

# HETEROGENEOUS USE FOR MULTIPLE PURPOSES: A POINT OF CONCERN TO IS USE MODELS' VALIDITY

*Completed Research Paper*

**Antti Salovaara**

Aalto University School of Business  
P.O. Box 21220, 00076 Aalto, Finland  
antti.salovaara@aalto.fi

**Anssi Öörni**

Department of Information Processing  
Science, University of Oulu  
P.O. Box 3000, 90014 University of  
Oulu, Finland  
anssi.oorni@oulu.fi

**Bertta Sokura**

Aalto University School of Business  
P.O. Box 21220, 00076 Aalto, Finland  
bertta.sokura@aalto.fi

## **Abstract**

*It is frequently observed in information systems (IS) research that technologies are used for multiple uses (i.e., heterogeneously). Quantitative IS use theories, including acceptance and task–technology fit models and their revisions, however do not take heterogeneity into account and address IS use on a general level. We present a study (N=116) on Microsoft Excel spreadsheet software's heterogeneous use. We demonstrate that Excel does not have one generally agreed-upon purpose of use but instead several important uses, each user exhibiting a different interpretation and pattern of use. Perceived usefulness (PU) and extent of use (U) receive different values depending on the purpose of use that is addressed, leading to PU–U correlations ranging between -.39 and .45 in the extremes. We find that the ambiguity on what PU and U measure is a serious threat to the existing IS use theories' validity. We discuss the implications and suggest corrective actions.*

**Keywords:** IS models, IS theory, Critical perspective, Heterogeneous use

This is an author copy of the paper to be published in Proceedings of the Thirty Fourth International Conference on Information Systems (ICIS 2013), F. Pennarola, J. Becker, R. Baskerville, and M. Chau (eds.), to be presented in Milan, Italy, 15–18 December 2013.
---

## Introduction

Theories and models of information system (IS) use comprise an important track of research and a highly regarded domain in our field (e.g., Burton-Jones, 2005; Straub and del Giudice, 2012). Inside this domain, predictive (quantitative) theories at the level of the individual user have been studied most actively. Among the prominent models in this tradition are the technology acceptance model (TAM; Davis, 1989; Davis et al., 1989), and its more comprehensive revisions (e.g., TAM2, UTAUT, UTAUT2; Venkatesh and Davis, 2000; Venkatesh et al., 2003; Venkatesh et al., 2012), the task–technology fit model (TTF; Goodhue and Thompson, 1995), and post-adoption (i.e., continued use) models (e.g., Bhattacharjee, 2001; Karahanna et al., 1999; Limayem et al., 2007).

Some of these studies have produced astonishing findings. For example, in studies on voluntary IS acceptance, most TAM variants explain 37–76% of the variance in the acceptance intentions (Venkatesh et al., 2003). UTAUT has been reported to explain 53% of the variance in actual use (ibid.). Such results have earned these models a wide recognition in the IS research domain. Habit-based models of continued use, in turn, have explained 26% of the variance in actual use of a voluntary IS (Limayem et al., 2007). Bearing in mind that these models make predictions on naturalistic human behavior that is known to be highly situated, these are remarkable levels of predictive power, especially for the most parsimonious models that only have 2–3 predictive variables. Due to the continuing success, TAM and UTAUT have been widely adopted as the prime means for explaining IT adoption. Habit-based models have rapidly acquired a similar recognition in continued IS use theories.

Despite the empirical breakthroughs, many researchers have expressed dissatisfaction with the current conceptual understanding of IS use construct. Orlikowski and Iacono (2001) assert that IS researchers have poorly reflected on the nature of the IS artifact they have studied, which has led to theoretical and conceptual simplifications. Burton-Jones (2005) observes that the studies presenting predictive IS use theories have rarely included any theoretical examination of the nature of IS use itself. Rather, IS use has been operationalized as a single quantitative variable, most often as an extent or frequency of use. The critique suggests that while the predictive theories may be statistically convincing, their predictions may not address the true nature of IS use, due to the naivety of their dependent variable. The need for theoretical reflection is exemplified by the definition of IS use offered by Burton-Jones and Straub (2006): IS use is an activity consisting of the three elements of (1) a user, (2) a system, and (3) the task. Hence, any model that attempts to capture instrumental use of IT and fails to explicate the task should be treated with caution. The problem of omitting the task context and focusing on too broad tasks was also one of the conclusions of Lee et al.'s (2007) review of TAM studies.

Conceptualizing IS use as a single measure of an amount or frequency is not fully supported by empirical observations. Several qualitative studies have shown that IS may be used in different ways, for example, to serve different purposes (DeSanctis and Poole 1994; Lassila and Brancheau, 1999; Majchrzak et al., 2000; Orlikowski, 1996; Robey and Sahay 1996). Thus, while the predictive models portray high levels of statistical power qualitative studies show that IS use is more varied than how it is captured in the aforementioned models. One may, therefore, justifiably ask how is it possible that high predictive power is consistently obtained when the dependent variable of the models omits the task for which the technology is being used (Salovaara and Tamminen, 2009). One way to answer to this question is to investigate empirically the *boundary conditions* for the validity of predictive theories. This is a different approach than what is common in IS research. Most TAM studies, for instance, have been *confirmatory* in a sense that their goal has been either to find increasingly better-fitting models or to extend them to new domains (e.g., from utility-oriented IS to so-called hedonistic IS; van der Heijden, 2004).

Studies that focus on boundary conditions have a different approach: their goal is to attempt to disprove the existing theories and find the conditions in which they do not hold. This approach borrows from Popper's (1963) *falsificationist* philosophy<sup>1</sup> of scientific progress. According to Popper, science advances

---

<sup>1</sup> For Popper, falsifiability is a central requirement of any scientific claim. That is, if there is no way to prove a claim to be untrue, it does not represent scientific theory. Popper also noted that the strongest theories are those that make surprising predictions (which should therefore be easy to falsify) but which nevertheless stand the repeated tests of falsification. Weakest theories, correspondingly, only present predictions that are obviously true.

through attempts to falsify (i.e., refute) existing theories. A falsification of a theory provides information on where the theory holds true and where it requires improvement. Although falsificationist research is not common in IS research, a few examples exist that address IS use models. Burton-Jones and Hubona (2005) evaluated the claim that individual differences (e.g., age, staff seniority, and educational level) would be fully mediated by perceived usefulness (*PU*) and perceived ease of use (*PEOU*) in TAMs, removing a need to include them as moderating variables. They found no support to this claim. Straub et al. (1995) studied the validity of self-reported IS usage measures by comparing the values with automatically logged measures. They found no correlation between these two measures, raising a question whether IS use models based on self-reported data are at least partially artifactual.

This paper follows the falsificationist approach. We will examine critically the implications of *heterogeneous use* (i.e., IS use for multiple different purposes) on IS use models. Our focus is on the possible ambiguities in *PU* and actual use (*U*) constructs with respect to the different purposes of IS use. More specifically, we focus on the errors that result when *PU* (or a related construct, such as performance expectancy; Venkatesh et al., 2003) and *U* are measured on a general level and not separately for each individual purpose of use.

Our findings suggest considerable problems with the validities of *PU* and *U* constructs resulting from unacknowledged heterogeneous use. Given the central role of *PU* in many IS use models, the falsifying evidence of our findings suggests severely delimiting boundary conditions for all IS use models that contain these constructs (especially TAM and UTAUT variants). Also other frameworks, such as TTF and system feature based approaches (e.g., Al-Natour and Benbasat, 2009; Barki et al., 2007; Griffith, 1999; Jaspersion et al., 2005) may need to be modified.

## Heterogeneous Use

We define heterogeneous use as *use of an IS for multiple purposes*. An underlying idea here is that the IS is used as a means, or a tool, to reach a certain goal or to serve some end. System's use for multiple purposes is one of the aspects that Burton-Jones and Gallivan (2007) list in their description of "configural use" (p. 667). In our definition of heterogeneous use, we adopt the "tool view" in Orlikowski and Iacono's (2001) classification of different IS artifact conceptualizations.

A "purpose" in our definition can be any task that the user wants to accomplish. In spreadsheet program use, such purposes may be, for instance, maintenance and editing of large tables of data (e.g., directories of phone numbers and addresses) or calculation of financial figures. Successful adoption of a program for these purposes depends on the features that it provides. In this case, the table editing and calculation of finances both depend especially on spreadsheet program's basic feature that allows for tabularization of data. Digital camera use is another good example of a single feature being used for several purposes: for "typical" picture taking, of course, but very often also creatively for scanning, note-taking, and use as a mirror or a periscope, for instance (Salovaara et al., 2011). These examples demonstrate that the same feature may be in use in different kinds of tasks.

Features, in other words, do not fully determine the purposes for which they may be used. In our view, features only afford different uses and it is the users who determine their purposes. The same applies to feature sets: they may similarly be used for multiple purposes. By maintaining that features are used for multiple purposes, the heterogeneous use viewpoint is related to many feature-level IS use models (e.g., Al-Natour and Benbasat, 2009; Griffith, 1999; Jaspersion et al., 2005; Sun, 2012) that suggest that extensions to feature use are part of ordinary IS use.

Heterogeneous use may manifest itself both in interpersonal and intrapersonal dimensions. At interpersonal level, individuals differ in their ways of using the same IS. At intrapersonal level, an individual uses the system for multiple different purposes in different situations. Furthermore, an individual may have different purposes for the system at different times, depending on the context of IS use. For example, when a user is at work, she may use a spreadsheet program for calculating finances, but at home, maintaining a list of people attending to a fellow couple's wedding may be the primary use at a certain point of time.

Therefore, although we conceptualize heterogeneous use in a rather limited manner by considering IS primarily as a tool (Orlikowski and Iacono, 2001), this conceptualization shows how much variance there

may exist for use of any IS artifact. In the following sections, we will first review the empirical evidence for this kind of heterogeneous use (i.e., is it common that systems are used heterogeneously), and then will present our research hypotheses, motivated by a suggestion that the meanings of *PU* and *U* constructs will clearly depend on what purpose of use is considered at each time.

### **Empirical Evidence**

In the introduction we already mentioned that a number of studies have shown that technologies are used heterogeneously. Their use practices may undergo changes over long periods of time when users discover new uses or adapt to new circumstances in their work. Orlikowski's (1996) longitudinal study reports how the use of a Lotus Notes based customer call tracking system transformed dialectically in five stages during a 2-year observation period. The use practices changed through series of opportunistic changes made by technical service specialists as well as through monitoring and work management changes by their managers. At the end of the two-year period the system was used in considerably different ways than at the time of its adoption. Also Boudreau and Robey (2005), Majchrzak et al. (2000), Lassila and Brancheau (1999), and Tyre and Orlikowski (1994) have written about similar dynamic, situated change processes, basing their analysis on ethnographically gathered data. DeSanctis and Poole (1994) have suggested that on a more detailed level, workers discover adaptations and new uses through "appropriation moves" (p. 129) during their courses of action.

Workarounds are a subclass of heterogeneous use that is frequently mentioned in IS use studies. Boudreau and Robey (2005) mention them as one of the three ways in which workers may react to exogenous changes in their work practices. In their study on implementation of a new financial processing system, workers "tweaked" the system's free-text lines to convey additional information and entered vendors' records multiple times in the system to allow for a possibility for multiple addresses. Such workarounds may become quickly the new standard practice within the firm, because they help bypass problems that affect many workers experiencing the same organizational change. Also, even if the work is not being restructured, workarounds are a way to cope with anomalies in one's ordinary daily work (Gasser, 1986). Researchers have also noted that enacting workarounds and adapting IS use in other ways may be unfaithful to the system's spirit at the organizational level (e.g., Wilkin and Davern, 2012).

The most direct evidence for heterogeneous use comes from content analyses of computer-mediated communication in workplaces. Several papers have classified instant messaging (IM) communication, identifying functions such as quick questions and clarifications, coordination and scheduling work tasks, coordinating impromptu social meetings, and keeping in touch with friends and family (Isaacs et al., 2002; Nardi et al., 2000; Salovaara and Tuunainen, 2013). On the level of individual messages, IM may be used for communication of work (69%), availability (13%), greetings (7%), humor (5%), non-work topics (3%) and other issues (4%), for instance (Handel and Herbsleb, 2002). In the 1980s, similar studies were carried out about employees' email use (Mackay, 1988). Overall, in human-computer interaction and computer-supported cooperative work research – fields in which research is not limited to work-oriented IS use – studies on novel IS uses are rather common (e.g., blogging, micro-blogging, and using social networking sites; DiMicco et al., 2009; Nardi et al., 2004; Zhao and Rosson, 2009).

### **How Heterogeneous Use Has Been Acknowledged in IS Use Models**

While emergent uses of different ISs in different contexts have therefore been observed, attempts to quantify the amount of heterogeneous use at individual or higher level have been missing (see e.g., Burton-Jones and Gallivan, 2007, p. 667). Our previous study (Salovaara et al., 2011) was a step towards this direction. It examined individual users' tendency to invent eight creative uses (mirror, scanner, map, note-taking tool, periscope, memory storage, lamp, and instruction tool) for digital cameras. We found that almost half (48%) of the respondents ( $N = 2,379$ ) had at least tried out such uses or adopted them in more frequent use. However, our study did not address a clearly utility-oriented IS. To our knowledge, no attempts to quantify heterogeneous IS use – neither creative nor typical – has been presented before the present paper. The above-listed studies, however, provide evidence that heterogeneous IS use is an important aspect of IS use in general.

**Table 1. Operationalizations for *PU* and *PEOU* in Venkatesh et al. (2003, pp. 448 and 451)**

Perceived usefulness ( <i>PU</i> )	Perceived ease of use ( <i>PEOU</i> )
1. Using the system in my job would enable me to accomplish tasks more quickly.	1. Learning to operate the system would be easy for me.
2. Using the system would improve my job performance.	2. I would find it easy to get the system to do what I want it to do.
3. Using the system in my job would increase my productivity.	3. My interaction with the system would be clear and understandable.
4. Using the system would enhance my effectiveness on the job.	4. I would find the system to be flexible to interact with.
5. Using the system would make it easier to do my job.	5. It would be easy for me to become skillful at using the system.
6. I would find the system useful in my job.	6. I would find the system easy to use.

Feature-based models have been already mentioned as a set of approaches that are related to heterogeneous use. These models focus on feature-level use, but do not attempt to predict the extent to which feature's (or feature sets') purposes of use are extended. Instead, they address the interaction-borne changes in user's and system's relationship (Al-Natour and Benbasat, 2009), the organizational impacts of post-adoptive individual IS use (Jasperson et al., 2005), or IS use at a general level (Barki et al., 2007). These models do not, however, incorporate heterogeneity as a predictor for such IS use outcomes.

Infusion models (e.g., Fadel 2012; Saga and Zmud, 1994; Wang and Hsieh, 2006), in turn, are focused on the problem of users failing to use ISs to their fullest potential. Users may adopt only a small subset of features into their use, which leads to suboptimal performance outcomes on an organizational level. While these studies also consider IS use at a feature level, they do not address the different purposes of use; their main interest is instead on the extent that different features are in use.

By far the most significant attempt to improve sensitivity to heterogeneity has been the development of the task-technology fit model (TTF; Goodhue and Thompson, 1995). TTF focuses on users' tasks and how the technology supports them. However, while the model itself is sensitive to different uses, its operationalization considers each technology as being used for one purpose only (ibid, p. 222).

As a result, most IS use theories have been formulated at a general level that does not specify what users do with the studied technologies. This applies especially to TAMs and their later revisions (e.g., UTAUT), but also to some IS continuance models (Bhattacharjee, 2011; Karahanna et al., 1999; Limayem et al., 2007) in which *PU* plays a role of an important predictive construct of use. As an example, Table 1 presents the statements used in UTAUT (Venkatesh et al., 2003) to measure *PU* and *PEOU*. As can be noticed, the statements do not mention the purposes for which the system in question is being used. By doing this, UTAUT effectively disregards the possible implications resulting from multiple purposes of use.

Such a disregard can be a problem. In a mixed-methods acceptance study, Wu (2012) noticed that interviews revealed that the respondents understood the concepts of usefulness and ease of use in different ways, depending on their personal interpretive frames. In another previous work of ours (Salovaara and Tamminen, 2009) we argued that without acknowledging heterogeneous use, the meaning of the questionnaire responses are likely to suffer from ambiguity. When respondents fill in questionnaires, they orient to the statements about *PU*, *PEOU* and other factors with different mindsets, yielding incommensurable answers. In sum, although heterogeneous use has been noted in several previous studies, its effect to the validity of predictive IS use models has not been analyzed. We present three hypotheses to start this examination.

### **Hypothesis 1: Prevalence of Heterogeneity**

As explained above, qualitative empirical studies have reported heterogeneous use for a variety ISs and in different contexts. Yet, due to their interpretive and qualitative nature, the question remains – how frequent is heterogeneous use at an individual level? We initiate our study with a measurement of this

frequency. This is a necessary first step in assessing the effect of heterogeneous use on IS use models' validity. Should we find that, at best, a modest share of individuals have more than one use for ISs, absence of task-specificity would probably not be detrimental to the validity of IS use models and theories. We therefore state the first hypothesis as follows:

H1: Information systems are used for multiple different purposes by most users.

We do not expect that all ISs would be used heterogeneously. For instance, automatic teller machines (ATMs) can be rarely used for anything else than drawing cash, and in some cases, topping up money on prepaid mobile phone cards or other pay cards (e.g., ones used in public transportation). Using ATMs for any other purposes is prevented by the system's design. Open-ended knowledge work tools, on the other hand, may be in the other extreme of the continuum of heterogeneous use. Because knowledge work often involves problem solving in tasks for which no predefined solutions are available (e.g., Blackler, 1995) the tools may be frequently adapted and appropriated for multiple purposes. In the empirical part of this paper, we will evaluate the hypothesis with respect to a spreadsheet software (Microsoft Excel) use as a starting point for exploration of this continuum.

## Hypothesis 2: Heterogeneity in Interpretations of Primary Use

An implication from H1, given that empirical data support it, is that when IS use models are applied, a mixture of different uses is being measured. The dominant IS use models do not currently specify the purposes of use that the models are intended to predict. Instead, concerning the purposes of use, the constructs are defined on an unspecified and general level. This was already noticed in the discussion of UTAUT's operationalization of *PU* and *PEOU* constructs.

Our second hypothesis is a methodological implication of H1. When users participate in IS use studies and are inquired about their perceptions of usefulness, the heterogeneity of use (given that H1 is taken to be true) makes them interpret the *PU*s, *PEOU*s, *Us* and other constructs' items in different ways. Because the items are imprecise regarding the purpose of use, the responses obtained will represent a mix of different interpretations which may or may not be the same as those of the researcher. Respondents may provide their answers on the basis of at least four different interpretations:

1. *Aligned interpretations*: the user and the researcher will understand the questionnaire's items along the same interpretation of purpose of use. This is the situation implicitly assumed in practically all acceptance studies (for a contrast, see Wu, 2012).
2. *Misaligned interpretations*: Respondent interprets the questions from an unintended use point of view which, however, for this respondent is the system's primary purpose of use. Thus, respondent's and researcher's interpretations differ.
3. *Possibly misaligned interpretations*: Upon reading a questionnaire item, the respondent may realize that there are different ways to answer to the items. The respondent may decide to make a guess about the system's intended use without knowing what that might be. This use may be different from the respondent's own primary purpose of use. The guess can go right or wrong, leading to a possible misalignment of user's and researcher's interpretations.
4. *Weighted average*: Similarly with the third situation, the respondent realizes that the questions are ambiguous. Upon realizing this, the respondent makes an attempt to answer to the questions at a general level. The general assessment will be a kind of weighted average of the different purposes of use, some of which are representative of this respondent's own use practices and some that the user is only aware of but does not apply actively.

Interpretations 2–4 are methodologically problematic, because they introduce undesired and unknown noise and bias in the data. If the standard questionnaire is not amended with control questions, it is impossible to detect which of the four situations is realized in each respondent's case.

Regarding *U* the same problems can be noticed. Typically *U* is measured by using self-report questions (e.g., by asking how frequently the system is being used; Venkatesh et al., 2012) or directly as a behavioral variable (e.g., by studying system logs; Venkatesh et al., 2003). Both of these measures can be collected quite straightforwardly. However, if the purpose of use for each use event is not captured as a part of data collection, it will be impossible to assign them to different purposes at the analysis stage. We formulate

the confusion related to different interpretations as the following hypothesis:

H2: Users differ in their main purpose of use and their interpretations of the intended purpose of use.

Our goal with this hypothesis is to find evidence for the existence of the interpretations 2–4 in the data. Positive evidence gives a rise to our next research hypothesis.

### **Hypothesis 3: Severity of Incongruence of Interpretations**

Assuming that H1 and H2 are found to be empirically plausible, their implication is that models' antecedent constructs will represent a different purpose of use than the dependent variable. For instance, in a case of spreadsheet use, if the questionnaire items have not contained a specification of what purpose should be evaluated, the misaligned interpretations (see H2) will lead the respondent to answer differently to different parts of the questionnaire. We call this *incongruence of interpretations*.

Incongruent interpretations may be fatal to the validity of correlations and covariances between different constructs. In an incongruent situation, apples are compared to oranges: the pair of variables represents two different kinds of use, and their comparison will lead to meaningless results. We formulate our hypothesis as follows:

H3: The ambiguity in measured purpose of use introduces a severe measurement incongruence into those IS use models where *PU–U* relationship plays an important role.

With this hypothesis, we will test whether incongruent comparisons will lead to severe measurement errors. We have particularly in mind the situation in which a regression or structural equation model will have invalid beta coefficients due to the incongruence of measured variables.

## **Study**

We carried out a study on Microsoft Excel spreadsheet software (hereafter Excel) to investigate the three hypotheses. Our decision to focus on Excel was motivated by several reasons: i) Excel is a widely used knowledge work tool and researching its use therefore has relevance in many contexts; ii) Excel is rather open-ended in its use, meaning that heterogeneous use should be easily observable if such use exists; and iii) Excel is a productivity tool and therefore provides a good match with our interest in the *PU* construct.

With these choices, we wanted to focus on an IS that has a high likelihood to support our hypotheses. We reasoned that if hypotheses are not supported in Excel use, their support is unlikely in any other contexts either. Support for the hypotheses, on the other hand, would mean that heterogeneous use has relevance, and future research would be necessary in order to assess its impact also in other IS use contexts.

## **Method**

Our data sample consisted of business school students who participated in an introductory course on information and communication technology (ICT) skills. The 1.5-month course between September and November 2012 consisted of lectures and hands-on teaching of Microsoft Excel, Word and PowerPoint. This course is one of the first that new students take in the school. Our reason for studying first-year students was their initial homogeneity as a group: when they enter the school, they usually have very little experience of Excel. This makes them an easily comparable group of subjects.

When the course was in progress, we carried out a *preparatory data collection* stage using a paper-based questionnaire. This included collection of demographic information (age, gender, which year at our school) of our students ( $N = 211$ ), their initial Excel experience at the start of the course as well as the *PU* and *U* measures at the end of the course in four different ways: as a general purpose-unspecific assessment and specifically with respect to calculation in tables, making charts, and working with lists. We used UTAUT's six *PU* items (see Table 1) for measuring *PU*. *U* was operationalized as a frequency of use on a 5-point interval scale (daily, weekly, monthly, yearly, more rarely or never). As a gratuity for their effort, the students participated in a draw of six 20€ gift cards.

Table 2. The Six Specific Purposes of Use and the General Use Inquired in the Study	
Purpose	How the Purpose Was Explained to Respondents (Heading : Description of an example situation) <sup>a</sup>
Making charts	Making charts and visualizing information: Imagine that you have to <b>make a chart or an exemplifying graph</b> from a table of numbers.
Calculation in tables	Calculation in tables (e.g., averages and sums): Imagine that you have to carry out a calculation on a <b>tableful of numbers</b> , e.g., calculate their average.
Taking notes	Taking notes: Imagine that you are at work writing a <b>check list</b> (e.g., a shopping or to-do list).
Long tables	Information retrieval and sorting in long tables with more than 100 lines of numbers and text: Imagine that you are at work editing a <b>table with more than 100 lines</b> containing both numbers and text, Your task is to enter more lines, search for information and keep the contents sorted.
Complex calculations	Complex calculations: Imagine that you have to carry out a <b>complex calculation</b> , e.g., an investment calculation involving discounting, or fitting a trend curve to a table of statistical data.
Simple calculations	Simple calculations: Imagine that you end up in a situation in which you have to carry out a <b>very simple calculation</b> , e.g., $2300 * 1.23$ or some other calculation including only addition, subtraction, multiplication, division, or percentages,
General use	General questions: Think about situations in which you have <b>worked on numbers, tables or lists</b> lately.

<sup>a</sup> The words in boldface appeared in boldface also in the questionnaire.

The *primary data collection* took place 10 weeks after the course when the students had had more experience (e.g., coursework in different courses) on situations in which they might have a need for Excel. This questionnaire was administered online. We collected data on the following pieces of information:

- *U* and *PU*, inquired in seven different ways (at a general level, as in the previous round, followed by inquiries with respect to six specific uses; see Table 2). The increase in repetition created however a danger that the questionnaire would become excessively long and lead to a low return rate. We therefore shortened the *PU* scale down to four items by removing items 3 and 4 (see Table 1) that had high correlations with item 2 ( $r > .75$ ) and therefore were redundant.
- The use purpose that the respondent had in mind when responding to *PU* questions at the general level. The respondent could choose one of six specific purposes or provide a free-text answer.
- The respondent's interpretation of Excel's intended uses (selection of two options from a list of the six specific uses, added with an optional text field).

When we devised the list of Excel's possible uses we attempted to remain agnostic of Excel's possible "intended purpose of use". We started out by considering the affordances provided by the tool (e.g., that data can be entered into a tabular structure and that calculations can be performed on the data) and thinking about plausible ways of using them. Because of this we included also note-taking as one of the uses. Even if this may be a rare use for Excel, in fact, we had witnessed it in an earlier study on supermarket shopping (Nurmi et al., in press). We also made no attempt of being exhaustive in the list. For example, we decided not to include the option of using Excel as a calendar, even if this was one of the



uses suggested by Microsoft in its web pages on Excel's 2010 version<sup>2</sup>. Given our goal of proving heterogeneity, it was not necessary to be exhaustive; it sufficed to reach a level that would have implications for existing IS use models.

In the primary data collection stage, we unexpectedly deemed it necessary to change our method for measuring  $U$ . During the preparations for the primary data collection stage, we had already analyzed the data from the preparatory stage and had observed  $U$  having low values: in the median, Excel had been used only "less than monthly" among the students.  $U$ 's five-point interval scale proved too coarse for analysis of such an infrequent use. We hypothesized that maybe the users had not encountered enough situations in which Excel would be needed. Even if they would use Excel every time when they perceived a need for it, the  $U$  would stay low because of the scarcity of relevant events in general. If this were true,  $U$  would not be a sign of Excel's low acceptance, but in fact a product of two components: the frequency of situations in which Excel was relevant, and the probability of selecting Excel as a tool for solving those situations. A product like this reminded us of the way in which the calculation of conditional probabilities is carried out in mathematics:

$$p(A \text{ and } B) = p(A) \cdot p(B \text{ with a condition that } A),$$

In our case, the  $U$  of using Excel for drawing graphs, for instance, would be the frequency of situations in which one needs to draw graphs multiplied with the probability that the person chooses Excel as a tool for making charts. As a formula, this translates to

$$U(\text{must make charts and uses Excel}) = \text{freq}(\text{must make charts}) \cdot p(\text{uses Excel if must make charts}).$$

By using the conditional formula, we considered it possible to study also such uses of Excel that would occur in infrequent situations (i.e., with low *freq*) but where the IS has a high relevance (i.e., high  $p$ ) for the user. Should such a use be measured in the standard way,  $U$  would appear as unimportant due to the situation's rarity. Using the conditional formula, the  $p$ -part would signal about the system's importance also in such situations and be thus a more valid measure of Excel's importance for the given purpose.

Because we had obtained low  $U$  values in the preparatory data collection stage, we wanted to be prepared for low values also in the primary data collection. We decided to explore the measurement of  $U$  using also the conditional formula. Our assumption was that this approach would provide us with more granularity in the scale's low-frequency extreme where the 5-point Likert scale had proved too coarse. On the other hand, we also took a deliberate risk of added measurement bias that would result from subjective estimations that our respondents would provide for the formula's  $p$ -part (i.e., as a probability of choosing Excel). To our knowledge  $U$  had not been conceptualized and measured in this manner before, and we deemed this a good opportunity to examine its suitability, given the opportunity of potentially better analyzability of infrequent use.

We assigned the respondents randomly into two equal-sized groups and used the following alternative wordings for measuring  $U$  (using the chart-making as an example):

- Standard- $U$  condition: "How often have you used Excel for making charts during the last 7 days?" (a text field permitting only values between 0 and 100, decimals allowed)
- Conditional- $U$  condition:
  - The *freq*-part: "Overall, how often have you had a need for making charts during the last 7 days? The situations do not need to be related to using Excel as such. Please think about all situations like this." (a text field permitting only values between 0 and 100, decimals allowed)
  - The  $p$ -part: "How often in these situations have you used or you would use Excel?" (a text field permitting only values between 0 and 100, to designate percentage, decimals allowed)

As can be seen, we abandoned the interval scale in this new wording for  $U$ . In both of the conditions,  $U$  was now operationalized as a scalar (i.e., times of use during the last 7 days).

We administered all the questionnaires in Finnish, because the course was taught to Finnish-speaking students. We carried out the translations for  $PU$ 's items ourselves. The students participated in a draw of

<sup>2</sup> See <http://office.microsoft.com/en-us/excel-help/basic-tasks-in-excel-2010-HA101829993.aspx> (accessed 28 August 2013).

six 20€ gift cards in the preparatory data collection, and a draw of ten 20€ gift cards in the primary data collection.

## Results

A total of 304 students had signed up for the course. We carried out the preparatory data collection stage during lectures, gathering data from those students who were present. We obtained 190 responses at the start of the course and 137 responses at its end from altogether 200 different students, with data from 116 students who were present both at the beginning and the end of the course.

In the primary data collection stage, we approached all the 200 students who had already participated in the study in one way or another, and partitioned them randomly into two groups of 100 students each, which we invited for the primary data collection stage. We obtained 61 and 55 fully completed responses from the two conditions, yielding final  $N = 116$ . Table 3 presents the basic data of these two samples. We tested with  $t$ -tests and Mann-Whitney non-parametric  $U$  tests that the two samples did not have statistically significant differences with respect to any of Table 3's basic data variables. In the following subsections, we will present the results for the three hypotheses with Excel as our case system.

Table 3. Basic Data About the Two Samples		
	Condition	
	Standard- $U$	Conditional- $U$
$N$	61	55
Gender, % female	54%	59%
Age, avg. years ( <i>s.d.</i> )	22.6 (4.9)	21.6 (3.2)
Year in the university, year ( <i>s.d.</i> )	1.2 (0.76)	1.1 (0.35)
Excel use before ICT course, median frequency <sup>a</sup>	Less than monthly	Less than monthly
Excel use before ICT course, median years <sup>b</sup>	1–5	0–1

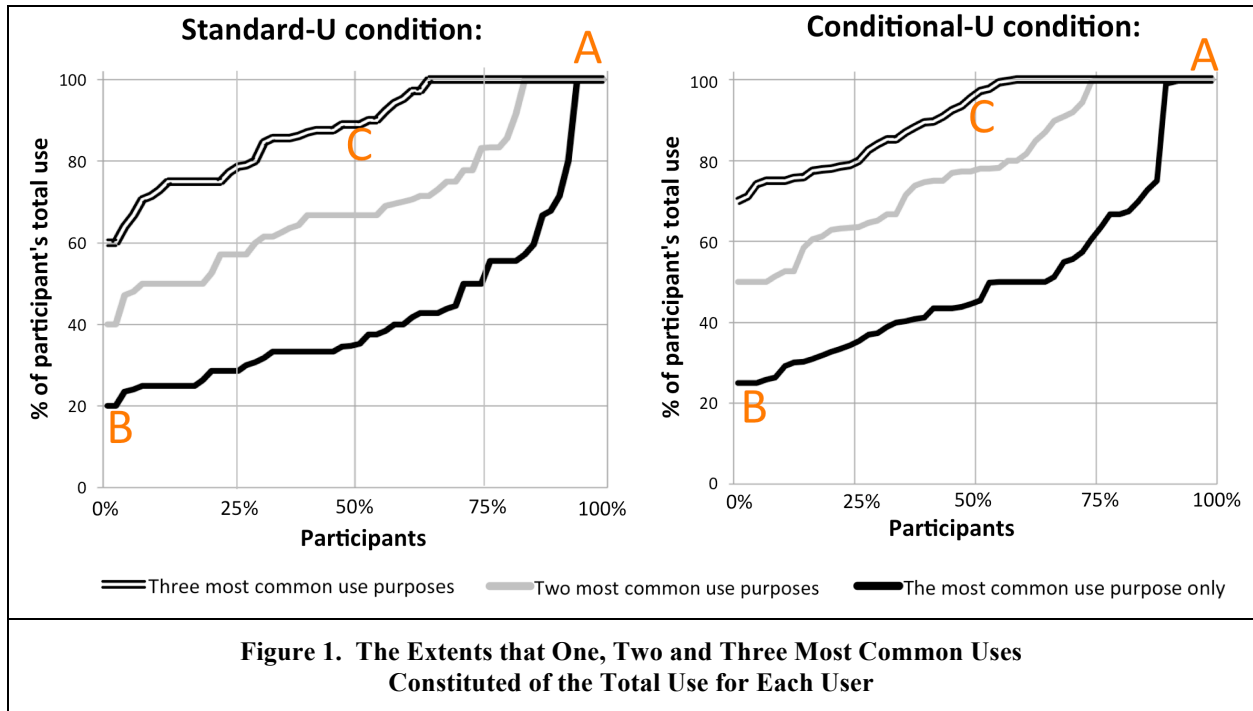
<sup>a</sup> Range for frequencies was: not at all, less than monthly, monthly, weekly, daily.

<sup>b</sup> Range for years was: no experience, 0–1 years, 1–5 years, over 5 years.

### H1: Prevalence of Heterogeneity

The first hypothesis stated that ISs are used for multiple different purposes by most users. To evaluate this hypothesis, we calculated the percentage that the most common use purpose, the two most common use purposes, and the three most common use purposes represented of each student's total Excel use. The total Excel use was calculated as a sum of the six separate purpose-specific  $U$ s.

The results are presented in Figure 1 separately for the two conditions, with the students ordered by the percentage of their most common use purpose out of the total use. Respondents reporting zero use for Excel altogether had to be removed from this analysis due to divisions by zero in the percentage calculations, yielding  $N = 110$ . With this plotting, the left extremes of the charts present the users whose most common purpose of use (whatever it is of the six alternatives) amounts to only a small percentage of their total use. In other words, the most heterogeneous users are presented in the left, and heterogeneity decreases as one moves towards the charts' right ends. In the Conditional- $U$  condition the amounts of use were calculated using the conditional formula.

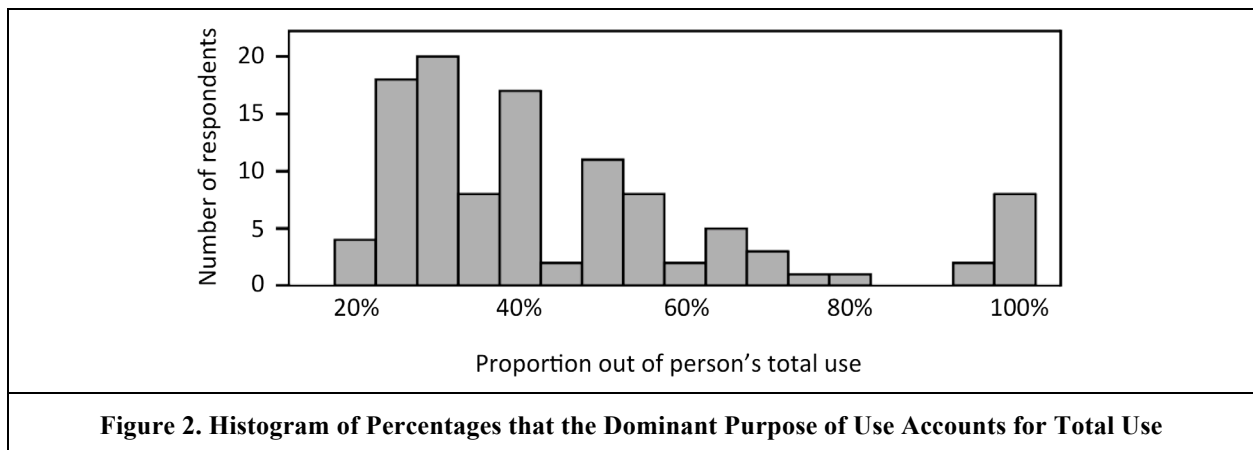


In both conditions, the following three observations can be made:

- In the most homogenous extreme (marked with A), as few as approximately 10% of all the users use Excel for only one purpose.
- In the most heterogeneous extreme (marked with B), Excel is used for its most common purpose only 20–25% of the time.
- The median user (at 50% mark, marked with C) uses Excel for more than 3 different purposes. The three most common purposes account for approximately 90% of the total use.

At face value, the graphs are a clear evidence of heterogeneous use. We are not aware of a test suitable of proving heterogeneity statistically in the kind of data that we have. Nevertheless, we carried out a set of analyses to obtain a better understanding on the extent of heterogeneous use among our respondents.

Figure 2 shows how the participants' most frequently applied (i.e., with highest  $U$ ; hereafter called "dominant") use purposes' percentages out of their total use were distributed. The distribution's mean and median are at 47% and 41% marks, respectively. If Excel use were homogeneous (i.e. the dominant use would represent a large share of total use for most participants), these measures should be close to



100% (e.g., 75%). For the purposes of exploration, ignoring the fact our data is not normally distributed (Shapiro-Wilk's  $W = .852$ ;  $p = .000$ ), we tested which distributions with mean values higher than our 47% would significantly differ from our data. Let us assume that a distribution of observations of an equal-sized sample ( $N = 110$ ) with a mean at 75% mark and with the same variance as our data would be taken as a proof for homogeneous use. A  $t$  test between our data and such a distribution clearly tells that the observed patterns of Excel use in our data cannot represent a population with prevalently homogeneous purposes of use ( $t(218) = 9.62$ ,  $p < .001$ ). The comparison remains significant until we relax our definition of homogeneous use as low as to a mean of 52.6% ( $t(218) = 1.95$ ,  $p = .05$ ).

Supported by both descriptive as well as explorative statistical evidence, we conclude that Excel was used overwhelmingly for multiple purposes. Hypothesis H1 holds true at least in the case of Excel use.

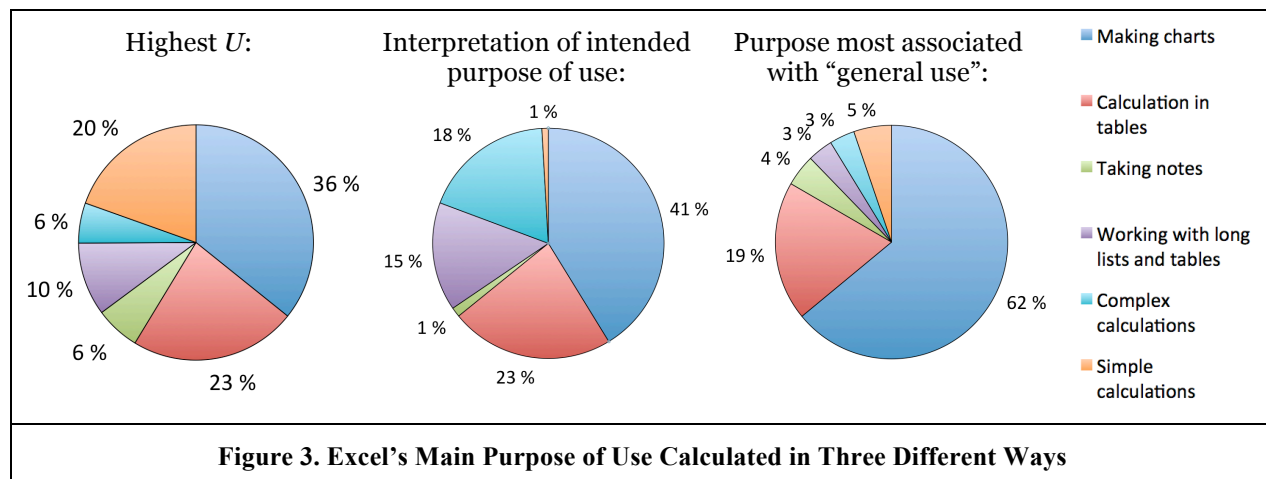
## H2: Heterogeneity in Interpretations of Primary Use

The second hypothesis stated that the main purpose of use would be different for different users and that the users' interpretations of the intended purpose of use would also be different. As described, we had collected information about Excel's main purpose of use in different ways in our survey. We therefore had three alternative metrics to define the main purpose of use:

- The dominant purpose of use: the purpose with the highest  $U$ .
- The interpretation of intended purpose of use: the two purposes that each respondent had stated as Excel's primary intended uses (with an omission of answering allowed).
- The general-level- $PU$  orientation: the interpretation that determined the responses to general-level  $PU$  items (i.e., the scale in which the purpose was not specified).

We identified what the main use purpose was for each respondent in these three measurement methods. Regarding the purpose with the highest  $U$ , we first analyzed the results separately for  $U$ 's both operationalizations (i.e., Standard- $U$  vs. Conditional- $U$ ) and compared them with a Chi-square test. Because no statistically significant difference was found ( $\chi^2(5) = 7.37$ ,  $p = .19$ ), we pooled the two conditions together. Regarding the interpretations of intended purpose of use, we had two votes from each respondent. We included both votes in our calculations.

We counted how many times each purpose of use was identified as the main purpose in each of the three methods. In the first and the third metric, one purpose of use was identified for each respondent. For the second metric, due to the varying number of answers given by each respondent (with two answers provided by 90% of the respondents),  $N = 225$  answers were included in the analysis. Figure 3 shows the results. Already by visual inspection one can see, for example, that when students answered to general-level  $PU$  items, 62% of them oriented to Excel's use for making charts. However, making charts was Excel's most used purpose only among 36% of the students when measured more behaviorally, as their dominant purposes of use. Therefore, what students used Excel for was different from what they took as its intended use.



However, observations like these do not directly prove the claims stated in hypothesis H2. As said, to refute H2 the proportions should show that across all the respondents, one purpose of use would have a considerably higher percentage than the others. Only in such a case it would be possible to state that a) users would *not* use Excel for multiple purpose and that b) the main purpose of use would also be the same for everyone. These two results would refute H2.

Looking at the charts in Figure 3 on sees that only chart-making would have a possibility of being considered as Excel's single main purpose of use. However, with the exception of general-level-*PU* orientation metric, its proportions were low. Intuitively, being able to call one purpose of use as a system's main purpose, its percentage should be at least above 50%. This was true only in one case out of three metrics, given the narrow error margins for the two other metrics. Assuming a normal distribution of error, the percentages for chart-making were  $36 \pm 9\%$  for the highest-*U* metric,  $41 \pm 6\%$  for the interpretation-of-intended-use metric, and  $62 \pm 9\%$  for the general-level-*PU* orientation metric.

Taken together, the results showed that the Excel's main purpose of use varies between users, both behaviorally as well as when considered as interpretations. The results also show that in our study, the participants provided their general-level *PU* and *U* ratings based on different grounds. Only in a study addressing *PU* alone, without any comparison to *U*, there would be a possibility to warrant a claim for uniformly agreed-on purpose for Excel's use. Therefore, hypothesis H2 was supported.

### H3: Severity of Incongruence of Interpretations

The third hypothesis stated that the ambiguity of what purpose of use is being measured and what the respondents think that the purpose is introduces a severe measurement incongruence into those IS use models where *PU*–*U* relationship plays an important role. Confirmation of H2 showed that this hypothesis was plausible, because users oriented to general-level *PU* inquiries in different ways, leading to incongruences between what the questions are intended to measure and what is being measured.

To evaluate the magnitude of incongruence-borne error, we compared the six purpose-specific *PU*s and the general-level *PU* with each other. Because each *PU* was represented with 4 items, we first took an average for these items, thus assigning a single scalar measure for each purpose and each user. We then calculated averages of these purpose-specific scalars over all the respondents and compared them with each other. Because *PU*s were measured identically in both of our *U* conditions, we pooled the respondents from both conditions together in this analysis. Table 4 presents the results.

Table 4. Differences of Mean <i>PU</i> for Each Pair of Purpose-Specific Uses							
	<i>PU</i>	$\Delta PU$					
	mean (s.d.)	Making charts	Calculation in tables	Taking notes	Long tables	Complex calculations	Simple calculations
Making charts	6.07 (1.44)						
Calculation in tables	6.14 (1.38)	-.07					
Taking notes	3.46 (2.11)	2.60***	2.67***				
Long tables	5.78 (1.70)	.29	.36	-2.31***			
Complex calculations	5.32 (1.79)	.74***	.81***	-1.86***	.45*		
Simple calculations	4.45 (1.93)	1.62***	1.69***	-.99***	1.33***	.87***	
General use	5.84 (1.09)	.23	.30	-2.37***	-.06	-.51**	-1.39***

\*  $\Delta PU$  is significant at the .05 level (2-tailed).

\*\*  $\Delta PU$  is significant at the .01 level (2-tailed).

\*\*\*  $\Delta PU$  is significant at the .005 level (2-tailed).

**Table 5. Correlations Between *PU*–*U* Pairs Where *PU* and *U* Represent Incongruent Purposes of Use**

Purpose of use measured in <i>PU</i>	Correlation ( <i>r</i> ) with different purpose-specific <i>Us</i>						
	Making charts	Calculation in tables	Taking notes	Long tables	Complex calculations	Simple calculations	Unspecified use
Making charts	.17	.15	-.07 <sup>ddd</sup>	.04	.17	.15 <sup>d</sup>	-.03 <sup>ddd</sup>
Calculation in tables	.16	.30	.02 <sup>ddd</sup>	.21	.19	.21	.18
Taking notes	-.15 / .24	.18	.39	.10	.09 <sup>d</sup>	.24	.10 <sup>d</sup>
Long tables	.10	.23	-.05 <sup>ddd</sup>	.24	-.04 <sup>ddd</sup>	.09 <sup>ddd</sup>	.11 <sup>d</sup>
Complex calculations	.13	.21	.01 <sup>dd</sup> / -.39 <sup>ddd</sup>	.21	.39	.22	.08 <sup>d</sup>
Simple calculations	.16	.14	.01 <sup>ddd</sup>	.23	.10 / .45	.44	.08 <sup>d</sup>
Unspecified use	.21	.37	.09 <sup>d</sup>	.24	.36	.27	.38

If correlations calculated with Standard-*U* and Conditional-*U* differed statistically significantly (by more than .35 in this data), both correlations are presented (Standard-*U* / Conditional-*U*). Otherwise a pooled correlation is given.

The shaded cells are cases in which *PU* and *U* have measured the same purpose of use (i.e., *PU* and *U* are congruent).

<sup>d</sup> Correlation differs significantly at the .05 level (2-tailed) from the congruent correlation in the same column.

<sup>dd</sup> Correlation differs significantly at the .01 level (2-tailed) from the congruent correlation in the same column.

<sup>ddd</sup> Correlation differs significantly at the .005 level (2-tailed) from the congruent correlation in the same column.

From the table, we can read that the largest difference between *PU*s was observed between table calculation and note-taking (2.67). Given that the scale range was 1...7 (i.e., scale width 6), this is a difference of 44.5% of the scale's width! If our study had been focused on Excel's use for calculation in tables and we had presented *PU* items with a general-level wording without specifying what we intend to measure, the respondents with a note-taking orientation would have provided responses erring by 44.5% in the *PU* variable. This source of error would have been impossible to notice without extraneous information about these users' different Excel use habits. Luckily for such a study, this incongruence would have concerned only 4% of the respondents (see the Figure 3, rightmost chart).

The example above is an extreme case; it is useful to consider also more ordinary scenarios. Thus, comparing only the three most common uses (as defined by the metric of highest *U*; see Figure 3) – making charts, calculation in tables, and simple calculations – the largest *PU* measurement error would result from incongruence between chart-making and simple calculations ( $\Delta PU = 1.62 = 27\%$  of the scale's width). Using *t*-tests, we found that most of the differences shown in Table 4 are also statistically highly significant.

For those IS use models in which *PU*–*U* correlation plays an important role (e.g., TAM and UTAUT), the possibility of incongruence is a threat to the models' validity. Incongruence will lead to incorrect *PU*–*U* correlations every time when *PU* and *U* represent different purposes of use. Table 5 shows the correlations between each pair of incongruent *PU*s and *Us*. The *congruent* (i.e., valid) correlations are presented in the diagonal, highlighted with a shaded background.

To generate the contents for Table 5, we first computed the pairwise *PU*–*U* correlations separately for the Standard-*U* and Conditional-*U* conditions. Using a procedure developed by Fisher (1921), we then tested for each cell whether the correlations obtained in different conditions were significantly different from each other. Table 5 shows the pooled correlations in all the cells except for those in which a difference was found.

The correlation typically used in IS use studies is the one in which the purpose of use is specified neither for *PU* nor *U*. This correlation (.38) can be found in Table 5's bottom-right corner. Our analysis of hypothesis H2 already showed (see the second chart in Figure 3) that different users have different purposes of use in mind when they answer to unspecific *PU* items. H2 suggests that what really is being measured in a typical IS use study is therefore a mixture of correlations in the same column where .38 is located in the bottom. We have added superscripted ds <sup>(d)</sup> to indicate all the incongruent correlations that differ significantly from the congruent correlations of their respective columns. Continuing our consideration of a typical IS use study whose purpose is to measure *PU* and *U* on a general level, we observe that several of the incongruent correlations in the same column are statistically significantly different from .38. One way to estimate what the correlation mixture would yield as an aggregate measure is to calculate a weighted average of the individual correlations. In our case, the weighted average, when the percentages in Figure 3's second chart are used as weights (i.e., 41% for making charts, 23% for calculation in tables, and so on), is .06 – very different from .38. A conclusion is that in addition to being ambiguous for many respondents, the correlation .38 is not even an approximation of the expected value of the correlation mixture that may be measured in reality. Instead, in our data, its value seems irreducible to the underlying *PU–U* correlations on which its value could be based on.

Table 5 allows for similar analyses also for situations in which the research intentions are related to more specific uses. For example, if respondents are inquired about *PU* without specifying a particular use and are then led to think about calculations in tables (either explicitly thought a direct question about calculations, or indirectly through questions related to budgeting, for instance), the resulting incongruence errors may be close to the ones shown in the column with “Calculation in tables” as its heading. In this case the errors are quite small and not even statistically significant. In this case the measurements would provide a valid *PU–U* correlation even if its measurements were flawed due to incongruent interpretations. In principle, however, if unspecific *PU* and *U* measurements are used, all the incongruent correlations shown in Table 5 play a part in determining the correlations and covariances that are obtained. Reading from the table, when studying Excel, the range goes from -.39 to .45. This is a major spread and a source of error.

Confirmation of hypothesis H2 already proved that users will interpret the purposes of use differently, thereby orienting to questions on *PU* and *U* differently, which leads to incongruence between *PU* and *U* measurements. The analysis above demonstrates that the incongruence has a considerable effect on the resulting *PU–U* correlations. In other words, the data proves that if the purpose of use is left unspecified in a questionnaire, the resulting incongruence will lead to *PU–U* correlations that differ significantly from each other in a range from -.39 to .45 in the extremes. We find that the evidence presented above is sufficient for accepting hypothesis H3.

## Discussion

By finding support to all of the paper's three hypotheses, our study shows that IS research should better acknowledge heterogeneous use in its theorizing on IS use. By proving that measurement incongruence may lead to correlations that range from -.39 to .45 at worst, the empirical results of studies that only measure IS use on a general level may provide arbitrary results. Erroneous measurements will introduce errors in structural equation models' covariance matrices, leading both to false rejections of empirically valid models and false support for empirically invalid models. Incongruence, if not anticipated, is impossible to notice without adding purpose-gauging control questions in the survey. This is the case with practically all IS use model studies to date. We understand that this is a radical implication. However, we have shown that this is a possibility at least in studies on Excel use. It remains an unexplored issue whether such errors are possible more generally also in other contexts.

Our study had also weaknesses. First, despite our original intention, we could not find a significant use for the two measurement conditions for *U* (i.e., Standard-*U* and Conditional-*U*) in the study. As the analyses showed, the differences obtained using the two measurements were rarely statistically significantly different. The use of Conditional-*U* measurement therefore only introduced unnecessary complexity in our analysis without added value. This did not however compromise the findings of our study. Second, our data sample may have been biased towards heterogeneous Excel use. Because of self-selection sampling, it is possible that students with positive attitudes towards Excel took part in the study more actively than

others. The positive attitude may have resulted from respondents' contentment of finding many uses for Excel after the introductory course. Third, our study did not quantify the effect that incongruence of interpretations will have on the most common IS use models' fit measures (e.g., path coefficients,  $R^2$ , RMSEA, and so on). Doing this would have required us to collect a complete dataset of the selected models' relevant constructs at a purpose-specific level, to serve as a basis for comparative analysis. Such an examination will need a dedicated effort and will therefore be left for future research.

The empirical findings suggest that boundary conditions of validity exist to those IS use theories and models that postulate a strong  $PU-U$  relationship. Before further empirical studies are carried out, we do not know whether these theories are in fact valid only in contexts in which heterogeneous use is unlikely.

One way to re-establish the validity of acceptance models in the face of this paper's critique is to adapt the models so that the heterogeneity-related threats to validity will be taken into account. Much of the ambiguity can be removed by operationalizing the measurements of  $PU$  and  $U$  on the level of different purposes of use. On an operationalized level, a statement such as "Using the system in my job would enable me to accomplish tasks more quickly" can be converted into a set of statements where the notions "in my job" and "tasks" are replaced with wordings that are relevant for each purpose of use. The set of different uses can be found out with a qualitative pre-study among the users, determined *a priori* by the researchers themselves, or by using both of the methods. The dependent variable –  $U$  – should be also operationalized in a way that allows for distinguishing the different purposes of use within the data. However, in practice this may be challenging, especially if one wants to avoid common method bias and wants to apply direct behavioral measures such as log data. The crucial part of devising surveys is that researchers reflect on the use or uses that they intend to measure and then make sure that the questions in the survey address only those uses.

However, on one hand, theorizing on IS use on a level of idiosyncratic uses leads to poor generalizability. On the other hand, we concur with Burton-Jones and Straub's (2006) suggestion that IS use should be conceptualized based on the characteristics of user, task, and the technology. Our empirical findings are a proof that the task should not be forgotten in IS use theorizing. As a first attempt to theorize IS use with heterogeneous use taken into account, we would like to suggest reconsidering the merit of having  $PU$  as a predictor, and the frequency of use as the dependent variable. In different contexts of use, the desired outcomes of IS use are different and cannot be evaluated by measuring an extent of use (i.e.,  $U$ ). For instance, in knowledge work,  $U$  is unlikely to be a useful dependent variable. What is more important than  $U$  is the question whether IS helps the worker solve knowledge-intensive problems. Thus, a measure of problem-solving performance could be a better dependent variable in a knowledge work context. Alternatively, the dependent variable could measure increasing skillfulness in IS use, from repetitive through adaptive to innovative use (Sokura, 2012).  $PU$ , in turn, could then be framed with respect to the cognitive support provided by the system's features. Other contexts than knowledge work may require other kinds of conceptualizations for predictors and dependent variables. Conceptualizing theories for different tasks will be an important future work.

A third direction for novel theorizing is to take heterogeneous use itself as a topic of interest, and use it as a dependent variable. As we remarked earlier, we are not aware of earlier studies that would have attempted to quantify heterogeneity or attempted to predict it. In our previous study (Salovaara et al., 2011), we explored quantification of unexpected creative use of digital cameras, and predicted it with individual-level constructs such as use history, understanding of the underlying technology, intention to explore, and other possibly relevant constructs. Sun's (2012) feature-level model of adaptive system use is a similar attempt in this research direction. His model addresses triggers that initiate revisions of feature-level use. However, we want to make a note that dropping the level of analysis to the level of features may not necessarily solve the problem related to incongruence of interpretations. Similarly to an IS use in general, also system's individual features or their combinations may be used for multiple purposes. Features and purposes do not have a one-to-one correspondence by necessity. The correspondence depends on the heterogeneity of tasks that users set out to carry out with an IS, either by using individual features or their combinations.

More generally, we hope that the findings presented in this paper encourage other researchers also to carry out empirical studies with a goal of critically evaluating existing theories and specifying their boundary conditions of validity. Such falsificationist studies have had little foothold in IS research to date. As this study has exemplified, attempts to refute existing theories may give a rise to novel understanding



of their relative strengths and weaknesses and also provide a basis for new theories that have a better ecological fit with the phenomena of interest. Such an approach has been applied successfully not only in natural sciences but also in psychology (e.g., in a comparison of theories of creativity; see Dominowski, 1981; Weisberg and Alba, 1981) and economics (e.g., Allais 1953) in which falsification is part of the research tradition and a refutation of a theory can make a major contribution. Also IS research would benefit from such a practice.

## Acknowledgments

The authors want to thank Senior Fellow Tomi Seppälä for helpful discussions. This work was supported by Academy of Finland through funding for *Creative use of ICT* project.

## References

- Al-Natour, S. and Benbasat, I. 2009. "The Adoption and Use of IT Artifacts: A New Interaction-Centric Model for the Study of User-Artifact Relationships," *Journal of the Association for Information Systems* (10:9), pp. 661–685.
- Allais, M. 1953. "Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école Américaine," *Econometrica* (21:4), pp. 503–546.
- Barki, H., Titah, R., and Boffo, C. 2007. "Information System Use-Related Activity: An Expanded Behavioral Conceptualization of Individual-Level Information System Use," *Information Systems Research* (18:2), pp. 173–192.
- Bhattacharjee, A. and Barfar, A. 2011. "Information Technology Continuance Research: Current State and Future Directions," *Asia Pacific Journal of Information Systems* (21:2), pp. 1–18.
- Blackler, F. 1995. "Knowledge, Knowledge Work and Organizations: An Overview and Interpretation," *Organization Studies* (16:6), pp. 1021–1046.
- Boudreau, M.-C. and Robey, D. 2005. "Enacting Integrated Information Technology: A Human Agency Perspective," *Organization Science* (16:1), pp. 3–18.
- Burton-Jones, A. (2005). *New perspectives on the system usage construct*. Ph.D. dissertation, Georgia State University.
- Burton-Jones, A. and Hubona, G. S. 2005. "Individual Differences and Usage Behavior: Revisiting a Technology Acceptance Model Assumption," *The DATA BASE for Advances in Information Systems* (36:2), pp. 58–77.
- Burton-Jones, A. and Straub, Jr., D. W. 2006. "Reconceptualizing System Usage: An Approach and Empirical Test," *Information Systems Research* (17:3), pp. 228–246.
- Burton-Jones, A. and Gallivan, M. J. 2007. "Toward a Deeper Understanding of System Usage in Organizations: A Multilevel Perspective," *MIS Quarterly* (31:4), pp. 657–679.
- Davis, F. D. 1989. "Perceived Usefulness, Perceived Ease of User, and User Acceptance of Information Technology," *MIS Quarterly* (13:3), pp. 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), pp. 982–1003.
- DeSanctis, G. and Poole, M. S. 1994. "Capturing the Complexity of Advance Technology Use: Adaptive Structuration Theory," *Organization Science* (5:2), pp. 121–147.
- DiMicco, J., Millen, D. R., Geyer, W., Dugan, C., Brownholtz, B., and Muller, M. 2008. "Motivations for Social Networking at Work," in *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work (CSCW 2008)*, B. Begole and D. W. McDonald (eds.), New York, NY: ACM Press, pp. 711–720.
- Dominowski, R. L. 1981. "Comment on "An Examination of the Alleged Role of 'Fixation' in the Solution of Several 'Insight' Problems" by Weisberg and Alba," *Journal of Experimental Psychology: General* (110:2), pp. 193–198.
- Fadel, K. J. 2012. "User Adaptation and Infusion of Information Systems," *Journal of Computer Information Systems* (52:3), pp. 1–10.
- Fisher, R. A. 1921. "On the "Probable Error" of a Coefficient of Correlation Deduced from a Small Sample," *Metron* (1:1), pp. 3–32.
- Gasser, L. 1986. "The Integration of Computing and Routine Work," *ACM Transactions on Office*

- Information Systems* (4:3), pp. 205–225.
- Goodhue, D. L. and Thompson, R. L. 1995. "Task-Technology Fit and Individual Performance," *MIS Quarterly* (19:2), pp. 213–236.
- Griffith, T. L. 1999. "Technology Features as Triggers for Sensemaking," *Academy of Management Review* (24:3), pp. 472–488.
- Handel, M. and Herbsleb, J. D. 2002. "What Is Chat Doing in the Workplace?," in *Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work (CSCW 2002)*, E. F. Churchill, J. McCarthy, C. Neuwirth, and T. Rodden (eds.), New York, NY: ACM Press, pp. 1–10.
- Isaacs, E., Walendowski, A., Whittaker, S., Schiano, D. J., and Kamm, C. 2002. "The Character, Functions, and Styles of Instant Messaging in the Workplace," in *Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work (CSCW 2002)*, E. F. Churchill, J. McCarthy, C. Neuwirth, and T. Rodden (eds.), New York, NY: ACM Press, pp. 11–20.
- Jasperson, J. S., Carter, P. E., and Zmud, R. W. 2005. "A Comprehensive Conceptualization of Post-Adoptive Behaviors Associated with Information Technology Enabled Work Systems," *MIS Quarterly* (29:3), pp. 525–557.
- Karahanna, E., Straub, Jr., D. W., and Chervany, N. L. 1999. "Information Technology Adoption across Time: A Cross-Sectional Comparison of Pre-Adoption and Post-Adoption Beliefs," *MIS Quarterly* (23:2), pp. 183–213.
- Lassila, K. S. and Brancheau, J. C. 1999. "Adoption and Utilization of Commercial Software Packages: Exploring Utilization Equilibria, Transitions, Triggers, and Tracks," *Journal of the Management Information Systems* (16:2), pp. 63–90.
- Lee, Y., Kozar, K. A., and Larsen, K. R. T. 2003. "The Technology Acceptance Model: Past, Present, and Future," *Communications of the Association for Information Systems* (12:5), pp. 752–780.
- Limayem, M., Hirt, S. G., and Cheung, C. M. K. 2007. "How Habit Limits the Predictive Power of Intention: The Case of Information Systems Continuance," *MIS Quarterly* (31:4), pp. 705–737.
- Mackay, W. E. 1988. "Diversity in the Use of Electronic Mail: A Preliminary Inquiry," *ACM Transactions on Office Information Systems* (6:4), pp. 380–397.
- Majchrzak, A., Rice, R. E., Malhotra, A., King, N., and Ba, S. 2000. "Technology Adaptation: The Case of a Computer-Supported Inter-Organizational Virtual Team," *MIS Quarterly* (24:5), pp. 569–600.
- Nardi, B. A., Whittaker, S., and Bradner, E. 2000. "Interaction and Outeraction: Instant Messaging in Action," in *Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work (CSCW 2000)*, W. A. Kellogg and S. Whittaker (eds.), New York, NY: ACM Press, pp. 79–88.
- Nardi, B. A., Schiano, D. J., Gumbrecht, M., and Swartz, L. 2004. "Why We Blog," *Communications of the ACM* (47:12), pp. 41–46.
- Nurmi, P., Salovaara, A., Forsblom, A., Bohnert, F., and Floréen, P. In press. "PromotionRank: Ranking and Recommending Grocery Product Promotions Using Personal Shopping Lists," *ACM Transactions on Interactive Intelligent Systems*.
- Orlikowski, W. J. 1996. "Improvising Organizational Transformation Over Time: A Situated Change Perspective," *Information Systems Research* (7:1), pp. 63–92.
- Orlikowski, W. J. and Iacono, C. S. 2001. "Research Commentary: Desperately Seeking the "IT" in IT Research – A Call to Theorizing the IT Artifact," *Information Systems Research* (12:2), pp. 121–134.
- Popper, K. R. 1963. *Conjectures and Refutations: The Growth of Scientific Knowledge*. London, UK: Routledge.
- Robey, D. and Sahay, S. 1996. "Transforming Work Through Information Technology: A Comparative Case Study of Geographic Information Systems in County Government," *Information Systems Research* (7:1), pp. 93–110.
- Saga, V. L. and Zmud, R. W. 1994. "The Nature and Determinants of IT Acceptance, Routinization, and Infusion," in *Diffusion, Transfer and Implementation of Information Technology*, L. Levine (ed.), Amsterdam, The Netherlands: Elsevier, pp. 67–86.
- Salovaara, A. and Tamminen, S. 2009. "Acceptance or Appropriation? A Design-Oriented Critique on Technology Acceptance Models," in *Future Interaction Design II*, P. Saariluoma and H. Isomäki (eds.), London, UK: Springer, pp. 157–173.
- Salovaara, A., Helfenstein, S., and Oulasvirta, A. 2011. "Everyday Appropriations of Information Technology: A Study of Creative Uses of Digital Cameras," *Journal of the American Society for Information Science and Technology* (62:12), pp. 2347–2363.
- Salovaara, A. and Tuunainen, V. K. 2013. "Software Developers' Online Chat as an Intra-Firm Mechanism for Sharing Ephemeral Knowledge," in *Proceedings of the Thirty Fourth International Conference on*

- Information Systems (ICIS 2013)*, F. Pennarola, J. Becker, R. Baskerville, and M. Chau (eds.), Milan, Italy.
- Sokura, B. 2012. "Flow Experiences in Learning to Use a Spreadsheet Application," in *Proceedings of the 45th Hawaii International Conference on System Sciences (HICSS 2012)*, New York, NY: IEEE Computer Society, pp. 4613–4622.
- Straub, D., Limayem, M., and Karahanna-Evaristo, E. 1995. "Measuring System Usage: Implications for IS Theory Testing," *Management Science* (41:8), pp. 1328–1342.
- Straub, D. and del Giudice, M. 2012. "Use," *MIS Quarterly* (36:4), pp. iii–viii.
- Sun, H. 2012. "Understanding User Revisions When Using Information System Features: Adaptive System Use and Triggers," *MIS Quarterly* (36:2), pp. 453–478.
- Tyre, M. J. and Orlikowski, W. J. 1994. "Windows of Opportunity: Temporal Patterns of Technological Adaptation in Organizations," *Organization Science* (5:1), pp. 98–118.
- van der Heijden, H. 2004. "User Acceptance of Hedonic Information Systems," *MIS Quarterly* (28:4), pp. 695–704.
- Venkatesh, V. and Davis, F. D. 2000. "A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies," *Management Science* (46:2), pp. 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), pp. 425–478.
- Venkatesh, V., Thong, J. Y. L., and Xu, X. 2012. "Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology," *MIS Quarterly* (36:1), pp. 157–178.
- Wang, W. and Hsieh, J. J. P.-A. 2006. "Beyond Routine: Symbolic Adoption, Extended Use, and Emergent Use of Complex Information Systems in the Mandatory Organizational Context," in *Proceedings of the Twenty-Seventh Conference on Information Systems (ICIS 2006)*, paper 48.
- Weisberg, R. W. and Alba, J. W. 1981. "An Examination of the Alleged Role of "Fixation" in the Solution of Several "Insight" Problems," *Journal of the Experimental Psychology: General* (110:2), pp. 169–192.
- Wilkin, C. L. and Davern, M. J. 2012. "Acceptance of Post-Adoption Unanticipated IS Usage: Towards a Taxonomy," *The DATA BASE for Advances in Information Systems* (43:3), pp. 9–25.
- Wu, P. F. 2012. "A Mixed Methods Approach to Technology Acceptance Research," *Journal of the Association for Information Systems* (13:3), pp. 172–187.
- Zhao, D. and Rosson, M. B. 2009. "How and Why People Twitter: The Role that Micro-Blogging Plays in Informal Communication at Work," in *Proceedings of the ACM 2009 international conference on Supporting group work (GROUP 2004)*, S. Teasley, E. Havn, W. Prinz, and W. G. Lutters (eds.), New York, NY: ACM Press, pp. 243–252.