repeatedly placed the same candidate items in these cluster. Following the recommendations of Moore and Benbasat (1991), we demonstrate reliability of

² Results for the remaining item pools can be requested from the contact author.

the cluster scheme by assessing the percentage of items placed in the target cluster across all panel members, which in turn indicates the degree of inter-judge agreement. Also, the items that obtained high placement percentages across the panel show high potential for high construct validity and reliability. Similar to the identification of the overall clusters, we placed items in a cluster if at least seven panel members (equalling 44 %) categorized the item accordingly.

The ranking and categorization data for the PU construct are summarized in Table 3. In similar fashion, the ranking and categorization tasks for PEOU and ItU were conducted and resulted in ranked items that fall into three, two and two clusters, respectively. More precisely, similar to (Davis, 1989), it was found that the ease of use construct embraces the domain substrata *effort of using* (100 % support) and *effort of learning* (75 % support). In addition, the categorization task resulted in the identification of a third cluster, *effort of understanding* (50 % support), that refers to the ease with which users find the modelling process and the resulting model clear and understandable. This, in hindsight, seems only reasonable and logical in the context of process modelling. With regards to ItU, we identified the two domain substrata *preference* (81 % support) and *intention* (88 % support), with the latter referring to an individual's plan or intent to use a process modelling language and the former referring to the fact that an intention to use may also be affected by alternative available process modelling languages. In such a case the decision to start or continue using a language may involve a reflection or reasoning about the advantages or disadvantages of a given process modelling language in comparison to others.

Item #	Ranking average	Rank	Cluster	Placement ratio	New Item #
PU1	5	5	-		dropped
PU2	4.1875	6	Overall usefulness	44 %	nPU1
PU3	5.53125	1	Relevance to modelling purpose	69 %	merged: nPU3
PU4	5.53125	1	Relevance to modelling purpose	63 %	merged: nPU3
PU5	5.21875	4	Relevance to modelling purpose	75 %	merged: nPU4
PU6	5.46875	3	Relevance to modelling purpose	63 %	merged: nPU4
PU7	3.4375	10	-		dropped
PU8	4	7	Overall usefulness	44 %	nPU2
PU9	3.4375	10	Overall usefulness	44 %	dropped
PU10	3.5	9	-		dropped
PU11	3.3125	12	-		dropped
PU12	3.3125	12	Overall usefulness	44 %	dropped
PU13	3.6875	8	Overall usefulness	44 %	dropped

Table 3: Pre-test results: perceived usefulness

The results obtained allow us to select from the initial item pool candidate items that show a high potential for validity and reliability. In terms of PU, for instance, items PU1, PU7, PU10 and PU11 were dropped because they failed to receive priority rankings and did not cluster with other items. As to the identified clusters, the two top-ranked items were selected for 'overall usefulness', *i.e.*, PU2 and PU8. The other items that fell into this cluster were dropped due to low priority rankings. As to the cluster 'relevance to modelling purpose' we had to consider that the initial item pool contained several items for several purposes (items PU3-PU6). Based on the responses obtained and the cluster identified as well as to pick up the content of all these items we decided to merge these items into two new items, "I find BPMN useful for the task of designing process models that serve my modelling

purpose” and “I find BPMN useful for the purpose of serving my modelling objective”. The creation of two items was done to be able to pick up different conceptions about the similarity or dissimilarity of the notions ‘modelling purpose’ and ‘modelling objective’. In summation, we were able to identify from our pool of thirteen candidate items four items, corresponding to two identified substrata of the PU construct, that appear to be suitable and promising candidates for measurement scales in an empirical instrument.

At this stage it has to be noted that the scale development procedure described here is more of a qualitative analysis than a rigorous statistical test of validity and reliability of the scales (Moore & Benbasat, 1991). Without full-scale tests of the complete scales inventory there is no way of establishing whether or not the scales in fact measure what they intend to measure. However, the procedures applied have been found very helpful in determining ‘good’ candidates for measurement scales that have the potential of obtaining acknowledged levels of reliability and validity. In addition, it allowed us to scale down the (quite extensive) list of candidates that can be obtained from reviewing related literature and thus to focus our work on the adoption of an existing theory (TAM) to a new domain (process modelling). Our findings already show that some of the items that previous TAM studies, *e.g.*, (Davis, 1989; Davis et al., 1989; Moore & Benbasat, 1991; Venkatesh & Davis, 1996; Venkatesh & Davis, 2000), found to be very explanatory and useful, *e.g.*, PU10 and PU11, appear not to be of required adequateness to the domain of process modelling. This in turn provides some empirical evidence in support of the arguments in (Fichman, 1992; Segars & Grover, 1993) that ‘blind’ adoptions of measurement inventories to research domains other than the original can lead to biased results that do not have appropriate levels of adequateness to the particularities of the domain.

3.3 Instrument testing

The next stage in this research is to conduct a pilot test of the scales inventory developed with a small convenient sample of process modellers. The objective is to ensure that the mechanics of the compiling the questionnaire had been adequate and to obtain further initial indications for scale reliability and validity. In this test, not only the most promising candidate items will be included but also further candidates that, forthcoming from the scale development procedures, seem to be candidates for elimination, *e.g.*, PU1, PU13. This is done to validate and test the initial findings obtained from scale development and to ensure that we did not accidentally drop ‘good’ items. The pilot test will result in a first formal reliability assessment that can then be used to finally cull ‘poor’ candidate items. Forthcoming from the revisions stemming from the pilot test will be the final field test that will then be used to obtain conclusive empirical proof for the validity and reliability of the developed scales. The final test will be performed by means of a survey, which is traditionally a typical method for testing scale development efforts in IS (Grover et al., 1993).

4 CONCLUSIONS

The instrument development research outlined in this paper provides several contributions. First and foremost this paper reported on the process of rigorously creating an overall scales inventory to measure the perceptions of individuals towards process modelling artefacts and associated phenomena. The procedure described ensures high levels of confidence in construct and content validity of the scales. The method employed in this research has been found both helpful and rigorous and should motivate researchers to adopt this design in related empirical studies. The results obtained, *i.e.*, the resulting scales inventory, can now be used in various studies to investigate how perceptions affect individual’s behavioural intention and affection towards process modelling.

Some words of caution ought to be said. The scales inventory development process described in this paper is not yet complete. Without testing of the overall questionnaire only initial indications of reliability and validity could be obtained. However, with this paper it was sought to follow the guideline of carefully documenting, and reporting on, every step in a research project. We also feel

that with the work to-date we have already contributed significantly by reporting on rigorous procedures of empirical study design. A second noted limitation is related to the fact that our instrument development drew heavily on existing frameworks and theories and thus potentially lacks other endogenous constructs that may pose relevance to the context of process modelling. Researchers working in this area thus have to carefully observe whether or not to consider such factors in addition to the ones discussed here.

References

- Ajzen, I. and Fishbein, M. (1980) *Understanding attitudes and predicting social behavior*. Prentice Hall, Englewood Cliffs, New Jersey.
- Ambler, S. W. (2004) *The object primer: Agile model-driven development with uml 2.0*. Cambridge University Press., Cambridge, England.
- Anastasi, A. (1986) Evolving concepts of test validation. *Annual Review of Psychology* 37 (1), 1-15.
- Bandara, W., Gable, G. G. and Rosemann, M. (2005) Factors and measures of business process modelling: Model building through a multiple case study. *European Journal of Information Systems* 14 (4), 347-360.
- Bohrstedt, G. W. (1970) Reliability and validity assessment in attitude measurement. In *Attitude measurement* (Summers, G. F., Ed), pp 80-99, Rand-McNally, Chicago, Illinois.
- Bpmi.Org and Omg (2006) *Business process modeling notation specification*. Final adopted specification. Object Management Group.
- Bubenko, J. A. (1986) Information systems methodologies - a research view. In *Information systems design methodologies: Improving the practice* (Olle, T. W. and Sol, H. G. and Verrijn-Stuart, A. A., Eds), pp 289-318, North-Holland, Amsterdam, The Netherlands.
- Burton-Jones, A. and Straub, D. W. (2006) Reconceptualizing system usage: An approach and empirical test. *Information Systems Research* 17 (3), 228-246.
- Chau, P. Y. K. (1996) An empirical assessment of a modified technology acceptance model. *Journal of Management Information Systems* 13 (2), 185-204.
- Curtis, B., Kellner, M. I. and Over, J. (1992) Process modeling. *Communications of the ACM* 35 (9), 75-90.
- Davenport, T. H. (1993) *Process innovation: Reengineering work through information technology*. Harvard Business School Press, Boston, MA.
- Davies, I., Green, P., Rosemann, M., Indulska, M. and Gallo, S. (2006) How do practitioners use conceptual modeling in practice? *Data & Knowledge Engineering* 58 (3), 358-380.
- Davis, F. D. (1989) Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly* 13 (3), 319-340.
- Davis, F. D., Bagozzi, R. P. and Warshaw, P. R. (1989) User acceptance of computer technology: A comparison of two theoretical models. *Management Science* 35 (8), 982-1003.
- De Vaus, D. A. (2001) *Surveys in social research*. Taylor & Francis Books, London, England.
- Dickson, G. W., Leitheser, R. L., Wetherbe, J. C. and Nechis, M. (1984) Key information systems issues for the 1980's. *MIS Quarterly* 8 (3), 135-159.
- Doll, W. J. and Torkzadeh, G. (1991) The measurement of end-user computing satisfaction: Theoretical and methodological issues. *MIS Quarterly* 15 (1), 5-10.
- Doll, W. J. and Torkzadeh, G. (1998) Developing a multidimensional measure of system-use in an organizational context. *Information & Management* 33 (4), 171-185.
- Fichman, R. G. (1992) Information technology diffusion: A review of empirical research. In *13th International Conference on Information Systems* (De Gross, J. I. and Becker, J. D. and Elam, J. J., Eds), pp 195-206, Association for Information Systems, Dallas, Texas.
- Froehle, C. M. and Roth, A. V. (2004) New measurement scales for evaluating perceptions of the technology-mediated customer service experience. *Journal of Operations Management* 22 (1), 1-21.
- Grover, V., Lee, C. C. and Durand, D. (1993) Analyzing methodological rigor of mis survey research from 1980-1989. *Information & Management* 24 (6), 305-317.

- Karimi, J. (1988) Strategic planning for information systems: Requirements and information engineering methods. *Journal of Management Information Systems* 4 (4), 5-24.
- King, W. R. and He, J. (2006) A meta-analysis of the technology acceptance model. *Information & Management* 43 (6), 740-755.
- Kung, C. H. and Sølvsberg, A. (1986) Activity modeling and behavior modeling of information systems. In *Information systems design methodologies: Improving the practice* (Olle, T. W. and Sol, H. G. and Verrijn-Stuart, A. A., Eds), pp 145-171, North-Holland, Amsterdam.
- Kwon, T. H. and Zmud, R. W. (1987) Unifying the fragmented models of information systems implementation. In *Critical issues in information systems research* (Boland, R. J. and Hirschheim, R., Eds), pp 227-251, John Wiley & Sons, New York, New York.
- Lee, A. S. (2001) Editor's comments. *Mis quarterly's* editorial policies and practices. *MIS Quarterly* 25 (01), iii-vii.
- Lee, Y., Kozar, K. A. and Larsen, K. R. T. (2003) The technology acceptance model: Past, present, and future. *Communications of the AIS* 12 (50), 752-780.
- Mcgrath, J. E. (1981) Dilemmatics: The study of research choices and dilemmas. *American Behavioral Scientist* 25 (2), 179-210.
- Moody, D. L. (2003) The method evaluation model: A theoretical model for validating information systems design methods. In *11th European Conference on Information Systems*, Naples, Italy.
- Moody, D. L. (2005) Theoretical and practical issues in evaluating the quality of conceptual models: Current state and future directions. *Data & Knowledge Engineering* 15 (3), 243-276.
- 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), 192-222.
- Mylopoulos, J. (1992) Conceptual modelling and telos. In *Conceptual modelling, databases, and case: An integrated view of information system development* (Loucopoulos, P. and Zicari, R., Eds), pp 49-68, John Wiley & Sons, New York, New York.
- Nunnally, J. C. and Bernstein, I. H. (1994) *Psychometric theory*. McGraw-Hill, New York, New York.
- Rescher, N. (1973) *The primacy of practice*. Basil Blackwell, Oxford, UK.
- Segars, A. H. and Grover, V. (1993) Re-examining perceived ease of use and usefulness: A confirmatory factor analysis. *MIS Quarterly* 17 (4), 517-525.
- Sherif, M. and Sherif, C. W. (1967) Attitudes as the individual's own categories: The social-judgment approach to attitude and attitude change. In *Attitude, ego-involvement and change* (Sherif, C. W. and Sherif, M., Eds), pp 105-139, Wiley, New York, New York.
- Siau, K. (2004) Informational and computational equivalence in comparing information modeling methods. *Journal of Database Management* 15 (1), 73-86.
- Straub, D. W. (1989) Validating instruments in mis research. *MIS Quarterly* 13 (2), 147-169.
- Venkatesh, V. and Davis, F. D. (1996) A model of the antecedents of perceived ease of use: Development and test. *Decision Sciences* 27 (3), 451-481.
- Venkatesh, V. and Davis, F. D. (2000) A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science* 46 (2), 186-204.
- Venkatesh, V., Morris, M. G., Davis, G. B. and Davis, F. D. (2003) User acceptance of information technology: Toward a unified view. *MIS Quarterly* 27 (3), 425-478.
- Wand, Y. and Weber, R. (2002) Research commentary: Information systems and conceptual modeling - a research agenda. *Information Systems Research* 13 (4), 363-376.
- Weber, R. (2003) Conceptual modelling and ontology: Possibilities and pitfalls. *Journal of Database Management* 14 (3), 1-20.
- Zmud, R. W. and Boynton, A. C. (1991) Survey measures and instruments in mis: Inventory and appraisal. In *The information systems research challenge: Survey research methods* (Kraemer, K. L., Ed), pp 149-180, Harvard Business School, Boston, Massachusetts.