Advancing the Theory of Adaptive Structuration: The Development of a Scale to Measure Faithfulness of Appropriation

Wynne W. Chin • Abhijit Gopal • W. David Salisbury
College of Business Administration, Melcher Hall, University of Houston, Houston, Texas 77204-6282
wchin@uh.edu
Faculty of Management, University of Calgary, Calgary, Alberta, Canada T2N 1N4
gopal@acs.ucalgary.ca
Department of Management and Information Systems, College of Business and Industry, Mississippi State University, Mississippi State, Mississippi 39762-9851
dsalisbury@cobilan.msstate.edu

Adaptive Structuration Theory (AST) is rapidly becoming an influential theoretical perspective in research on advanced information technologies. However, there still exists a paucity of methods to capture critical AST constructs. This paper describes the development of an instrument to capture the extent to which users of an advanced information technology believe they have appropriated its structures faithfully. The development of such instruments is considered critical if the theoretical base provided by AST is to be fully exploited in understanding the use of advanced information technologies. The development procedure, which occurred in the context of the use of an electronic meeting system, was carried out in three phases that began with initial item development and proceeded through an exploratory to a confirmatory phase. Three experiments, two in the exploratory phase and one in the confirmatory phase, were performed. In the final phase, structural equation modeling techniques were used to confirm the convergent, discriminant, and nomological validity of the resulting five-item scale.
(*Adaptive Structuration Theory; Technology Appropriation; Electronic Meeting Systems; Structural Equations Modeling; Scale Development*)

Introduction

Information technologies (IT) are evolving along several dimensions at an increasingly rapid rate, with commensurate increases in the complexity of the social, procedural, and communicational arrangements associated with their use. While older technologies merely supported business transactions, *advanced information technologies* (Huber 1990, DeSanctis and Poole 1994) aim also to support coordination and interpersonal communication.

One form of advanced information technology, the electronic meeting system (EMS),[1] was conceived as a rational response to both a problem and an opportunity related to group work in general and to organizational meetings in particular. The problem

[1] Several terms have been used to label this kind of information technology to support group activity, including group decision support systems (GDSS), group support systems (GSS), and electronic meeting systems (EMS). We have chosen to adopt the latter term (following Dennis et al. 1988) for this paper, as it highlights the fact that these technologies relate to the support of groups in *meetings*.

1047-7047/97/0804/0342$05.00

was that group work and meetings were fraught with pitfalls and, consequently, less than desirable results. Groups encounter a variety of phenomena that diminish their effectiveness, such as groupthink (Janis 1982) and other process losses (Nunamaker et al. 1991). Conjecture and research suggested that these impediments to effective group process result in groups underachieving their potential. In view of the acknowledged importance of groups to organizations (Hackman and Kaplan 1974, Drucker 1988), these suboptimal outcomes were seen as a cause for concern, and research into techniques to improve group process gained prominence as a domain of scientific inquiry; a range of interventions were developed aimed at dampening the negative influences on and of group process and emphasizing its potential gains (Maier and Maier 1957, Dalkey and Halmer 1963, Van de Ven and Delbecq 1974).

The use of EMS, information technology to support the work of groups, was based on the essential premise that if effective information and interaction support were provided, the quality of group outcomes would improve (DeSanctis and Gallupe 1987, Nunamaker et al. 1991). Consequently, improved structuring of group interaction and information processing features have been the key elements in EMS design (Clapper and Prasad 1993).

Several EMS researchers, to whom DeSanctis and Poole (1994, p. 122) apply the label "decision theorists," assumed that features that enhanced groups' information processing capabilities would predictably lead to improvements in outcomes for all groups that used these features (DeSanctis 1993). Unfortunately, the results of the application of this input-output framework in EMS research have been inconsistent (Dennis and Gallupe 1993), resulting in the recent rethinking of the nature of groups, group work, and the type of support that an EMS can or should provide (Poole and DeSanctis 1990, Rao and Jarvenpaa 1991; Clapper and Prasad 1993, DeSanctis and Poole 1994). As a consequence, there has emerged an appreciation for the processes that intervene in the relationship between an EMS and the outcomes of its use.

Poole, DeSanctis, and their colleagues (Poole and DeSanctis 1990, DeSanctis and Poole 1994, Sambamurthy 1989, DeSanctis et al. 1994) have proposed a promising approach to the study of group process that accounts for the differences in outcomes that occur even when the same conditions (i.e., EMS) exist. Their approach, *adaptive structuration theory* (AST), suggests that groups are not merely information processing entities and that they have a social existence that must be considered when they use an EMS. This social aspect will determine how groups use, or *appropriate*, the EMS for their own purpose and will mediate any influence that improved information processing and interaction features have on meeting outcomes. Specific constructs that may be used to represent this social aspect are *faithfulness* of appropriation of an EMS, the degree of *consensus* on its appropriation, and group members' *attitudes* toward its use (Poole and DeSanctis 1990).

While AST appears to have considerable explanatory potential, the frequency of its use compared to the larger body of EMS literature has been low. The dearth of research on AST may be attributed partly to its complexity and to the fact that conveniently usable measures exist for only one of its constructs, attitude toward EMS use (Sambamurthy and Chin 1994, Gopal et al. 1993). While the other constructs can be measured using microcoding methods (Poole and DeSanctis 1992, DeSanctis and Poole 1994), these methods are time-consuming (Sambamurthy and Poole 1992, Sambamurthy and Chin 1994) and sometimes impractical to apply.

Measurement scales for faithfulness and consensus on appropriation, therefore, could prove extremely useful for the application of AST in EMS research. Even when AST is not explicitly applied or when it is applied in conjunction with other frameworks, its key constructs may help explain variations in EMS outcomes as well as or better than constructs currently applied, thereby addressing some major inconsistencies in EMS research. From this perspective, usable scales based on AST constructs could prove invaluable for EMS research in general.

More importantly, however, conveniently usable scales based on AST can facilitate the study of a wide range of advanced information technologies, many of which are undergirded, like EMS, by coordination and communication processes. The rapid deployment of these technologies makes the development of concise, convenient, and effective measures of key constructs a matter of considerable urgency. As DeSanctis and Poole (1994, p. 143) point out, "A critical challenge is to systematize the research so that technologies and interaction processes can be meaningfully assessed and comparative analysis is possible."

This paper reports on an effort to take up this challenge through the development of an instrument to measure the faithfulness of appropriation of an EMS. The development process is embedded within a carefully conducted procedure concentrated upon establishing the construct validity of the scale. As such, the paper restricts itself to construct validation rather than theory testing and is therefore not concerned with specifying a model that represents the propositions of AST and then testing it; all theoretical models used should be viewed merely as *plausible* representations of AST intended to validate the faithfulness construct. However, given the extent to which AST is implicated within the instrument, AST as theory is carefully considered in the development of the scale itself. In order to adapt an interpretive concept to a more positive framework, key assumptions and positions in AST are examined and extended in a process that may well be characterized, in the argot of AST, as structurational in character.

The paper begins with an overview of AST, then outlines the theoretical basis for the instrument development. It goes on to describe the development and validation of the instrument, which includes a multistage development/testing approach involving the use of confirmatory factor analysis to identify the appropriate factor structure and the use of structural equation modeling to examine the nomological validity of the instrument. The paper then offers conclusions about the importance of faithfulness of appropriation and suggestions for future work on enhancing adaptive structuration theory.

Adaptive Structuration Theory

Adaptive structuration theory has its roots in Anthony Giddens' structuration theory (1979, 1984, 1993), in which he rejects the extreme positions in the agency-structure debate, positing in the process that human action, while institutionally constrained, also influences and alters institutional arrangements, or structures, in the process of ongoing use. Structuration theory, which readily lends itself to "adaptation" due to Giddens' resistance to making an epistemological commitment (Bryant 1992), has assumed several guises in the organizational (see Ranson et al. 1980, Riley 1983, Pentland 1992, Whittington 1992) and information systems literature (see Orlikowski and Robey 1991, Orlikowski 1992). Poole and DeSanctis have adapted the theory to study EMS in particular (Poole and DeSanctis 1990, 1992) and advanced information technologies in general (DeSanctis and Poole 1994).

AST eschews the "technocentric" (DeSanctis 1989) view of technology use and emphasizes its social aspects because groups "mediate technological effects, adapting systems to their needs, resisting them, or not using them at all" (Poole and DeSanctis 1990, p. 177). Provided with an EMS to perform their work, groups dynamically create perceptions about the role and utility of the technology (Sambamurthy and Chin 1994, cf. Weick 1990, Sproull and Goodman 1990), and how it may be applied to their tasks. These perceptions can vary dramatically across groups, influencing how the EMS is used and hence mediating its impacts on group decision-making outcomes (Poole and DeSanctis 1990, DeSanctis and Poole 1994).

Central to AST is the notion of *structuration,* the process by which groups create and maintain a social system through the application of structures, which are rules and resources provided by the EMS, the task, organizational culture, group norms, and the knowledge represented by participants (DeSanctis and Poole 1994). The structures available to group members from these sources comprise the *structural potential* and the specific structures used are *structures in use* (Poole and DeSanctis 1990).

The choice of structures used depends on how groups *appropriate* the structures provided through the EMS and its context (henceforth, the term EMS will be used to refer to both the system and its context); groups select specific structures from the structural potential and adapt these to their needs (Poole and DeSanctis 1990). Appropriation, therefore, creates structures in use that may differ between groups even when the structural potential is constant (Poole and DeSanctis

1990, 1992). Appropriations are characterized by faithfulness (the extent to which structures provided to a group are used in a manner consistent with the *spirit* of the EMS), consensus (the extent of the agreement among group members on how the EMS should be used), and attitude (the views about using the EMS held by group members). While scales exist for the three attitude constructs of comfort, respect and challenge (Sambamurthy 1989, DeSanctis and Poole 1994), convenient measures for faithfulness and consensus on appropriation are lacking.

This paper restricts itself to a consideration of faithfulness alone in order to treat the scale development problem systematically: it is desirable to validate each such scale incrementally in relation to established constructs only (Chin and Todd 1995); the inclusion of two "unproven" constructs is likely to compromise the validity of the validation procedure itself for each individual construct. While either construct could have been selected for scale development (and, to an extent, the choice is arbitrary), the decision to select faithfulness was based on the reasoning that it would have greater applicability to advanced information technologies in general. Consensus on appropriation, by its very nature, is restricted to technologies that call for explicit agreement regarding use; while such explicit agreement might be considered relevant in the context of an electronic meeting system, it may not be applicable in the case of other advanced information technologies.

The Concept of the "Spirit"

The *spirit* may be described as the general objectives and procedures that the EMS aims to promote (Poole and DeSanctis 1992) and can be interpreted in a manner analogous to the spirit of the law. The concept of spirit, however, is a subtle one and the manner in which it will be used in this paper deserves elaboration, primarily because it serves as the "anchor" for a construct such as faithfulness, which can be evaluated *only with respect to* the spirit.

DeSanctis, Poole, and their colleagues (DeSanctis et al. 1994, DeSanctis and Poole 1994) propose that determination of the spirit may be accomplished by analyzing the design metaphor underlying the system, the system's features, the user interface, reference materials, and information from sources such as the system's designer or observations on how users perceive the system. Implicit among these sources is an externally defined, "objective" spirit and an internally (to the participant) defined, "subjective" spirit (see Figure 1).[2] We use the term *objective* to refer to the signifiers that the user interprets as indicative of the prescribed or "correct" approach, many of which are presented to the user by the facilitator or other "expert." The expert as observer similarly interprets a group of signifiers, many of which overlap those presented to and interpreted by the user, to develop an understanding of how the technology "ought" to be used. This interpretation is then used as a standard against which observed user behavior is evaluated in terms of technology usage. Our use of the terms is intended to skirt the issue of the subjectivity of the expert and emphasize the manner in which an externally imposed conception of the spirit is likely to be experienced by the user. We take the position that it is in this subjective guise that spirit is actually encountered, as an explicit or implicit construction in the mind of the individual.

A useful illustration of this notion can be found in the academic journal review process. As reviewers, we possess some conception of the aims and objectives, or spirit, of the review process. In this respect, it is difficult at best to conceive of an objective spirit: reviewers might (and will) conceive of the spirit in different ways and it is against this internal standard that each reviewer evaluates an unfolding review. Someone to whom the review process is new is likely to seek out or be provided with one or more such subjective conceptions of spirit and this person is likely to encounter these conceptions, initially, as objective. Eventually, however, this objectified spirit, in conjunction with the reviewer's growing experience, is internalized and becomes a subjective conception.

For the user of the EMS, too, initial encounters may call for an explicit description of the system's aims and objectives but over time the spirit of the system becomes internal and subjective. Even when the system is initially encountered, the explicit spirit is interpreted by the user and forms the basis for the eventual subjective conception. It stands to reason, therefore, that

[2] We thank Reviewer 1 for giving us the idea for Figure 1.

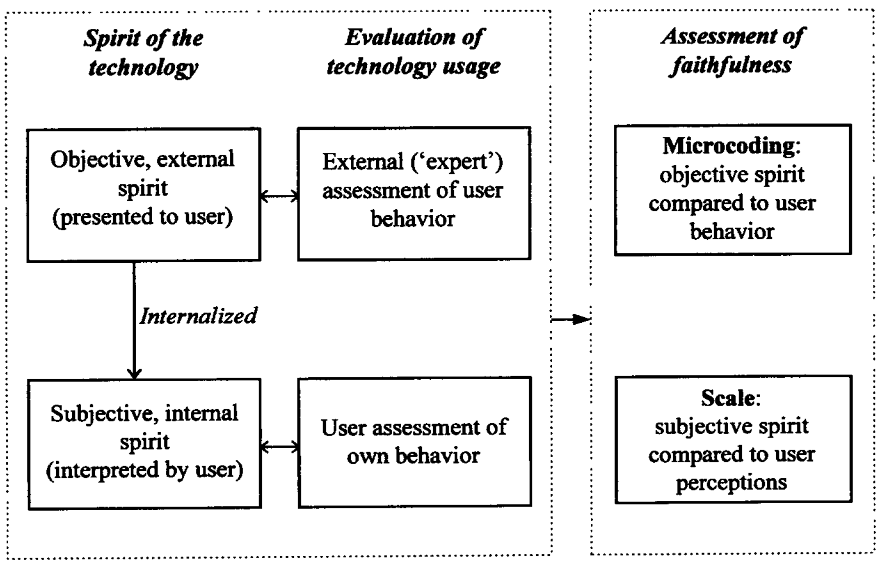
any assessment of *faithfulness*, which is an assessment of usage of a technology with respect to its spirit (see Figure 1), must take into account the internal *interpretation* of spirit against which it is evaluated. The only person capable of making such an assessment is the individual.

Two implications of this conceptualization of spirit should be noted. First, the spirit as encountered by the individual changes over time; and while it needs to be anchored initially to provide an explicit account of the system's aims and objectives, such anchoring will *eventually* be inappropriate because once the spirit is internalized, it cannot be pinned down as issuing from one or more specific sources. A scale designed to assess faithfulness early in an individual's experience with a system, therefore, must anchor faithfulness in an explicit conception of spirit. Later, faithfulness may be assessed relative to the individual's evolving interpretation of spirit, reflecting its internalization and subjectification. In the experiments reported in this paper, low experience with EMS was presumed (and later confirmed) and an explicit conception of spirit was therefore evoked.

Second, to assess faithfulness, it may be *insufficient* to employ external approaches such as microcoding (DeSanctis and Poole 1994) or a facilitator's assessment of faithfulness because these approaches assume the existence of an objective spirit against which the researcher or facilitator, rather than the individual user, assesses faithfulness. Such an ontological commitment (i.e., the objective spirit), by itself, may be untenable. While faithfulness can, for the purpose of triangulation, be assessed against the spirit as *presented*, the argument stated earlier suggests that faithfulness evaluated with respect to the spirit as *internalized* must also be assessed and should even be privileged over the other means. Thus, while the scale would have the obvious advantage of being a *convenient* means of assessing faithfulness, it is also a *crucial* means, in the absence of interview data, of understanding faithfulness at the level of interpretation.

It should be noted, however, that the fact that spirit refers to a phenomenon encountered by the individual rather than to an objective reality does not mean that some common conception of spirit cannot be discerned. Just as reviewers depend on their peers sharing their conception of the spirit of the review process, EMS users can usually be expected to have a shared meaning of the spirit of the system. A shared conception is especially likely if some representation of the spirit is adequately and explicitly communicated to group members (DeSanctis et al. 1989). Therefore, external measures could be premised on such a shared

Figure 1 **Relationship Between Spirit, Usage, and Faithfulness**



conception and could be used to augment the results obtained through an individually completed scale, as implied in Figure 1. What is achieved through such augmentation, however, is not validation of the scale but an evaluation of whether the observer and the user have similar interpretations of the aims and objectives, or spirit, of the system. Dissimilar assessments of faithfulness are most likely to indicate dissimilarities in the understanding of the aims of the system while similar assessments would indicate similar conceptions of spirit. As the purpose of this paper is exclusively one of scale validation and as external approaches do not, as argued above, contribute to such validation, the results of microcoding have not been incorporated in the paper.[3]

In assessing faithfulness, appropriations may be characterized on a range from *faithful* to *unfaithful* (DeSanctis and Poole 1994). A faithful appropriation occurs when those using the EMS follow the spirit of the system, while an unfaithful appropriation results when the EMS is used in a manner inconsistent with its spirit. Faithfulness of appropriation is not necessarily concerned with the precise duplication of the procedures provided by the EMS. Rather, it is concerned with whether the EMS is used in a manner consistent with its overall goals and objectives. A unique or innovative use of EMS by a group may well be a faithful appropriation as long as the use is consistent with the spirit that the system is intended to promote.

Development of the Faithfulness of Appropriation Scale

The process of the development of the faithfulness scale can be divided into three phases. The initial phase was *developmental;* items reflecting faithfulness of appropriation (designated FOA for all statistical analyses) were developed based on the work of Poole and DeSanctis (1990). These items were reviewed for content by several EMS researchers in order to ascertain whether they did indeed reflect the construct of interest.

[3] All meetings in the third experiment reported in this paper were videotaped and microcoding data will be available for further analysis of that data set.

The second phase was *exploratory;* the items were administered to individuals participating in an EMS experiment and analyzed using principal components analysis. Individual rather than group level data was considered appropriate because faithfulness, as has been pointed out earlier, is evaluated by the individual. However, the issue of whether group membership influenced individual responses should be considered; the existence of this potential influence was statistically tested. Those items that did not load well on the underlying factor (FOA) were revised and the modified scale was distributed to groups in a second EMS experiment. This time, items that did not load well on the underlying factor were dropped.

The final phase was *confirmatory;* convergent, discriminant, and nomological validity were assessed at this time. The scale was administered to participants in a third EMS experiment and was subjected to confirmatory factor analysis and structural equation modeling, resulting eventually in a scale of five items. Table 1 provides a summary of the phases of development.

In all three experiments carried out (two in Phase 2 and one in Phase 3), meetings were conducted by one of two trained and experienced facilitators (facilitators differed between experiments, with two specific individuals used in the first two experiments and two others in the third). Assignment of facilitators was balanced between treatments to control for facilitator effects. To ensure that facilitators expressly and consistently communicated the spirit of the set of structures that represented the EMS used in each experiment, guidelines were prepared that identified its aims and objectives. While the specific wording regarding spirit differed between the first two experiments on the one hand and the third experiment on the other, the content of what was described emphasized four aspects: each EMS was presented as computer-based, designed to encourage participation, based on structured procedures, and intended to improve group decision-making. This formulation of the spirit was seen by the researchers as applicable to all three EMS technologies used and was therefore invoked consistently across experiments. Moreover, the facilitators who conducted the meetings in each experiment agreed in advance of the experiment on the spirit that would be communicated and could therefore be expected to have been

Table 1 Project Summary

	Phase 1	Phase 2	Phase 3	
Purpose	Developmental	Exploratory	Confirmatory	
Data Collection	None	Experiments 1 and 2	Experiment 3	
Activity	Initial item development	Instrument testing and refinement	Model testing	Model confirmation

consistent. Consistency, in this regard, was understood to mean the explicit communication of the four aspects of spirit. No specific process for communicating spirit was enforced other than to ensure that the four aspects were communicated at the beginning of the meeting and, as necessary, at other times during the meeting. Facilitators were expected to remain true to these guidelines, but were not required to adhere to them as they would to a script: the inflexibility of strictly scripted interaction with participants may have detracted from the external validity of each experiment and was therefore avoided. Subsequent analysis revealed that no significant facilitator-based differences existed for any of the variables reported in this paper. While this finding does not guarantee consistency in the communication of spirit, it lends some credence to the expectation that a reasonable degree of consistency did indeed exist.

Phase 1: Initial Item Development

The initial challenge in developing the faithfulness scale was to create items that would elicit responses relative to the spirit of the system. In view of the novelty of EMS to the users and the poor likelihood of their already having developed an internal conception of the spirit, it was considered important to explicitly convey the concept of the spirit in a concise, understandable manner. The following working definition of faithfulness of appropriation was adopted:

> The extent to which a group's use of the EMS structures was consistent with the original design intent of the system developers.

This definition is modified from that found in Poole and DeSanctis (1990; cf. DeSanctis et al. 1994). The term "spirit" was replaced with "original design intent of the system developers." While DeSanctis and Poole (1994, p. 126) state that "it is impossible to wholly realize [designers'] intent," therefore making the substitution of *spirit* with *original design intent* seemingly problematic, the substitution is considered appropriate for several reasons.

First, the purpose was not to adopt an alternative definition of spirit but rather to *represent* the concept to respondents without using the term explicitly (due to the possibility of misinterpretation). Original design intent can serve as an effective symbol that *represents* the spirit of an EMS or any advanced information technology. It is, moreover, a parsimonious symbol that incorporates most if not all the factors that DeSanctis and Poole (1994) identify as sources from which the spirit of a system can be ascertained: its design metaphor, its features, its user interface, and its training and help facilities. In fact, they go on to state that, "Early on, when a technology is new . . . spirit is put forth by the designers . . ." (p. 127). Second, the choice of a single such symbol of spirit is justified from a methodological standpoint: by focusing on original design intent, all respondents would be exposed to the same explicit rendition of the concept. While any subsequent interpretation of the concept may not be consistent, these interpretations would at least be anchored in a consistent source. Third, the idea of the original design intent is easily understandable and is not diffuse; respondents can focus on original design intent as the source of the system's aims and objectives. Moreover, original design intent does not detract in any way from the meaning of spirit available through other potential sources.

To develop indicators which would reflect faithfulness, items were created by using various combinations of the wording for the definition of the construct. In view of the novelty of the technology, the relative complexity of its use, and the intricate nature of AST, it was felt that the development of indicators would need to be undertaken by individuals conversant with these issues. The process, therefore, was undertaken by the two researchers who acted as facilitators in the first two experiments. Both individuals had experience with facilitation, EMS use, and the application of AST. The items they developed were worded in a way intended to evoke the spirit of the technological intervention with statements such as, "... the designers of the EMS would regard our use of the system as inappropriate." Care was taken to ensure that each item reflected some aspect of the definition of the construct. This approach led to the development of the 11 items in Table 2. These items were then reviewed for content and deemed acceptable by one other EMS researcher at the same university and two EMS researchers at other universities. All three researchers were conversant with AST.

Phase 2: Instrument Testing and Refinement

The primary activity conducted during this phase was exploratory analysis to test the applicability of the items developed. Consequently, principal components analysis was utilized to make an initial assessment about which items appeared to adequately represent the FOA construct. Two sets of analyses were conducted, each corresponding to one of two experiments.

In the first experiment, the 11 FOA items were administered to students at a western Canadian university after they had completed a session in which they performed a task using an EMS (see Table 3 for descriptive statistics). The task performed by the groups was "A Case of Student Misconduct" (Gopal 1991, adapted from Beauclair 1987). The spirit of the EMS intervention was communicated by the facilitators as they conducted the meetings. The two EMS used were Ventana Corporation's *GroupSystems* and Option Technologies' *OptionFinder.*

Exploratory factor analysis using principal components extraction was performed to test the data from the first experiment. Using a combination of the scree plot and eigenvalue-greater-than-one rule, a three-component solution was considered most appropriate.

Table 2 Faithfulness of Appropriation Items

Item	Exp. 1	Exp. 2	Exp. 3	Final Items
My group's use of the EMS was consistent with the goals and attitudes the system *aims [aimed]* to promote.	*foa1*	*foa1*		
My group used the EMS as originally intended by the creators.	*foa2*	*foa2*		
My group's use of the EMS probably differs considerably from the purpose for which it was originally designed.	*foa3*	*foa3*	*foa1*	
The developers of the EMS would *probably be shocked at [disagree with]* how our group used the system.	*foa4*	*foa4*	*foa2*	*foa1*
Our group's use of the EMS was likely at odds with its original intent.	*foa5*	*foa5*	*foa3*	
Our group probably used the EMS *in relatively novel ways [improperly]*.	*foa6*	*foa6*	*foa4*	*foa2*
Our group's use of the EMS is faithful to its original design.	*foa7*	*foa7*		
It would be ironic to have the original developers of the EMS *see how our group used the system [would view our group's use of the system as inappropriate]*.	*foa8*	*foa8*	*foa5*	*foa3*
Our group failed to use the EMS as it should have been used.	*foa9*	*foa9*	*foa6*	*foa4*
We did not use the EMS in the most appropriate fashion.	*foa10*	*foa10*	*foa7*	*foa5*
Ideally we should have used the EMS in a different way.	*foa11*	*foa11*	*foa8*	

Between the first and second phases, italicized terms outside brackets [] were substituted with the terms inside brackets (e.g., the altered form of foa6 *in Phase 2 and thereafter was, "Our group probably used the EMS improperly").*

Entries in the last four columns are variable names assigned to items

Table 3 Descriptive Statistics

			Exp. 3	
	Exp. 1	Exp. 2	Testing	Confirmation
Number of Subjects	114	284	90	228
Number of Groups	27	13	20	50
Group Size Range	3–5	17–32	4–5	4–5
Gender Distribution (%male/female)	60/40	53/47	57/43	52/48

Though the scree plot suggested a smaller number of factors than the eigenvalue rule, we opted to err on the conservative side by including more factors to avoid the possibility of missing relevant ones. The results of a varimax rotation yielded a less interpretable result. Thus, Table 4 presents the results of an unrotated solution. All but four items loaded on the first factor, designated as faithfulness of appropriation. As the goal was to be inclusive during initial item development, items that loaded higher than 0.60 on the faithfulness factor but less than 0.40 on any other factor were considered acceptable and were kept without modification. Of the five items that did not meet these criteria, four were modified (Items 1, 4, 6, and 8) in the hope of improving their psychometric properties. While item 7 did not meet the criteria either, it was left unchanged on the theoretical grounds that it was closest in wording to the construct definition adopted.

The revised scale (Table 2) was subjected to further testing in a second experiment at the same university (descriptive statistics are presented in Table 3). As in the previous study, the questionnaire was administered to student groups after they had completed a task using an EMS. The task in this experiment called for subjects to help close a substantial gap between the projected income and expenses of a business school by recommending budget reductions. The EMS software used was *GroupSystems*.

The results of the principal components analysis, in conjunction with a scree plot and the eigenvalue rule, suggested a two-component solution. Varimax rotation, in this situation, did yield a more interpretable result (presented in Table 4). The same item selection criteria used for the data set for the first experiment were applied. Items 1, 2, and 7 had higher loadings on the second factor than on the first and were therefore dropped from further consideration. Item 3 was retained because it was considered close enough to the 0.60 cutoff. The result was the eight items (Table 2) that provided the basis for the third experimental study.

Phase 3: Confirmatory Analysis

The primary activity during this phase was the assessment of convergent, discriminant, and nomological validity, with each successive validity measure providing a more rigorous test of the scale than the previous one. Moreover, the data set was split into two to facilitate cross-validation (the covariance matrices for the two resulting data sets are presented in Table 5). The result of this phase was a set of five items that could be used to measure faithfulness.

The eight items that resulted from the exploratory analyses in the second phase were administered in a third experiment to 330 undergraduate subjects (70 groups) at the same university. The data set was split such that the first 20 groups (94 students) represented the model testing sample, while the remaining 50 groups (236 students) represented the model confirmation sample. After eliminating cases due to missing responses, the actual sample sizes were 90 for the model testing set and 228 for the confirmation of the model. Descriptive statistics for this experiment are presented in Table 3.

Procedures. The degree of restrictiveness (Silver 1990, DeSanctis et al. 1989; cf. Wheeler et al. 1993 and DeSanctis and Poole 1994) of the EMS was manipulated in order to induce variance in faithfulness. As the creation of variance was the only objective of this manipulation, it was not considered of any relevance in the context of the scale validation and did not, therefore, figure in the validation procedure. A restrictive treatment was designed that was expected to increase the likelihood of a faithful appropriation; the nonrestrictive treatment, on the other hand, allowed for the possibility of less faithful appropriations. In the restrictive treatment, administered to half the groups, a facilitator led each group through an on-screen agenda, thereby limiting the range of options in using the system. In the nonrestrictive treatment, the remaining groups were allowed to use (or not use) the EMS in

Table 4 Principal Components Analysis Loadings for Faithfulness Items

Item	Exp. 1 (unrotated) Factor 1	Exp. 1 (unrotated) Factor 2	Exp. 1 (unrotated) Factor 3	Exp. 2 (varimax) Factor 1	Exp. 2 (varimax) Factor 2	Exp 3 (varimax) Factor 1
foa1	0.595	−0.396	−0.024	−0.139	0.898	
foa2	0.685	−0.352	−0.019	−0.190	0.882	
foa3	0.764	0.003	0.246	0.598	−0.369	0.696
foa4	0.707	0.430	0.111	0.714	−0.048	0.912
foa5	0.757	0.181	0.117	0.704	−0.100	0.726
foa6	0.356	0.591	0.297	0.720	−0.281	0.872
foa7	0.620	−0.531	−0.083	−0.460	0.485	
foa8	0.417	−0.205	0.741	0.701	−0.239	0.907
foa9	0.702	−0.091	−0.229	0.824	−0.278	0.857
foa10	0 704	0.218	−0.398	0.712	−0.278	0.891
foa11	0.647	0.229	−0 386	0.776	−0 171	0.847
Eigenvalue	*4.57*	*1 28*	*1.09*	*4.43*	*2.29*	*5.67*

any manner. The spirit of the EMS was communicated consistently irrespective of treatment. The EMS used was *VisionQuest*, a product of Collaborative Technologies Corporation.

Group participants were seated around a table, with an IBM Thinkpad 700C notebook computer available to each person. A training session was conducted for each group prior to its performance of the experimental task. The experimental task used was an adaptation to the Canadian context of the "School of Business Policy Task" (Wheeler and Mennecke 1992), a hidden profile task (Stasser 1992) in which each group member is made aware of only a portion of the task information, requiring the group to work together to reach an optimal solution.

Validity Testing Procedure. As with the previous phases, an initial principal components analysis was performed on the model testing data. Both the scree plot and eigenvalue rule suggested a single factor solution. All eight items demonstrate acceptable loadings with six at or above 0.85 (see Table 4).

Even though the eight items in the instrument appeared to load well on FOA, a more rigorous analysis of this scale was considered critical. To establish its validity, it was important to establish that (a) the item measures related strongly to the FOA construct and not to any other construct (convergent validity), (b) the FOA construct differed from all other constructs (discriminant validity), and (c) the FOA construct related to other constructs of interest as predicted while exhibiting convergent and discriminant validity (nomological validity).

The use of structural equation modeling facilitates validity testing. Using it, models that more closely resemble the hypothesized construct and its relationship to other constructs can be specified and tested. As the purpose of constructing such a model is to validate a scale rather than to test a theory, it is sufficient to construct a model that *plausibly* represents the theoretical context of the construct for which the scale is developed. Moreover, the scales that represent the theoretical context should themselves have been tested in order to more accurately test the scale under construction. The constructs selected to test the FOA scale, therefore, were represented by validated scales that could be (and, in fact, had been) plausibly used in the AST context.

Initial Test for Convergent Validity. Structural equation modeling was used to conduct an initial test of convergent validity through confirmatory factor analysis (CFA) (see Long 1983 for the distinction between exploratory and confirmatory factor analysis). Rather than merely specifying the number of components and items to be analyzed, CFA allows specification of the exact relationship between the common

Table 5 Covariance Matrix

(a) Model Testing Data (Phase 3)

foa1	2.883																									
foa2	1.990	2.142																								
foa3	2.119	1.880	2.288																							
foa4	2.072	1.808	1.927	2.846																						
foa5	2.435	1.806	2.072	2.348	3.262																					
eou1	0.758	0.575	0.544	0.761	0.400	2.939																				
eou2	0.893	0.535	0.656	0.931	0.742	2.350	3.033																			
eou3	0.768	0.647	0.695	0.701	0.421	2.400	2.414	2.851																		
eou4	0.612	0.366	0.426	0.872	0.419	2.417	2.409	2.241	2.806																	
eou5	0.417	0.366	0.263	0.620	0.168	2.817	2.563	2.573	2.709	3.381																
eou6	0.635	0.597	0.533	0.705	0.238	2.589	2.278	2.583	2.404	2.841	2.933															
ufl1	0.802	0.403	0.531	1.041	0.840	1.975	2.107	1.785	2.206	2.286	2.109	2.992														
ufl2	0.245	0.111	0.139	0.376	0.270	1.514	1.547	1.397	1.734	1.805	1.572	1.991	2.868													
ufl3	0.385	0.186	0.225	0.408	0.293	1.750	1.793	1.566	1.861	2.037	1.825	2.261	2.651	2.988												
ufl4	0.355	0.163	0.179	0.387	0.193	1.656	1.716	1.486	1.844	1.994	1.790	2.251	2.811	2.896	3.204											
ufl5	0.397	0.079	0.150	0.582	0.252	1.744	1.819	1.410	1.922	1.953	1.817	2.406	2.430	2.576	2.702	2.929										
ufl6	0.340	0.167	0.216	0.548	0.240	1.922	1.898	1.754	2.096	2.179	2.049	2.251	2.781	2.893	2.981	2.730	3.262									
dss1	0.517	0.426	0.395	0.444	0.625	0.233	0.263	0.315	0.207	0.300	0.295	0.308	0.402	0.339	0.413	0.334	0.314	1.005								
dss2	0.590	0.575	0.477	0.627	0.732	0.361	0.362	0.322	0.380	0.443	0.399	0.293	0.281	0.267	0.290	0.179	0.277	0.668	0.915							
dss3	0.415	0.395	0.496	0.372	0.398	0.222	0.241	0.232	0.204	0.193	0.240	0.227	0.235	0.281	0.234	0.217	0.253	0.466	0.458	0.925						
dss4	0.645	0.471	0.522	0.758	0.680	0.194	0.389	0.097	0.291	0.193	0.155	0.365	0.147	0.207	0.295	0.392	0.225	0.414	0.350	0.270	0.967					
dss5	0.461	0.351	0.319	0.573	0.432	0.433	0.407	0.354	0.585	0.544	0.552	0.615	0.452	0.464	0.477	0.494	0.521	0.520	0.538	0.456	0.448	0.962				
ss1	−0.044	0.145	0.108	0.106	0.060	−0.006	−0.088	−0.185	−0.024	−0.113	−0.092	−0.013	0.365	0.222	0.343	0.333	0.360	0.213	0.078	0.105	0.288	0.221	0.885			
ss2	0.138	0.259	0.211	0.065	0.165	0.153	0.117	0.018	0.092	0.156	0.155	0.284	0.407	0.416	0.514	0.404	0.454	0.205	0.119	0.144	0.222	0.206	0.542	0.814		
ss3	0.245	0.408	0.354	0.360	0.315	0.244	0.145	0.184	0.200	0.267	0.321	0.243	0.267	0.264	0.269	0.296	0.359	0.219	0.230	0.187	0.188	0.279	0.403	0.449	0.596	
ss4	0.166	0.162	0.292	0.225	0.177	0.050	0.070	−0.095	0.080	−0.048	0.026	0.191	0.284	0.263	0.295	0.414	0.288	0.122	0.035	0.107	0.318	0.211	0.525	0.489	0.366	0.759
ss5	0.380	0.377	0.310	0.295	0.350	0.136	−0.016	0.010	0.053	0.101	0.132	0.136	0.205	0.171	0.188	0.202	0.206	0.259	0.267	0.154	0.266	0.329	0.305	0.346	0.380	0.255	0.556
	foa1	foa2	foa3	foa4	foa5	eou1	eou2	eou3	eou4	eou5	eou6	ufl1	ufl2	ufl3	ufl4	ufl5	ufl6	dss1	dss2	dss3	dss4	dss5	ss1	ss2	ss3	ss4	ss5

(b) Model Confirmation Data (Phase 3)

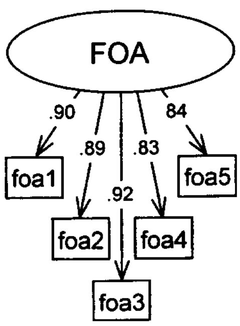
foa1	2.381																											
foa2	1.528	2.060																										
foa3	1.444	1.474	1.680																									
foa4	1.537	1.531	1.441	1.995																								
foa5	1.413	1.430	1.333	1.827	2.324																							
eou1	0.265	0.352	0.216	0.294	0.224	2.402																						
eou2	0.233	0.386	0.323	0.433	0.436	1.947	2.251																					
eou3	0.210	0.293	0.249	0.325	0.235	2.058	1.907	2.336																				
eou4	0.271	0.305	0.325	0.330	0.264	1.949	1.869	2.074	2.372																			
eou5	0.156	0.198	0.259	0.203	0.119	2.069	1.926	2.094	2.156	2.560																		
eou6	0.223	0.247	0.258	0.267	0.148	2.004	1.927	2.116	2.109	2.320	2.413																	
ufl1	0.124	0.270	0.138	0.344	0.223	1.451	1.443	1.443	1.449	1.537	1.495	2.572																
ufl2	0.265	0.389	0.260	0.371	0.274	1.686	1.570	1.545	1.611	1.719	1.627	2.247	2.650															
ufl3	0.241	0.355	0.226	0.394	0.293	1.521	1.485	1.468	1.561	1.624	1.541	2.286	2.447	2.696														
ufl4	0.180	0.366	0.218	0.389	0.305	1.578	1.449	1.527	1.517	1.575	1.530	2.187	2.383	2.421	2.558													
ufl5	0.121	0.313	0.179	0.354	0.241	1.517	1.395	1.536	1.494	1.490	1.498	2.116	2.260	2.259	2.309	2.516												
ufl6	0.179	0.357	0.246	0.410	0.339	1.744	1.624	1.688	1.715	1.808	1.752	2.272	2.539	2.507	2.518	2.406	2.871											
dss1	0.474	0.436	0.379	0.449	0.458	0.258	0.327	0.297	0.290	0.268	0.277	0.402	0.413	0.445	0.389	0.344	0.414	0.961										
dss2	0.445	0.585	0.429	0.475	0.476	0.190	0.263	0.218	0.209	0.185	0.167	0.350	0.392	0.458	0.401	0.372	0.384	0.572	0.907									
dss3	0.371	0.427	0.384	0.390	0.395	0.218	0.252	0.251	0.239	0.271	0.216	0.261	0.254	0.271	0.269	0.205	0.256	0.420	0.479	0.715								
dss4	0.310	0.363	0.348	0.358	0.285	0.270	0.265	0.305	0.244	0.246	0.243	0.355	0.451	0.431	0.424	0.382	0.400	0.294	0.323	0.283	1.071							
dss5	0.497	0.492	0.448	0.502	0.553	0.151	0.212	0.200	0.207	0.227	0.220	0.194	0.255	0.264	0.312	0.258	0.297	0.506	0.480	0.510	0.309	0.808						
ss1	0.218	0.238	0.218	0.239	0.209	0.029	0.005	0.025	0.036	0.106	0.027	−0.046	0.025	−0.057	−0.006	0.021	−0.030	0.188	0.218	0.224	0.189	0.267	0.598					
ss2	0.251	0.296	0.286	0.283	0.254	0.002	−0.002	0.031	−0.008	0.038	0.010	0.007	0.040	−0.008	0.058	0.072	0.045	0.240	0.245	0.273	0.215	0.305	0.452	0.676				
ss3	0.361	0.380	0.301	0.374	0.425	−0.040	−0.041	−0.079	−0.102	−0.068	−0.046	0.039	0.047	0.024	0.082	0.061	0.057	0.263	0.244	0.288	0.150	0.399	0.352	0.420	0.709			
ss4	0.145	0.200	0.194	0.228	0.228	0.076	0.082	0.076	0.064	0.090	0.101	0.020	0.057	−0.005	0.033	0.088	0.067	0.148	0.122	0.197	0.066	0.224	0.355	0.373	0.368	0.549		
ss5	0.351	0.364	0.342	0.456	0.570	−0.088	0.083	−0.053	−0.057	−0.061	−0.053	−0.079	−0.035	−0.026	−0.022	−0.039	−0.004	0.289	0.268	0.276	0.180	0.377	0.338	0.368	0.456	0.320	0.798	
	foa1	foa2	foa3	foa4	foa5	eou1	eou2	eou3	eou4	eou5	eou6	ufl1	ufl2	ufl3	ufl4	ufl5	ufl6	dss1	dss2	dss3	dss4	dss5	ss1	ss2	ss3	ss4	ss5	

factors and the items used to measure them as well as the linkages among the factors. Ideally, this model specification is based both on theory and substantive knowledge. Furthermore, various overall model fit indices are obtained to determine how well the model explains the sample data.

CFA was used to examine the convergent validity of the FOA scale by specifying a single factor model using AMOS 3.1 (Arbuckle 1993). While all eight items had loadings above 0.60 (see Figure 2), the minimal level above which convergent validity is suggested (Bagozzi and Yi 1988), the model goodness of fit was relatively poor. To improve goodness of fit and ensure greater reliability for each indicator, items with loadings below 0.80 were dropped. Although item 8 had a loading above 0.80, it was eliminated along with the other two weaker items because dropping it increased goodness of fit. Thus, items 2, 4, 5, 6, and 7 (Cronbach $\alpha = 0.94$) comprised the "final" set of measures for FOA (Table 2; see Table 6 for further item details). Overall, except for the χ^2 measure, the model fit indices (Table 7) surpass the recommended value for a good model, suggesting that the item measures reflect a single factor.

It should be noted that while the number of items desirable in the scale was not prespecified, two considerations were taken into account in this regard. First, to make for convenient usage, a *parsimonious* number of items was sought; this consideration favored the use of a smaller number of items. Second, to enhance its psychometric value, an adequately *overidentified* scale (in a structural equation modeling sense) was sought; this consideration favored the use of a larger number of items. As the scale needed three items to be *just identified,* it was felt that a scale of five or six items would acceptably meet both criteria.

Figure 2 **Confirmatory Factor Analysis**



Data set: Model testing
Purpose: Convergent validity
FOA Faithfulness of appropriation
All path estimates are standardized.

While this initial model provided an early indication of convergent validity, it tested the FOA scale in isolation. It was considered important to further investigate convergent validity by testing the scale (a) in relation to other explanatory constructs of interest to ensure that the FOA items did not relate better to these other constructs and (b) in the context of a model that related FOA to other, endogenous (i.e., dependent), constructs based on theory. This further testing, which provided a more rigorous assessment of convergent validity, was done in conjunction with the subsequent tests for discriminant and nomological validity.

Discriminant Validity. Having determined that the five items demonstrated convergent validity, the next step was to determine the extent to which they exhibited discriminant validity as well: to what extent did the items appear to measure the construct of interest and not relate more strongly to other constructs? To test this, and in keeping with the need to use validated scales for the testing, the FOA scale was tested against two other scales with established credentials, *perceived ease of use* (EOU) and *perceived usefulness* (UFL) (Davis 1989).

Davis (1989, p. 320) defines *perceived usefulness* as "the degree to which a person believes that using a particular system would enhance his or her job performance," and *perceived ease of use* as "the degree to which a person believes that using a particular system would be free of effort." The attitudes used in past EMS research have, to a large extent, used similar concepts and certain EMS studies have actually used these specific constructs (see Gopal et al. 1993, Sambamurthy and Chin 1994). The actual items are slight modifications of Davis' (1989) original items and are presented in Table 6.

Two sets of confirmatory factor analyses were performed using the model testing data in which FOA was modeled to correlate with either EOU or UFL (the

Table 6 Items Used to Assess Convergent, Discriminant, and Nomological Validity (Phase 3)

Construct	Item
Faithfulness of appropriation	The developers of the EMS would disagree with how our group used the system.
	Our group probably used the EMS improperly.
	The original developers of the EMs would view our group's use of the system as inappropriate.
$\alpha_{testing} = 0.94$	Our group failed to use the EMs as it should have been used.
$\alpha_{confirm} = 0.93$	We did not use the EMs in the most appropriate fashion.
The items above represent the final FOA scale They were measured on seven point Likert scales anchored by the following. extremely likely; quite likely; slightly likely; neither; slightly unlikely, quite unlikely; extremely unlikely.	
Ease of use	Learning to operate an EMS *is [would be]* easy for groups I work with.
$\alpha_{testing} = 0.97$	Groups I work with *find [would find]* it easy to get an EMS to do what they *want [would want]* it to do.
$\alpha_{confirm} = 0.97$	Groups I work with *find [would find]* their interaction with an EMs clear and understandable.
	Groups I work with *find [would find]* an EMS to be flexible to interact with.
	It *is [would be]* easy for groups I work with to become skillful at using an EMS.
	Overall, groups I work with *find [would find] an EMS easy to use.*
Usefulness	*Using an EMS enables [would enable]* groups I normally work with to accomplish their tasks more quickly
$\alpha_{testing} = 0.97$	Using an EMS *improves [would improve]* the performance of groups I work with.
$\alpha_{confirm} = 0.98$	Using an EMS *increases [would increase]* the productivity of groups I work with.
	Using an EMS *enhances [would enhance]* the effectiveness of groups I work with.
	Using an EMS *makes it [would make it]* easier for groups I work with to carry out their tasks.
	Groups I work *find [would find]* an EMS useful for group work
The two scales above were administered in exactly the form shown The italicized items within brackets were added to the original wording of the scales (Davis 1989) to account for the fact that participants were not expected to have used the EMS previously, thereby being able only to speculate on its usefulness or ease of use. All items were measured on seven-point Likert scales anchored by the following: extremely likely; quite likely; slightly likely; neither; slightly unlikely; quite unlikely; extremely unlikely.	
Decision scheme satisfaction	How would you describe your group's problem solving process?
	efficient/inefficient
	coordinated/uncoordinated
	fair/unfair
$\alpha_{testing} = 0.82$	*confusing/understandable*
$\alpha_{confirm} = 0.81$	*satisfying/dissatisfying*
Solution satisfaction	To what extent does the final solution reflect your inputs?
	To what extent do you feel committed to the group solution?
	To what extent are you confident that the group solution is correct?
$\alpha_{testing} = 0.87$	To what extent do you feel personally responsible for the correctness of the group solution?
$\alpha_{confirm} = 0.87$	How satisfied or dissatisfied are you with the process by which your group made the decision?

(The two scales above are from Green and Taber 1980 and were administered exactly as shown. The five decision scheme satisfaction items were measured on five-point scales anchored by the adjectives pairs shown. The solution satisfaction items were measured on five-point Likert scales. The anchors for the first four of these items were: not at all; to a little extent; to some extent; to a great extent; to a very great extent. The anchors on the last item were very dissatisfied; somewhat dissatisfied, neither; somewhat satisfied, very satisfied)

Cronbach α reliabilities of all scales are shown in Table 6). Assuming the overall model fit indices are adequate, discriminant validity is suggested if the correlation between constructs is not equal to 1.00. It can be more rigorously tested using a χ^2 difference test where the χ^2 measures for two analyses are compared. In one analysis, the correlation between the constructs of interest (e.g., FOA and EOU) is fixed at 1.00, thereby assuming that the constructs are identical. In the second analysis, the correlation between the constructs is allowed to be freely estimated. There are no other differences in the model specifications for the two anal-

Table 7 Model Fit Indices for Validity Testing Models (Phase 3)

Statistic	Rec. Value	Convergent CFA	Discriminant		Nomological	
			FOA/EOU	FOA/UFL	Testing	Confirm
χ^2		18.73	109.94	92.97	569.97	636.80
χ^2, independence model		425.03	1207.01	1286.18	2893.92	6762.09
Degrees of freedom		5	43	43	317	317
χ^2 significance (p-value)	>0.05	0.002	0.000	0.000	0.000	0.000
χ^2/d.f. (Wheaton, et al., 1977)	<5.00	3.75	2.56	2.16	1.80	2.01
RMR(C) (Hu and Bentler, 1995)	<0.10	0.03	0.05	0.06	0.08	0.06
δ-1 (Bentler and Bonett, 1980)	>0.90	0.96	0.91	0.93	0.80	0.91
ρ-1 (Bollen, 1986)	>0.90	0.91	0.88	0.91	0.78	0.90
δ-2 (Bollen, 1989)	>0.90	0.97	0.94	0.96	0.90	0.95
ρ-2 (Tucker and Lewis, 1973)	>0 90	0.93	0.93	0.95	0.89	0.95
CFI (Bentler, 1990)	>0.90	0.97	0.94	0.96	0.90	0.95
RNI (McDonald and Marsh, 1990)	>0.90	0.97	0.94	0.96	0.90	0.95

Abbreviation	Expansion
Rec. Value	Recommended value (see sources cited for recommendations)
CFA	Confirmatory factory analysis
FOA/EOU	Faithfulness of appropriation versus perceived ease of use
FOA/UFL	Faithfulness of appropriation versus perceived usefulness
Testing	Model testing data set
Confirm.	Model confirmation data set
RMR(C)	Root Mean Square Residual, calculated from the correlation matrix
CFI	Comparative Fit Index
RNI	Relative Noncentrality Index

yses. Thus, if the constructs of interest are truly different, the difference in their correlation between the first (correlation fixed at 1.00) and second (correlation estimated freely) models would be significant. As the difference in the degrees of freedom between the two models is 1 (i.e., the correlation between constructs), a χ^2 difference greater than 3.84 would suggest the two constructs are statistically different ($\alpha = 0.05$).

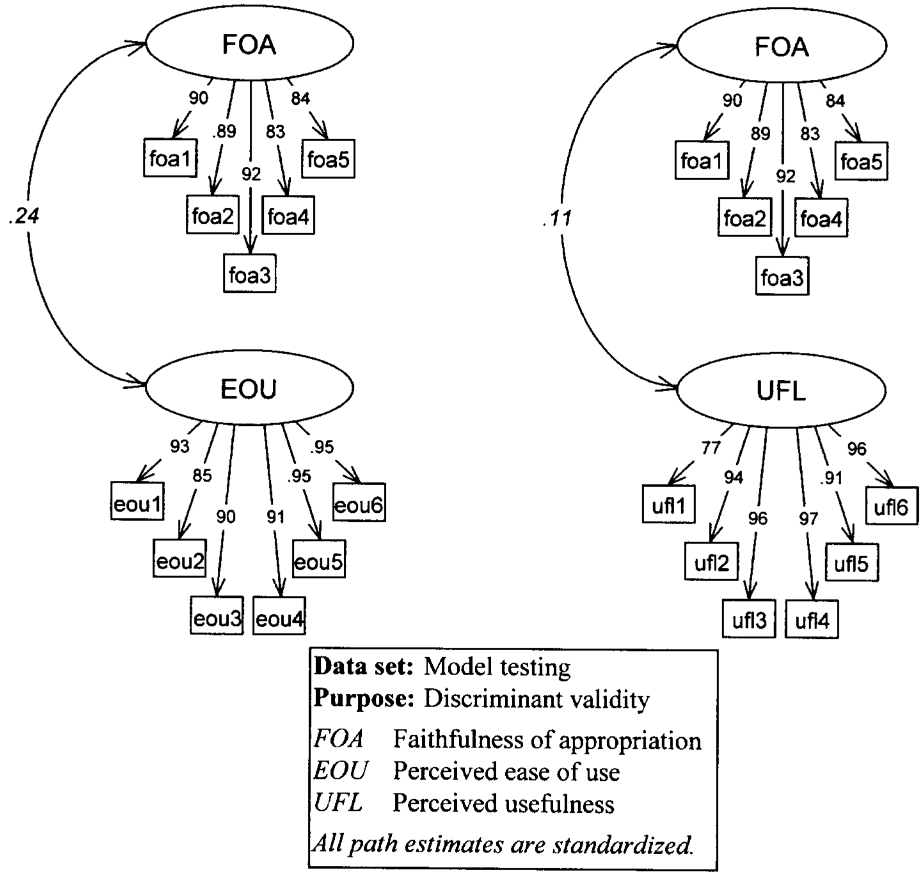
For the comparison between FOA and EOU (see Figure 3), the χ^2 value for the model with the fixed correlation of 1.00 was 498.53 (44 d.f.). When the correlation was allowed to be freely estimated, the χ^2 value (see Table 7) was 109.94 (43 d.f.). The χ^2 difference of 388.59 indicates that the two constructs are indeed distinct. Moreover, the loadings of the FOA items on the FOA construct in the second model (freely estimated correlation between constructs) provide further evidence of the convergent validity of the scale.

Similar conclusions are drawn when comparing FOA and UFL (see Figure 3). The model with fixed correlation of 1.00 resulted in a χ^2 value of 494.58 (44 d.f.), while the freely estimated model yielded a χ^2 value (see Table 7) of 92.97 (43 d.f.). Once again, the difference of 401.61 is much larger than the 3.84 (1 d.f., $\alpha = 0.05$) threshold, indicating that FOA and UFL are two different constructs and thereby establishing the discriminant validity of the former. Its convergent validity is further established by the high loadings of the FOA item measures.

The next step was to test the FOA scale in a more elaborate causal model that included other constructs theoretically related to it in order to assess its validity in a nomological context.

Nomological Validity. The nomological validity of the FOA scale was tested by constructing a causal model that included FOA, EOU, UFL, and two endog-

Figure 3 Freely Estimated Correlations

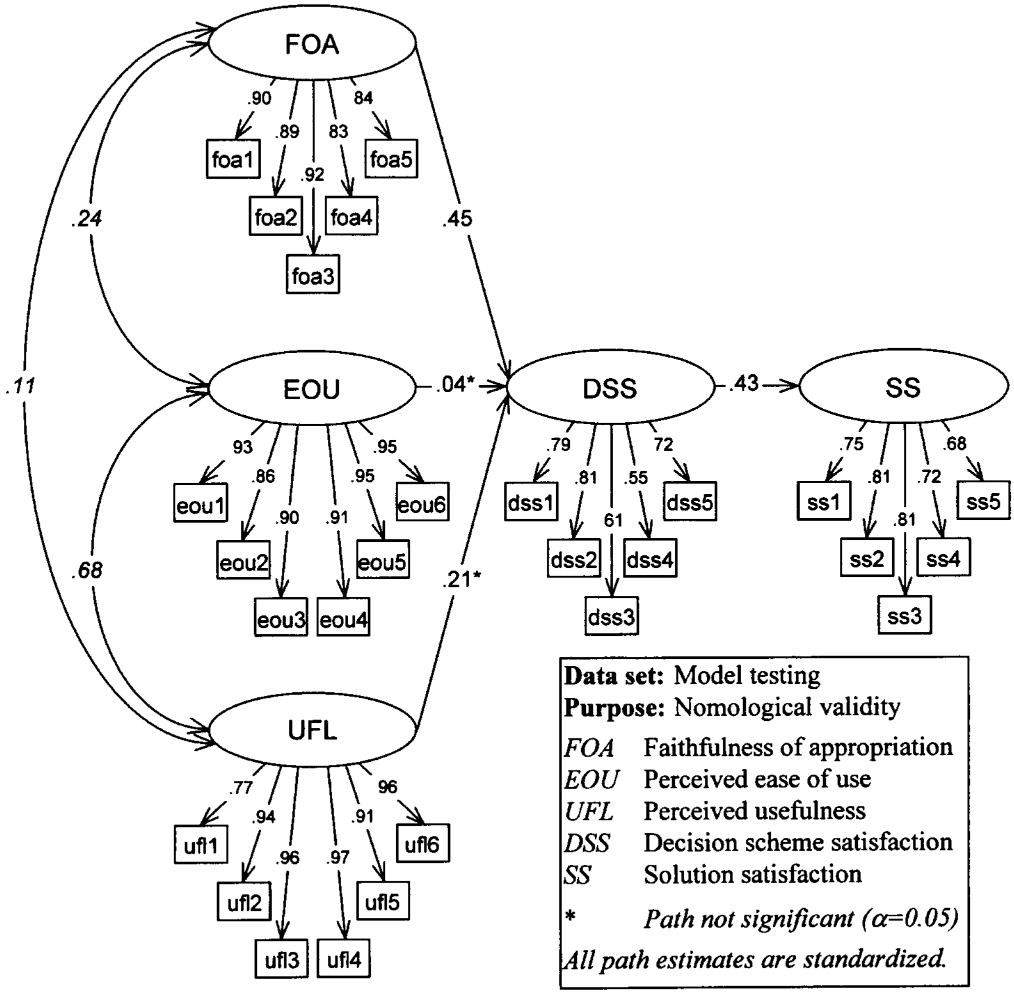


enous constructs, decision scheme satisfaction (DSS) and solution satisfaction (SS) (Green and Taber 1980). The inclusion of these scales met the previously specified criteria that the model be plausible and the constructs be established. As noted, EOU and UFL have been used in previous AST research to represent attitude (Gopal et al. 1993, Sambamurthy and Chin 1994), while DSS and SS have also been used in the AST context and in other EMS studies (see, for example, Watson 1987, Sambamurthy 1989, Gopal 1991, Wheeler et al. 1993); the model, therefore, was considered plausible. The fact that all four constructs had seen considerable use in EMS and IT research and had been subjected to validation procedures were evidence that they were established.

The aim of the analysis was to examine the usefulness of the FOA instrument in predicting meeting satisfaction. Thus, beyond examining how FOA correlated with EOU and UFL, the scale was further analyzed in a model in which all three exogenous factors were hypothesized to impact decision scheme satisfaction. Using temporal logic, DSS was seen as being causally antecedent to SS and modeled as such (see Figure 4).

Weick (1990) proposes that when groups are faced with novel technologies, the use of these technologies is influenced by attempts to make sense of them and their role in task activities. Thus, the attitudes that individuals in groups develop toward technologies such as EMS can influence the outcomes of its use. Hence, attitude is regarded as an important influence on meeting outcomes (Gopal et al. 1993, Sambamurthy and Chin 1994, DeSanctis and Poole 1994). Consistent with this expectation, a strong direct influence of attitudes

Figure 4 Causal Model (Model Testing)



on group performance has been found in EMS research (Gopal et al. 1993). Consequently, EOU and UFL were modeled as having direct influences on DSS.

Faithfulness of appropriation was also modeled as a direct influence on satisfaction based on the reasoning that it results from group members accepting the appropriateness of the EMS method and its potential to help them achieve their desired results. As a consequence of believing they are using the structures provided by the EMS in the "correct" manner (Collins 1992), group participants may be expected to be satisfied with the method.

While nomological validity was assessed using the model testing data set, as in the case of convergent and discriminant validity, it was recognized that the relatively low sample size for the testing data (90) might be insufficient in terms of the more complex model, resulting in a lack of the statistical power necessary to detect and estimate causal links. It was assumed, therefore, that the confirmation data set, with its higher sample size, would provide a better assessment of nomological validity than would the testing data set. Nevertheless, to maintain consistency with the earlier validity analyses and to make an initial assessment of whether the FOA scale performed well in a nomological context, the entire causal model was applied to the testing data.

The overall model fit indices (Table 7) indicate that the model is reasonably consistent with the testing data (Figure 4). The discriminant validity among the exogenous factors is apparent and the largest correlation is that between EOU and UFL. In the case of FOA, the correlation with EOU and UFL are 0.24 and 0.11, respectively. As might be expected, the causal link between DSS and SS is substantial at 0.43.

In terms of the hypothesized links between the exogenous factors (FOA, EOU, and UFL) and DSS, it can be seen that FOA has a strong direct effect on DSS (0.45). While the effect of UFL on DSS was of some magnitude, the path coefficient (0.21) was not found to be significant. This lack of significance can be attributed to the limited statistical power afforded by the low sample size (discussed later).

Also worth noting is the relatively low loading of one of the DSS items (dss4 = 0.55). This number falls below the 0.60 required for minimal reliability and convergent validity. This finding is especially interesting in light of the fact that the DSS scale, which was developed by Green and Taber (1980), has been used extensively in previous EMS research. The questionable performance of one of its items might signify the need to be more careful in adapting scales from other disciplines to the IT environment.

Cross Validation with the Confirmation Data Set. The foregoing assessment of convergent, discriminant, and nomological validity was conducted using the model testing data set. Cross-validation of the results of these analyses was achieved using the confirmation data set. It is possible to conduct this cross-validation using only the causal model used to examine nomological validity, as this model can be used to assess convergent and discriminant validity as well. In fact, as FOA is tested in the context in which it may plausibly be used, the causal model provides the most conservative and therefore the most rigorous test of the different forms of validity. Moreover, if the confirmation data set serves to confirm the validity assessment of the model testing data set, similar results based on the larger sample size of the former alone would be indicative of the robustness of the scale.

There are two reasons why, for the model testing data set, a series of confirmatory factor analyses was necessary to assess convergent and discriminant validity when the causal model used to assess nomological validity would have been enough to establish all three kinds of validity. In view of the objective of this paper to contribute to the discipline a viable approach to comprehensive scale development, these reasons deserve articulation.

The first reason for the battery of tests was to examine the FOA scale in detail to identify its characteristics; the construct was examined in isolation first, then in relation to each of two other constructs, and finally in a nomological network. Consequently, its individual characteristics and its specific relationships with other constructs that also influence outcomes were revealed. Once these characteristics were evident, cross-validation did not call for any further understanding in this regard; rather, the purpose of cross-validation was to determine whether convergent, discriminant and nomological validity for the scale existed even when extended to a different data sample.

The second reason for using a series of tests on the model testing data set relates to the fact that if the final (nomological) model had been found wanting, each of the other, less rigorous, tests for convergent and discriminatory validity would have had to have been conducted in order to identify the root cause for the model failure (i.e., is it due to poor measurement of the underlying construct or measurement overlap with other constructs?). Rather than initially test the FOA construct in a complete rigorous model, a systematic approach intended to establish each form of validity through initial analyses was adopted. Thus, it was *necessary* for the scale to "pass" each of these tests for validity to be established. Any failure would automatically result in a failure for the full nomological test. Yet, passing each individual test is not *sufficient* to prove validity. Only the results of the final causal model permit sufficient confidence to be placed in the validity of the scale.

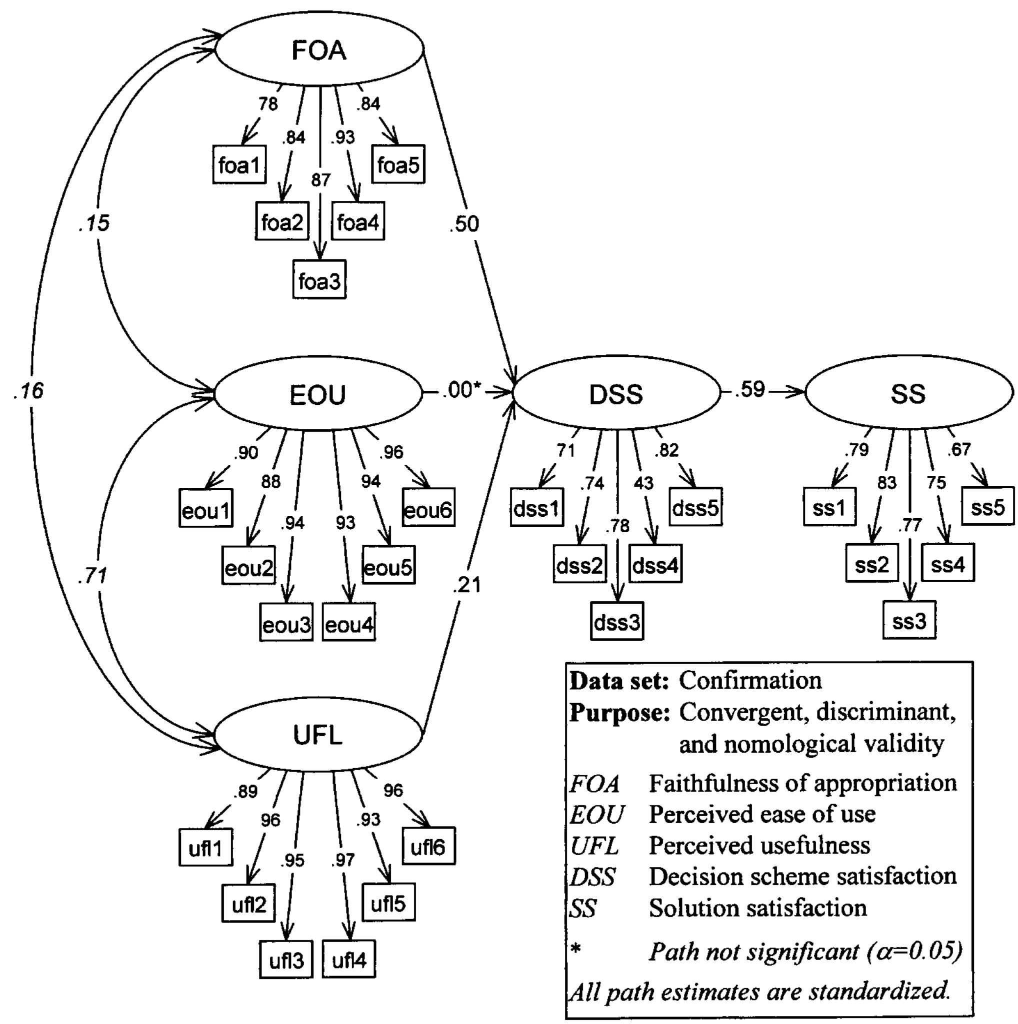
Once the series of tests was completed on the model testing data set, a simple examination of the goodness of fit measures (Table 7) and the path coefficients in the causal model for the confirmation data set (Figure 5) provided a rigorous assessment of the scale's validity. As is evident from the analysis of the confirmation data set, the convergent validity of the FOA scale is high (the lowest item loading is 0.78), discriminant validity is established (the correlation with EOU and UFL are 0.15 and 0.16, respectively), and the scale performs well in predicting process satisfaction. The goodness of fit measures are uniformly high. The overall effectiveness of the faithfulness scale, therefore, is confirmed.

It is worth noting that the DSS item that loaded poorly in the model testing data set (dss4) had an even poorer loading in the confirmation data set (0.43). This finding provides further evidence for the need to reassess the use of this scale in EMS research. Such reassessment, however, is considered beyond the scope of the present paper.

Threats to Validity
While considerable effort was expended on the validation of the FOA scale, the procedure is incomplete without the consideration of certain remaining threats to its validity. The specific threats are the question concerning the power of the models in phase 3 to detect significant effects, the potential for a biased reading of the FOA items due to the polarity of their wording, and the potential for group level influences on individual responses.

Power Analysis. The fact that the path from UFL to DSS is significant in the confirmation data set is

Figure 5 Causal Model (Confirmation)



worth noting. As suggested earlier, the lack of significance of this path using the testing data set may well have been attributable to the lack of statistical power resulting from the lower sample size. When the model was tested on the larger sample size of the confirmation data set, there was sufficient power to detect the significant influence of usefulness. Nonetheless, an explicit test for power is needed to determine whether the confirmation data set is capable of detecting other model misspecifications. Insufficient power would suggest that the resulting high goodness of fit for this data set can be attributed to an inability to detect misspecified models rather than to a confirmation of the fit between model and data.

To determine whether there was indeed a power effect between the two data sets, the procedure outlined by Saris and Stronkhorst (1984, p. 205) was applied. The approach is to take the original model tested (H_o) and modify it to create an alternative model (H_a) in terms of additional parameter(s) (e.g., additional causal links via structural paths between constructs or factor loadings) with specific effect size values. An implied population covariance matrix is generated from this alternative model and then used on the original model. In essence, the wrong model is tested with full awareness concerning the correct model paths and loadings. If the correct model, H_a, were run on this covariance matrix, a perfect fit would be obtained (i.e., χ^2 equal to zero). However, using H_o instead results in having the likelihood ratio value λ obtained from this analysis (i.e., the χ^2 statistic from standard covariance analysis programs) approximating a noncentral χ^2 distribution with degrees of freedom equal to the number of additional parameters included to create H_a. Given λ, the α level, and the degrees of freedom, the power can be determined using the noncentral χ^2 tables.

Table 8 shows the results of testing the power of the model testing and model confirmation data sets to detect three different misspecified models. The first test consists of creating an alternative model where a standardized path of 0.25 from EOU to DSS is fixed. The model is then tested for its power in detecting an incorrect model in which that path is excluded for estimation. The second test follows a similar procedure, but in this case a standardized cross-loading of 0.50 from the factor EOU to the indicator foa5 is specified.

Table 8 Results of Power Analysis (Phase 3)

		Testing		Confirmation	
Model	Effect Size	λ	Power	λ	Power
EOU-DSS	0.25	3.46	0.46	9 19	0.86
EOU-foa5	0.50	66.20	1.00	162.51	1.00
Correlated Errors	0.25	1 60	0.24	9.30	0.86

Calculated for $\alpha = 0.05$, degrees of freedom = 1 for all cases.

A model is then run where that loading is not included for estimation. The final test is to correlate the errors associated with two indicators of DSS (dss1 and dss2) and assess whether model misspecification can be detected when the correlation is not included.

The three tests were created to assess three different potential problem areas. The first test demonstrates that the zero path estimate for the structural path of EOU to DSS in the confirmation data set is probably correct since an effect size of 0.25 is likely to be detected. The second test examines the ability to detect a poor measure that taps into more than one construct at a level approaching that necessary to suggest convergent validity (e.g., 0.50). It shows that a measure that taps strongly into another construct beyond the one for which it was created will be detected. Finally, the third test examines the ability to detect additional latent variables. Correlated errors between two indicators can be interpreted as being influenced by one or more additional latent variables beyond the one modeled. A correlation of some magnitude (e.g., 0.25) between errors could indicate a shared method effect, latent trait, or any general influence that need not be unidimensional nor follow a congeneric structure. In all three tests, the power of the confirmation set to detect misspecificaton was above the generally accepted level of 0.80 (Cohen 1988), whereas that of the testing data set fell short in two of the tests. In other words, it is guaranteed that the confirmation data set allows the detection of model misspecifications of the types tested at least 80% of the time.

An alternative approach recently presented by MacCallum et al. (1996) provides a means for calculating power without specifying a specific alternative model. Instead, power estimations are calculated in terms of a null and alternative value of the root-mean-square error of approximation (RMSEA) fit index. Depending on the selection of these indices, power calculations can be made for testing hypotheses such as "close" or "not-close" fits as well as "exact fit." An exact fit assesses the power of detecting a difference for the RMSEA (suggested at 0.05) when it is actually 0. Yet, this assumption of obtaining a perfect fit of 0 has been argued as "not particularly useful in practice because the hypothesis being tested is implausible and is not empirically interesting and because the test will result in rejection of good models when N is large" (MacCallum et al. 1996, p. 132). Instead, MacCallum et al. recommend a test of not-close fit. This test examines the power of rejecting the hypothesis that the RMSEA is greater than 0.05 (i.e., not a good fit) when it actually is quite good (suggested value of 0.01). Interestingly enough, using our testing data with sample size of 90 and 317 degrees of freedom, the power was 0.822 and 0.742 for the exact and not-close fits, respectively. Thus, at a global level, even our test data had reasonable power to detect differences in overall model fit. A major reason is the large number of degrees of freedom in the model which compensate for the smaller sample size. For the confirmatory data set with the larger sample size, the power is essentially one.

Wording Effects. A question that might arise is whether positive wording effects may have led to the exclusion of the items eliminated during Phase 2.[4] Specifically, items 1, 2, and 7 are worded in the positive direction, whereas the remaining items are in the negative direction. At least two issues are of concern here. The first is whether the components analyses in these two phases inadvertently eliminated good measures of faithfulness because they were positively worded. The second and, from our perspective, more critical issue is whether there is method bias in either the positive or negative directions. In other words, to what extent does the orientation in wording increase the convergent validity (i.e., loadings)?

Two post-hoc confirmatory factor analyses were performed on the data from the first two experiments (this was not possible with the third experiment which used only the negatively oriented measures). For each data set, a three-factor model consisting of a single FOA construct and two method factors (positive and negative effects) was analyzed. While all indicators were modeled to load on the FOA factor, a positive method factor that influenced measures foa1, foa2, and foa7 and a negative method factor that influenced the other items were included. Overall model fits were quite strong for both the first (e.g., 0.930 and 0.933 for Bentler's CFI and Bollen's delta, respectively) and the second (e.g., 0.957 and 0.957 for Bentler's CFI and Bollen's delta, respectively) experiment.

The results indicate that the positively oriented measures consistently had a strong method bias whereas the negative ones varied. For the first experiment, the positive measures were all significant, whereas foa4 and foa6 were significant among the negatively phrased measures. In the second experiment, only the positively phrased measures had a significant method bias (i.e., significant loadings of 0.73, 0.87, and 0.46 for measures foa1, foa2, and foa7, respectively). Therefore, the elimination of the positively worded items at an early stage also served to eliminate the method bias in the original scale.

Group Effects. Another question that may arise is the possibility of group effects, another potential method effect that may reduce the generalizability of the study.[5] The question concerns whether the fact that subjects working together in groups represent an effect that may bias the individual responses in the study. While each subject's response to the questionnaire was *independent* of other subjects (i.e., subjects did not compare their responses with each other), each subject did share a group experience with other members in a meeting. Yet, due to the small group size, it is unlikely that any particular method effect was captured. The large number of groups would wash out any biases that may occur in any particular group session.

To confirm that group effects were not substantial, another post hoc analysis was performed, this time at the group level. The testing and confirmation data sets

[4] We thank the Associate Editor and Reviewer 3 for bringing this issue to our attention.

[5] We thank the Associate Editor for alerting us to this issue.

from the third experiment were combined to increase sample size. An average for each measure was obtained for each group, and the resulting 70 group cases were analyzed using the nomological model used earlier (i.e., Figures 4 and 5). The loadings were similar to the previous results (i.e., loadings above 0.80 except for the DSS and SS measures). The FOA to DSS path was 0.65, while paths from EOU and UFL were −0.05 and 0.10, respectively. A bootstrap analysis (2,050 resamples) was conducted to test for the significance of these paths and loadings. All the structural paths except those from EOU and UFL to DSS were significant ($p < 0.01$). The hypothesis that group effects biased the results, therefore, can be rejected.

Discussion and Conclusions

The primary aim of this paper was to report the development of a convenient scale to measure the faithfulness of appropriation of advanced information technologies. A multi-stage scale development and validation procedure was used to create a scale consisting of five items. Some methodological and theoretical results of the scale development process should be noted.

On the methodological front, the need for rigorous instrument validation beyond principal components factor analysis was demonstrated. Items that would have been acceptable when subjected to principal components analysis were less so when analyzed using the more rigorous CFA approach. Also, a detailed scale development and validation *procedure* was presented, the result of which was the FOA scale.

A theoretical finding concerns the importance of faithfulness of appropriation in understanding participant satisfaction with the use of the EMS (as noted earlier, we use this term to refer to the meeting process in its entirety). The magnitude of the path coefficient between faithfulness and decision scheme satisfaction (Figures 4 and 5) is evidence of the importance of faithfulness in predicting group outcomes. In other words, individuals' assessment of the degree to which their group uses the EMS in the "appropriate" manner strongly influences their subsequent satisfaction with the EMS.

The distinctiveness of faithfulness and satisfaction, however, should also be noted. The distinction warrants discussion because of the potential conflation of the idea of *using* a system in the manner specified with the idea of *being satisfied with using* it.[6] The question that might arise is whether there is a potential negative *evaluative* connotation of the manner in which the final FOA items were worded; could, for example, "inappropriate use" be translated by the user as unsatisfactory use. There is statistical evidence that such was not the case: the path coefficient from FOA to DSS in the confirmation data set (0.50), though high enough to suggest a strong causal relationship, was low enough to indicate that the two underlying constructs are distinct. More importantly, there is theoretical reason to reject this hypothesis. None of the five items retained in the scale refers to the *consequences* of using the system in the way in which it was used. All the items refer to the act of usage itself. Satisfaction, by its very nature, refers specifically to a consequence or outcome of use rather than the use itself.

An illustration may help reveal how the two constructs differ. Many students of IT are familiar with the practice, especially in the 1980s, of using a spreadsheet package for word processing tasks. The individuals who used spreadsheets in this manner were often informed that they were using the software "incorrectly" and almost certainly realized themselves that they were not using it as intended, recognizing, therefore, their unfaithful appropriation of the software. Nevertheless, many such individuals continued this manner of use and often grew quite comfortable doing so; satisfaction with the manner in which they used the software and with the outcomes of such use was, arguably, quite high. EMS users, similarly, might acknowledge that they have used the system "inappropriately" or "improperly" but might, nevertheless, be quite satisfied with having used it in this manner or with the outcomes of its use. Construed in this manner, faithfulness clearly differs from satisfaction.

The intent of the scale development procedure was to provide a measurement method that would *complement* rather than *supplant* external methods such as the microcoding approach advocated by the originators of

[6] We thank Reviewer 3 for pointing out the possibility of such an interpretation.

adaptive structuration theory (DeSanctis and Poole 1994). In view of the ambitious scope of application of the theory envisaged by DeSanctis and Poole and given the rapid rate at which advanced information technologies are being developed and deployed, the availability of a short five-item scale to measure faithfulness should prove useful to both researchers and implementors of new information technologies. However, it is important to consider the manner in which the scale relates to external approaches (see Figure 1), the ways in which it might be used in conjunction with such approaches, and the manner in which it can be extended for use with other advanced information technologies.

It was noted earlier that the failure of internal (e.g., the FOA scale) and external (e.g., microcoding) measures of faithfulness to converge should not be construed as either type of measure having invalidated the other. More likely, it indicates a differential understanding of the spirit of the system between the user and the researcher or other observer.

Nevertheless, a communality of understanding is more likely than a lack of it. In situations when a system is new to a user community, there is usually explicit communication of its aims and objectives (its spirit). Moreover, to the extent that the use of a construct such as faithfulness reflects a degree of sensitivity to the theory of adaptive structuration, researchers in such situations can be expected to have attended, in some fashion, to the task of ensuring such adequate communication of the spirit. Even in "live" organizational situations in which a system is smoothly incorporated into ongoing group work in a manner that precludes the possibility of intervening in order to communicate its spirit, individuals can be expected to have some common organizational reference point that informs their understanding of the intentions underlying the use of the system. Consequently, a common understanding of a spirit can be assumed, as can convergence between internal and external conceptions of spirit. The scale, therefore, can be unproblematically applied in the majority of situations in which a new system is introduced without having to resort to the more demanding activity of microcoding.

Microcoding and other external measures, however, may be useful in assessing whether the spirit as understood by the implementor is indeed similar to the spirit as understood by the user, especially in situations where such differences might be expected. For researchers to engage in microcoding, of course, all sessions will need to be videotaped, as was the case for the third experiment reported in this paper. In field settings, however, videotaping is not always a viable option for reasons such as resources being unavailable or group members failing to consent to being videotaped. In the latter case, an important reason for failure to provide consent may relate to the confidentiality issue—organizational users may consider the information exchanged during a meeting to be sensitive and may therefore consider a videotaped record of a meeting to represent a potential breach of confidentiality. Under such circumstances, researchers would be constrained to measure faithfulness by means of the scale rather than through microcoding.

As the use of a technology evolves, so do user interpretations of its spirit. The subjective understanding of spirit becomes increasingly unique to the individual, thereby making the assessment of faithfulness against an objectively defined spirit less meaningful. A scale, therefore, would be the most effective means to measure faithfulness, by having users assess their own usage in relation to the spirit defined subjectively (see Figure 1). The measurement of faithfulness in such situations of ongoing use is, however, unlikely to be well served by the FOA scale in its present form. As the spirit is internalized and subjectified, it will have to be decoupled from the design intent. Even though the spirit might initially reflect the designers' intent, it will eventually be inaccurate to define it solely in these terms—it will have evolved to represent considerably more than such initial intent. Careful scrutiny of the scale, however, reveals ways in which faithfulness can be captured relative to this evolving internalized spirit: only two of the five items were actually anchored in design intent (e.g., "The developers of the EMS would disagree with how our group used the system"); the other three items called for faithfulness to be evaluated with respect to a "general" spirit (e.g., "Our group probably used the EMS improperly"). To be sure, reference to the design intent in two of the items is crucial to the representation of and reference to spirit in all five items in situations where the system is new to the user community. However, the scale can be effectively modified to measure faithfulness in ongoing use situations by removing explicit references to design intent. We will leave the task of modifying the scale and testing it to future research.

Even though its diachronic use may be called into question, the applicability of the scale to other new advanced information technologies should be high. Generalizability, in this regard, is actually enhanced by the measurement of faithfulness from the individual's standpoint and by the anchoring of the scale in design intent. Beyond the argument that spirit is encountered only at the individual level, measurement at this level means that faithfulness can be assessed for advanced information technologies for which use is not observed or even unobservable. Therefore, use in natural settings such as workplaces can also be monitored for faithfulness. Moreover, minor modifications to the items will permit their use in relation to individual rather than group information technologies (an example is presented in Table 9). The anchoring of the scale in design intent will give individual users access to the aims and objectives of any new technology, even when it is difficult to explicitly communicate spirit (such as in 'live' organizational settings): design intent will serve as a representation of the spirit of the system.

The applicability of the scale in organizational settings, however, clearly should be tested in future research. It is in organizations that advanced information technologies are being introduced most rapidly and it is in those settings that the scale could prove most fruitful.

This is not to say, however, that the use of student subjects in the experiments reported in this paper serves to compromise the external validity of the scale, as is often assumed in research that utilizes student subjects (Gordon et al. 1986, 1987). When students are expected to assume the roles of organizational members or to make judgments concerning issues that organizational members would make, the use of student subjects could prove problematic, given their clearly different status. However, when students are expected to assume roles that do not remove them from their student frame of reference, their responses may be expected to be valid with respect to the objective of the study (DeSanctis 1989). In each of the experiments reported in this paper, students were asked to perform tasks in which they participated as students rather than in any other roles. Moreover, monetary rewards for performance were established in all three experiments, hence providing extrinsic motivation to perform. In addition, the task used in each experiment was relevant to a student population, thereby providing intrinsic motivation. Consequently, the use of student subjects is not considered problematic in this exercise.

Table 9 Example of FOA Items Adapted for Other Technologies

The developers of [the technology] would disagree with how I used it.
I probably used [the technology] improperly.
The original developers of [the technology] would view my use of it as inappropriate.
I failed to use [the technology] as it should have been used.
I did not use [the technology] in the most appropriate fashion.

Finally, it was noted earlier that the intent of this paper was to develop the FOA scale and not to test AST. Consequently, the model used to test the validity of the scale was intended only as a plausible rather than an "accurate" representation of AST. Beyond the fact that space limitations constrain our ability to develop the scale *and* test the theory, any test of the theory would have to include the other constructs implicated in it: consensus on appropriation and attitude. While attitude has been operationalized elsewhere (Sambamurthy 1989), a convenient scale to measure consensus on appropriation needs to be developed. Only when a complete suite of measures is available can the propositions of AST be conveniently tested in future research.

The aim of this project was not to provide a means to perfectly measure faithfulness of appropriation. Just as DeSanctis and Poole (1994, p. 141) point out that their "interest is in describing appropriation processes with sufficient refinement so that we can gain meaningful (though not perfect) insight into the connection between technology and action," the intent in this paper has been to put together a scale that taps into the faithfulness construct enough to provide insights into the appropriation process. Even though rigorous validation procedures were used to develop the scale, researchers using it should not expect to see uniformly high loadings and path coefficients every time it is used. Any attempt to assess human interpretive schemes must take into account the inherent variability of individual and social processes.[7]

[7]The research reported in this paper was funded in part by grants from the Alberta Energy Company and the Social Sciences and Humanities Research Council. We would like to express our appreciation for the part played by Barry Chute in conducting many of the experimental sessions reported in this paper. We would also like to acknowledge the detailed and insightful comments provided by the associate editor and three anonymous reviewers throughout the review process.

References

Arbuckle, J., *AMOS 3.1 Documentation Package*, Department of Psychology, Temple University, Philadelphia, PA, 1993.

Bagozzi, R. P. and Y. Yi, "On the Evaluation of Structural Equation Models," *J. Acad. Marketing Sci.*, 16 (1988), 74–94.

Beauclair, R. A., "An Experimental Study of the Effects of Group Decision Support System Process Support Applications on Small Group Decision Making," Unpublished Doctoral Dissertation, Indiana University, Bloomington, IN 1987.

Bentler, P. M., "Comparative Fit Indexes in Structural Models," *Psychological Bulletin*, 107 (1990), 238–246.

——, and D. G. Bonett, "Significance Tests and Goodness of Fit in the Analysis of Covariance Structures," *Psychological Bulletin*, 88 (1980), 588–606.

Bollen, K. A., "Sample Size and Bentler and Bonett's NonNormed Fit Index," *Psychometrika*, 51 (1986), 375–377.

——, *Structural Equations Modeling with Latent Variables*, Wiley, New York, 1989.

Bryant, C. G. A., "Sociology Without Philosophy? The Case of Giddens' Structuration Theory," *Sociological Theory*, 10, 2 (1992), 137–149.

Chin, W. W. and P. A. Todd, "On the Use, Usefulness, and Ease of Use of Structural Equation Modeling in MIS Research: A Note of Caution," *MIS Quarterly*, 19, 2 (1995), 237–246.

Clapper, D. L. and P. Prasad, "The Rationalization of the Organizational Meeting: Implications of Group Support Systems for Power, Symbolism and Face-Work," *Proc. Fourteenth Annual International Conf. on Information Systems*, ICIS, Orlando, FL, 1993, 321–329.

Cohen, J., *Statistical Power Analysis for the Behavioral Sciences (Second Ed.)*, Lawrence Erlbaum Associates, Hillsdale, NJ, 1988.

Collins, R., *Sociological Insight: An Introduction to Non-Obvious Sociology*, Second Ed., Oxford University Press, New York, 1992.

Dalkey, N. and O. Halmer, "An Experimental Application of the Delphi Method to the Use of Experts," *Management Sci.*, 6 (1963), 458–467.

Davis, F. D., "Perceived Usefulness, Perceived Ease Of Use, and User Acceptance of Information Technology," *MIS Quarterly*, 13, 3 (1989), 319–339.

Dennis, A. R. and R. B. Gallupe, "A History of Group Support Systems Empirical Research: Lessons Learned And Future Directions," in L. M. Jessup and J. S. Valacich (Eds.), *Group Support Systems: New Perspectives*, Macmillan, New York, 1993, 59–77.

——, J. F. George, L. M. Jessup, J. F. Nunamaker, and D. R. Vogel, "Information Technology to Support Electronic Meetings," *MIS Quarterly* 12, 4 (1988), 591–624.

DeSanctis, G., "Small Group Research in Information Systems: Theory and Method," *The Information Systems Research Challenge: Experimental Research Methods*, 2 (1989), 53–82.

——, "Shifting Foundations in Group Support Systems Research," in L. M. Jessup and J. S. Valacich (Eds.), *Group Support Systems: New Perspectives*, Macmillan, New York, 1993, 97–111.

——, M. D'Onofrio, V. Sambamurthy, and M. S. Poole, "Comprehensiveness and Restrictiveness in Group Decision Heuristics: Effects of Computer Support on Consensus Decision Making," *Proc. Tenth Annual International Conf. on Information Systems*, 1989, 131–140.

—— and R. B. Gallupe, "A Foundation for the Study of Group Decision Support Systems," *Management Sci.*, 33, 5 (1987), 589–609.

—— and M. S. Poole, "Capturing the Complexity of Advanced Technology Use: Adaptive Structuration Theory," *Organization Sci.*, 5, 2 (1994), 121–147.

——, J. R. Snyder, and M. S. Poole, "The Meaning of the Interface: A Functional and Holistic Evaluation of a Meeting Software System," *Decision Support Systems*, 11 (1994), 319–335.

Drucker, P. F., "The Coming of the New Organization," *Harvard Business Rev.*, 66, 1 (1988), 45–53.

Giddens, A., *Central Problems in Social Theory: Action, Structure, and Contradiction in Social Analysis*, Macmillan, London, 1979.

——, *The Constitution of Society: Outline of the Theory of Structuration*, University of California Press, Berkeley, CA, 1984.

——, *New Rules of Sociological Method: A Positive Critique of Interpretive Sociologies*, Second Ed., Polity Press, Cambridge, U. K., 1993.

Gopal, A., *The Effects of Technology Level and Task Type on Group Outcomes in a Group Decision Support System Environment*, Unpublished Doctoral Dissertation, University of Georgia, 1991.

——, R. P. Bostrom, and W. W. Chin, "Applying Adaptive Structuration Theory to Investigate the Process of Group Support Systems Use," *J. Management Information Systems*, 9, 3 (1993), 45–69.

Gordon, M. E., L. A. Slade, and N. Schmitt, "The 'Science of the Sophomore' Revisited: From Conjecture to Empiricism," *Acad. Management Rev.*, 11, 1 (1986), 191–207.

——, ——, and —— "Student Guinea Pigs: Porcine Predictors and Particularistic Phenomena," *Acad. Management Rev.*, 12, 1 (1987), 160–163.

Green, S. G. and T. D. Taber, "The Effects of Three Social Decision Schemes on Decision Group Process," *Organizational Behavior and Human Performance*, 25 (1980), 97–106.

Hackman, J. R. and R. E. Kaplan, "Interventions into Group Processes: An Approach to Improving the Effectiveness of Groups," *Decision Sci.*, 5 (1974), 459–480.

Hu, L.-T., and P. M. Bentler, "Evaluating Model Fit," in R. H. Hoyle (Ed.), *Structural Equation Modeling: Concepts, Issues, and Applications*, Thousand Oaks, CA, 1995, 76–99.

Huber, G. P., "A Theory of the Effects of Advanced Information Technologies on Organizational Design, Intelligence, and Decision-Making," *Acad. Management Rev.*, 5, 1 (1990), 47–71.

Janis, I. L., *Groupthink*, Houghton Mifflin, Boston, MA, 1982.

Long, J. S., *Covariance Structure Models: An Introduction to LISREL*, Sage University Paper Number 07-034, Sage, Newbury Park, CA, 1983.

MacCallum, R. C., M. W. Browne, and H. M. Sugawara, "Power Analysis and Determination of Sample Size for Covariance Structure Modeling," *Psychological Methods*, 1, 2 (1996), 130–149.

Maier, N. and R. Maier, "An Experimental Test of the Effects of 'Developmental' vs. 'Free' Discussions on the Quality of Group Decisions," *J. Applied Psychology*, 41 (1957), 320–333.

McDonald, R. P. and H. W. Marsh, "Choosing a Multivariate Model: Noncentrality and Goodness of Fit," *Psychological Bulletin*, 107, 2 (1990), 247–255.

Nunamaker, J. F., Jr., A. R. Dennis, J. S. Valacich, D. R. Vogel, and J. F. George, "Electronic Meeting Systems to Support Group Work," *Comm. ACM*, 34, 7 (1991), 40–61.

Orlikowski, W. J., "The Duality of Technology: Rethinking the Concept of Technology in Organizations," *Organization Sci.*, 3, 3 (1992), 398–427.

—— and D. Robey, "Information Technology and the Structuring of Organizations," *Information Systems Res.*, 2, 2 (1991), 143–169.

Pentland, B., "Organizing Moves in Software Support Hot Lines," *Admin. Sci. Quarterly*, 37 (1992), 527–548.

Poole, M. S. and G. DeSanctis, "Understanding the Use of Group Decision Support Systems: The Theory of Adaptive Structuration," in J. Fulk and C. Steinfield (Eds.), *Organizations and Communication Technology*, Sage, Newbury Park, CA, 1990, 173–193.

—— and ——, "Microlevel Structuration in Computer-Supported Group Decision Making," *Human Communication Res.*, 19, 1 (1992), 5–49.

Rao, V. S. and S. L. Jarvenpaa, "Computer Support of Groups: Theory-Based Models for GDSS Research," *Management Sci.*, 37, 10 (1991), 1347–1362.

Ranson, S., B. Hinings, and R. Greenwood, "The Structuring of Organizational Structures," *Admin. Sci. Quarterly*, 25 (1980), 1–17.

Riley, P., "A Structurationist Account of Political Culture," *Admin. Sci. Quarterly*, 28 (1983), 414–437.

Sambamurthy, V., "Supporting Group Performance During Stakeholder Analysis: The Effects of Alternative Decision-based Designs," Unpublished Doctoral Dissertation, University of Minnesota, Minneapolis, MN, 1989.

—— and W. W. Chin, "The Effects of Group Attitudes Toward GDSS Designs on the Decision-Making Performance of Computer-Supported Groups," *Decision Sci.*, 25, 2 (1994), 215–241.

—— and M. S. Poole, "The Effects of Variations in Capabilities of GDSS Designs on Management of Cognitive Conflict in Groups," *Information Systems Res.*, 3, 3 (1992), 224–251.

Saris, W. E. and L. H. Stronkhorst, *Causal Modeling In Nonexperimental Research: An Introduction To The LISREL Approach*, Sociometric Research Foundation, Amsterdam, 1984.

Silver, M. S., "Decision Support Systems: Directed and Nondirected Change," *Information Systems Res.*, 1, 1 (1990), 47–70.

Sproull, L. S. and P. S. Goodman, "Technology and Organizations: Integration and Opportunities," in P. S. Goodman, L. S. Sproull, and Associates, *Technology and Organizations*, Jossey-Bass, San Francisco, CA, 1990, 254–266.

Stasser, G., "Pooling of Unshared Information During Group Discussions," in S. Worchel, W. Wood, and J. A. Simpson (Eds.), *Group Process and Productivity*, Sage, Newbury Park, CA, 1992, 48–67.

Tucker, L. R. and C. Lewis, "A Reliability Coefficient for Maximum Likelihood Factor Analysis," *Psychometrika*, 38 (1973), 1–10.

Van de Ven, A. and A. Delbecq, "Nominal Versus Interacting Group Processes for Committee Decision Making," *Acad. Management J.*, 14 (1974), 203–213.

Watson, R. T., Unpublished Doctoral Dissertation, University of Minnesota, *"A Study of Group Decision Support System Use in Three- and Four-Person Groups for a Preference Allocation Decision,"* Minneapolis, MN, 1987.

Weick, K. E., "Technology as Equivoque: Sensemaking in New Technologies," in P. S. Goodman, L. S. Sproull, and Associates (Eds.), *Technology and Organizations*, Jossey-Bass, San Francisco, CA, 1990, 1–44.

Wheaton, B., B. Muthen, D. Alwin, and G. Summers, "Assessing Reliability and Stability in Panel Models," in D. Heise (Ed.), *Sociological Methodology 1977*, Jossey-Bass, San Francisco, CA, 1977, 84–136.

Wheeler, B. C. and B. M. Mennecke, *"The School of Business Policy Task Manual,"* Indiana University Working Paper 92-524, Bloomington, 1992.

——, —— and J. N. Scudder, "Restrictive Group Support Systems as a Source of Process Structure for High and Low Procedural Order Groups," *Small Group Res.*, 24, 4 (1993), 504–522.

Whittington, R., "Putting Giddens into Action: Social Systems and Managerial Agency," *J. Management Studies*, 29, 6 (1992), 695–712.

Marshall Scott Poole, Associate Editor. This paper was received on July 17, 1995 has been with the authors 4 months for 2 revisions.