



HAL
open science

Scale adaptation of MIS measures and their implications on MIS scholarship

Claudio Vitari, Jean-charles Pillet

► **To cite this version:**

Claudio Vitari, Jean-charles Pillet. Scale adaptation of MIS measures and their implications on MIS scholarship. Colloque de l'AIM, May 2023, Dijon (Bourgogne), France. hal-04118246

HAL Id: hal-04118246

<https://amu.hal.science/hal-04118246>

Submitted on 6 Jun 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Adaptation d'échelles de mesure en Systèmes d'Information et leurs implications pratiques

Claudio Vitari, Aix Marseille Univ, CERGAM, FEG, Aix-en-Provence, France

Jean-Charles Pillet, Information Management Department, TBS Business School, Toulouse, France

Résumé :

L'adaptation d'échelles fait référence au processus de changement d'un élément d'une échelle, tel que : modifier la formulation des phrases, revoir le type et le nombre de points d'ancrage, ou modifier la taille d'une échelle. L'adaptation des échelles est une pratique répandue en Systèmes d'Information (SI) depuis ses débuts et est particulièrement curant dans la recherche en SI en raison de la nécessité de maintenir un alignement étroit entre les concepts et leurs mesures, d'une part, et les contextes d'usage des SI qui évoluent rapidement. Dans cet article, nous remettons en question le présupposé méthodologique établi dans notre communauté SI, selon lequel l'adaptation d'échelles est une pratique inoffensive qui conduit finalement à l'amélioration des échelles existantes. Nous explorons les pratiques d'adaptation d'échelles dans la discipline SI, en examinant les échelles de mesure de la confiance, et nous contribuons à la conceptualisation de ces pratiques et nous avançons quelques recommandations concernant les pratiques identifiées.

Mots clés :

développement d'échelles ; adaptation d'échelles ; adaptation en cascade ; mixologie ; échelles de confiance

Scale adaptation of MIS measures and their implications on MIS scholarship

Claudio Vitari, Aix Marseille Univ, CERGAM, FEG, Aix-en-Provence, France

Jean-Charles Pillet, Information Management Department, TBS Business School, Toulouse,
France

Abstract :

Scale adaptation refers to the process of changing something about a scale, such as: alter item wording, modify the target referent, revisit the type and number of anchor points, or modify the size of a scale. Scale adaptation has been a widespread practice in Management Information Systems (MIS) research since its very early days and is especially prevalent in MIS research due to the need to maintain a close alignment between a construct's definition, its measurement, and the rapidly evolving contemporary Information Technology or behavioural contexts. In this paper, we challenge the established methodological assumption that scale adaptation is a harmless practice that ultimately leads to the betterment of existing scales. We explore the practices of scale adaptation in the MIS discipline, looking at the Trust scales, we contribute to the conceptualization of these practices and we provide some recommendations regarding the identified ones.

Keywords :

scale development; scale adaptation; cascading adaptation; mixology; trust scales

Scale adaptation of MIS measures and their implications on MIS scholarship

1. Introduction

Scale adaptation refers to the process of changing something about a scale, such as: alter item wording, modify the target referent (e.g., different IT), revisit the type and number of anchor points, or modify the size of a scale (Heggstad et al., 2019). Scale adaptation has been a widespread practice in Management Information Systems (IS) research since its very early days. MIS academics have mainly espoused a positivist research tradition (Lee, 1999) and quantitative research methods (Straub et al., 2005). Quantitative positivist researchers typically identify constructs, operationalize them through measurement variables and then articulate relationships among the identified constructs (Im & Wang, 2007). These measurement variables are developed in the form of scales to increase their reliability (Recker, 2021). Then, these scales are routinely adapted to overcome the constraints that MIS researchers face in the course of their work. Scale adaptation is especially prevalent in MIS research due to the need to maintain a close alignment between a construct's definition, its measurement, and the rapidly evolving contemporary Information Technology (IT) or behavioral contexts (Compeau et al., 2022). Hence the teachings coming from the MIS practices would be applicable to other quantitative positive researchers in other behavioural sciences.

In this paper, we explore these adaptations by answering the following research question: What are the MIS scholars' practices for scale adaptations? By exploring this question we challenge the established methodological assumption that scale adaptation is a harmless practice that ultimately leads to the betterment of existing scales. Of course, MIS researchers aim at preserving the construct qualities when they adapt existing scales. While adapting the scales, it is even tempting for researchers to try to improve them, for example rephrasing a sentence which sounds awkward. But these supposed improvements are only rarely empirically demonstrated, as we, as MIS community, largely assume that these supposed improvements are real betterments, not requiring any additional rigorous measurement (Alvesson & Sandberg, 2011).

But measurement is a foundation on which behavioural science is built (Anastasi, 1954). Without the use of rigorously validated measures, high quality research is at stake (DeVellis, 2017). We echo the repeated observations that the MIS field does not adequately engage with the issues of measures and measurement and the need to improve measurement effort beyond current practice (Boudreau et al., 2001; Burton-Jones & Lee, 2017; Schmitz & Storey, 2020). While we do not deny that scale adaptation is necessary or useful in MIS, we argue for the need to spend an adequate effort when adapting scales, to preserve (or even to improve) their qualities, rather than degrading them. Given the spread of scale adaptations in MIS, we consider that it is important to understand this practice better in order to conclude more firmly on its impact for MIS scholarship.

Moreover, the scale adaptation is largely used as a “catch-all term” (Heggestad et al., 2019) to resume that something has been changed in the scale, recent studies have started to conceptualize different kinds of adaptations (Heggestad et al., 2019; Newman et al., 2016). Newman et al. (2016) formalized construct mixology as the practice of combining different constructs. Existing constructs are combined, reshaped, compounded to deliver a single construct. While the authors sometimes ignore and fail to mention that the constituent constructs exist, in other cases, the authors acknowledge and mention the existence of constituent constructs (Newman et al., 2016). Complementary, Heggestad et al. (2019) noted the practice of adapting a scale that was already the result of a previous adaptation. They conceptualize this practice as cascading adaptation, “where one author adapts a scale, and then subsequent authors use the adapted scale, citing the work of the author who adapted it” (Heggestad et al., 2019, p. 2614). This cascading adaptation contrasts with the possibility of the author of the most recent adaptation to exploit the original source of the scale, as far as the authors who adapted the original scale explicitly cited their source.

In this paper, we explore the practices of scale adaptation in the MIS discipline and we contribute to the conceptualization of these practices and we provide some recommendations regarding the identified ones, to offset the risks and problems these practices entail. To better understand this phenomenon of scale adaptations, we shed light on the “filial relationship” that might exist between scales. In that sense, we adopt a longitudinal perspective of the scale adaptation process over time and across studies using a “backward citation” approach (Webster & Watson, 2002). By making the process through which constructs and their measures evolve over time more explicit and transparent, it is our hope that the MIS community will be able to embrace a cumulative tradition of scales adaptation more fully.

2.Methodology

To improve our understanding of the adaptation phenomenon, we examined how Trust measures have undergone successive adaptations. We chose Trust measures as Trust represents the third most cited construct cluster with 63 constructs after IT adoption (412 constructs) and IS development (102 constructs) (Larsen & Bong, 2016). It is our hope that by zooming in on adaptations of Trust measures, we will be able draw conclusions that apply to a wider range of MIS measures.

2.1 Sample construction: Individual-level trust scales

We started our review with a keyword search using the EBSCO Business Source Complete database for all articles that included the word “Trust” in the title. We restricted our search to articles published in the senior scholars’ basket of eight journals in MIS (European Journal of Information Systems, Information Systems Journal, Information Systems Research, Journal of AIS, Journal of Information Technology, Journal of MIS, Journal of Strategic Information Systems, MIS Quarterly). The search time frame was set to the year of 2000 until 2022, to appreciate the evolution of the measures of Trust over time. The EBSCO engine returned, 123 entries, after cleaning off duplicates.

As expected from the results of review of MIS constructs performed by Larsen & Bong (2016), where trust-related scales can be further categorized in four subcategories, the 123

articles returned very heterogeneous measures of Trust. Therefore, we decided to retain only individual-level Trust measures that assess the extent to which an IT-enabled object is trustworthy. The other measures, such as personal dispositions to trust or (forward oriented) intentions to trust, were excluded. This resulted in a relatively homogeneous sample of 13 articles employing individual-level measures of trust. These articles are strongly related with one another both conceptually (they all deal with individual-level trust) and operationally (they all measure the extent of trust in a IT-enabled object), which allows a systematic comparison.

We examined if the trust measures used in the 13 shortlisted articles has been adapted from previously validated scales. When the authors declared that their Trust measure had been adapted, we explored the single or several bibliographic sources that were cited to locate the source scale. We proceeded with this backward search strategy (Webster & Watson, 2002) from one adaptation to another adding the papers that employed a source scale to our initial sample. We stopped our backward search when we hit a paper that uses a self-developed scale or a paper published in a non-MIS journal. Indeed, several of the scales in our sample originate from marketing, but as far as we primarily wish to inform the MIS community, we do not investigate cascading adaptation further beyond the MIS field. Following this backward search, our initial sample of 13 papers was expanded to a total of 30 papers in the MIS field.

2.2 Coding scale adaptation

For the 30 papers using an adapted measure of trust, we systematically compared the adapted scale with the single or several sourced scales and examined the adaptations of: scale size, scale response format, scale target object, scale context and scale rater entity. We looked for the evidence of adaptations in these five areas, by thoroughly examining the scales and the items them-selves in conjunction to the measure development section of the papers.

Scale size refers to the number of items included in the instrument and expected to represent a specific conceptual domain (Smith et al., 2000). The adaptations of scales lead in many cases to shortening them (Heggstad et al., 2019). In cascading adaptation, hence the scale could become shorter and shorter after each round of adaptation, hence losing some of its ability to capture different facets of the construct. Moreover, for multidimensional scales, entire dimensions may be dropped at once, or the adaptation mixes multiple sources (Newman et al., 2016). We code the evolution of the number of items employed at each adaptation and the evolution of the number of dimensions declared at each adaptation.

Scale response format refers to the kind of responses accepted for the items of the scale. The scoring can be guided by different kinds of anchors (e.g., Likert, Guttman, semantic differential, etc.). Complimentary, the number of available scoring points is fixed by the authors (Weijters et al., 2010) and not necessarily complying to the previous version of the scale. To account for this, we take into consideration, for each adaptation, the kinds of anchors and the number of available scoring points.

Scale target object, scale context and scale raters refer to changes in wording in order to make consistent an existing scale with the object of the questionnaire in terms of IT or behavioral contexts (Compeau et al., 2022). Scale target object refers to the object of the questionnaire,

i.e. the IT artifact or system being rated. These changes may be inconsequential (e.g., substituting the name of the target technology) or they may be significant (e.g., modifying the object of the evaluation from a technology to a person or an organization, and vice-versa). Scale context refers to the application domain of the survey, such as online B2B relationships, telemedicine and so on. MIS measures are often context-specific in that their items carry implicit assumptions about the settings in which they have been developed. Substantial adaptations may be needed when transferring these measures across contexts. For example, the Task Adaptation measure introduced by Schmitz et al. (2016) bears the assumption of usage in a work-context in its wording (e.g., R-TASK02 : “I strive to find ways to take on new work responsibilities by using my current mobile phone.”). If the construct was to be used in a non-work context, which would be theoretically relevant, its wording would have to be significantly altered. Scale rater entity refers to the subject providing the judgment on the target object. The rater can change too, given the ubiquitous nature of IT products and services. MIS measures may be employed across a wide range of population characteristics, such as professional and cultural backgrounds, expertise with IT topics, or demographic characteristics. Several item-level changes driven by the characteristics of the rater entity may be made so that the items are processed more smoothly by the target population of the questionnaire. For example, Davis et al. (2009) developed two versions of the survey instrument, one for the IS department and another for the user base, where “phrases were substituted as appropriate based on the target respondents” (p. 30). An example of adaptation of the rater entity appears in Sun & Fang (2010), where the author indicates “I designed the questions to be from the seller’s viewpoint” (Sun & Fang, 2010, p. 194). For these three facets, target object, context and raters, we recorded each one of them for each scale.

At the conjunction of all these kinds of adaptations, we calculate the percentage of text similarity, via Twinword (<https://www.twinword.com/api/text-similarity.php>), between each adapted scale to each source scale as synthetic proxy of the overlap across adaptations (Table 1).

Scale adaptations	Operationalization
Size	Number of items and number of dimensions
Response format	Kind of anchors and number of scoring points
Target object	Kind of IT/IS rated
Context	Kind of application domain of the survey
Rater entity	Kind of respondent
Synthetic proxy	Percentage of text similarity

Table 1: coding scale adaptations and their operationalization

3. Findings

On the total of 30 papers in the IS field, 8 papers self-developed their scale of trust, 11 papers adapted their trust scale exclusively from articles in MIS outlets, 8 papers adapted their trust scale exclusively from articles outside MIS field, and 3 articles combined sources from MIS and beyond.

3.1 Scale size

Our sample show that the size of the scale shrinks during adaptation. On average, the original scales have 9,03 items and 2,02 dimensions, while the adapted scales have 5,86 items and 1,45 dimensions. A t-test confirms that this difference is statistically significant, for both, items ($t(20) = -2,8; p < ,05$) and dimensions ($t(21) = -3,1; p < ,05$) (Table 2).

(5 missing data points)	Number of items in the adapted scales	Number of items in the sourced scales	Number of dimensions in the adapted scales	Number of dimensions in the sourced scales
Mean (Standard Deviation)	5,86 (2,7)	9,03 (4,9)	1,45 (0,74)	2,02 (1,02)
Min	3	3	1	1
Max	11	19	3	4

Table 2: scale size differences between adapted and sourced scales

3.2 Scale response format

The seven-point Likert scale is the main scale response format of the sample, for both adapted and source scales (9 missing data points). All the adapted scales are Likert type, except 2 semantic differential types and all are seven-point scales, except 5 scales with 5, or 9 or 11 points. Among the sources, there is a slight higher heterogeneity. There are three differential semantic scales and 10 scales have other than a seven-point scale, such as 5, 6, 9, 10 points. During adaptations, scale format has been changed in 9 cases: the number of points increased and decreased, and semantic differential scales were converted into Likert ones and viceversa, too.

3.3 Scale target object, context and rater

Scale target objects, contexts and raters change at almost every adaptation. The main kind of change involves the translation of brick and mortar contexts, raters, and objects of trust to click and mortar or pure player contexts, raters and objects. As said before, half of our sample sourced its scales from outside MIS field, at least partially if multi-sourced, and many scales outside MIS come from marketing. It means that the classical buyer-seller relationships of a brick and mortal retail measured in marketing (Ganesan, 1994) is adapted, in MIS, to the online marketplace relationships between online buyers and sellers (Ba & Pavlou, 2002). If the scale is sourced within MIS, the changes in objects, contexts and raters are not systematic, but still the majority of the cases. The e-commerce relationships between online buyers and sellers (Gefen, 2000) can change into workplace collaboration in virtual world among virtual world users (Chandra et al., 2012). Or, rather than measuring the trust in the web vendor in general (McKnight et al., 2002), a specific system is targeted such as the recommendation agent system (Qiu & Benbasat, 2009).

3.4 Scale similarity

Among the 21 articles adapting scales, 10 articles cite a single source of adaptation, 5 articles cite 2 other articles as sources, 5 additional articles refer to 3 different sources, and the last article adapted its scale from 4 sources. When a scale is sourced from multiple scales, we measured the similarity of the adapted scale with each source. We found that the degree of similarity decreases when the number of sources increases, as the items of the adapted scale are a mash-up of different items from distinct scales, making the adapted scales distant from each source separately (Table 3).

Number of sources	Text similarity (mean; standard deviation)
1	0,57; 0,32
2	0,44; 0,22
3	0,48; 0,23
4	0,34; 0,14

Table 3: Text similarity depending on the number of sources

For the scales adapted from a single source the similarity equals 0,57 meaning that around half of the text of the items is shared between the original scale and the adapted one and the other half of the text of the adapted scale is a new one, not presented in the original sourced scale.

4. Discussions

From the sample of scales we gathered, some insights can be advanced, about changes in scale size, in response format, in target objects, contexts and raters and changes in wording, leveraging the text similarity as a synthetic measure of all these changes. We explored some particular practices we encounter in order to advance in their conceptualization and in recommendations. This exploration involves: scale mixology, single source adaptations, and cascading adaptations.

4.1 Changes in size, response format, objects, contexts and raters

Concerning scale size, in line with previous studies (Heggstad et al., 2019), we observed that during adaptation, scales shorten in both items and dimensions. Hence, during cascading adaptation the scale is shorter and shorter after each round of adaptation. The shorter scale risks to lose some of its ability to capture different facets of the construct: « however well validated an instrument may have been in its original form, excising selected items from a validated instrument does not result in a derivative instrument » (Straub, 1989). Moreover, for multidimensional scales or scale with multiple sources, when dimensions are dropped only part of the original sources and hence only part of the conceptual domains are retained (Newman et al., 2016). When adaptations shorten the scales, we agree to the importance to check the correspondence with the construct expected to be measured by the adapted scale (Heggstad et al., 2019), to maintain content validity.

Related to scale response format, the changes are relatively rare and as far as there is no evidence in research of the weakening of the scale validities when these kinds of changes are applied, we do not have any recommendation for authors.

About scale target objects, contexts and raters, we confirm that MIS scholars regularly adapt existing scale to face rapidly evolving contemporary IT or behavioural contexts (Compeau et al., 2022). Moreover, we wonder to what extent publication efforts and strategies push the MIS scholars toward this continuous scale adaptations and keep them away from replicability (Daneshvar Kakhki et al., 2021). Replicability has gained attention, by the MIS community, only recently, but its diffusion could reduce the number of adaptations. But changes in wording are not only directly related to changes in target objects, contexts and raters. We observed that changes in item wording takes place even when the attributes of the scale (the rater, the object, the context) are the same in the initial and in the adapted versions. We alert

that some alterations in item wording could undermine the cognitive properties of the initial scale. For example, Pavlou (2003) adapted Jarvenpaa (2000)'s items, and changed wording, although both versions deal with consumers trust in web stores. In terms of impact on cognitive properties, for example, the second item changed from "I trust this store keeps my best interests in mind" (Jarvenpaa et al., 2000) to "I trust this Web retailer because they keep my best interests in mind" (Pavlou, 2003). Adding the conjunction 'because' not only subtly transforms the intended focus of the item. It also creates a double-barreled situation where it is unclear whether respondents should report the degree to which they trust the web retailer, or whether they believe in the assertion that web retailers should be trusted for a particular reason: that they keep their customers' best interest in mind. Such differences are subtle to grasp, and it is unlikely that they would arise very clearly through a dedicated pretest. Nevertheless, they have important implications on the quality of the empirics. These kinds of changes that could undermine the scale qualities are widely spread in our sample. A final example concerns the risk of introducing cognitive biases depending on the degree of knowledge of respondents on the questionnaire topic: Teo et al. (2008) used the item "I feel that government acts in citizen's best interest" as adaptation of "I believe that LegalAdvice.com would act in my best interest" (McKnight et al., 2002). In this adaptation, Teo et al. (2008) assume that the rater knows what "the citizen's best interest" may be, while the McKnight (2002)'s version assumes that the rater more humbly knows where his or her best interest lies. By adapting McKnight (2002)'s scale, Teo et al. (2008) probably overestimated the ability of respondents to know with sufficient confidence where citizen's best interest lies. This could result in answers that are less accurate than the initial version of the scale. Basing on our observations, we discourage authors to change wording more than what is strictly needed to fit the new scale target objects, contexts and raters. Otherwise, cognitive validities of the adapted scale should be tested (Pillet et al., 2023).

4.2 Scale mixology

As already stated, more than half of the sample sources its scales from multiple articles and text similarity of the scales adapted from multiple sources is lower than the text similarity of the scale adapted from a single source.

In line with what has been proposed for construct mixology (Newman et al., 2016), we suggest to refer to these scale adaptations from multiple sources as scale mixology. Scale mixology requires extra attention than simple scale adaptation. For example, Dimoka (2010), when presenting its trust scale, stated that: "The nine measurement items for benevolence are based on Gefen (2002) and Pavlou (2002), and the nine items for credibility follow Ba and Pavlou (2002), Pavlou (2003), and Gefen (2002)". But, Dimoka (2010) reduced the number of items to four for each dimension, even if the sources have different numbers of items. Concerning Benevolence dimension, Gefen (2002) has a 3 item scale for Benevolence, and Pavlou (2002) has a 3 item scale for Benevolence, too. At the end, looking at the Benevolence items used by Dimoka (2010), the item "[Seller] will keep my best interests in mind during this transaction" does not find a clear source in the cited sources, while the other three items are closer to Pavlou (2002) items than Gefen (2002) one. Concerning credibility dimension, Ba and Pavlou (2002) have a 3 item scale for credibility, Pavlou (2003) has 3 item scale for

Trust in general, and Gefen (2002) has 3 item scale for Integrity, a 3 item scale for Ability and Gefen (2002) evokes that Integrity and Ability are the constituents of Credibility. Finally, the Dimoka (2010)'s Credibility items have only few words in commons with the sourced scales to the extent that we were not even able to establish correspondences between items (Table 4).

Dimoka 2010, dimensions	Dimoka 2010, final items	Ba and Pavlou 2002 sourced items	Pavlou 2003 sourced items	Gefen 2002 sourced items	Pavlou 2002 sourced items
Benevolence	[Seller] is likely to care about my well-being during this transaction. [Seller] will keep my best interests in mind during this transaction. If there is a problem with this auction, [Seller] will go out on a limb for me. [Seller] is likely to make sacrifices for me during this auction, if needed.			I expect that Amazon.com have good intentions toward me I expect that Amazon.com intentions are benevolent I expect that Amazon.com are well meaning	Sellers in this B2B marketplace are likely to care for buyers' welfare. If there is a problem, sellers in this B2B marketplace are likely to go out on a limb for buyers. I feel that sellers in this B2B marketplace are likely to make sacrifices for buyers if needed.
Credibility	[Seller] has the expertise to understand my needs and preferences. [Seller] has the ability to successfully undertake this auction. [Seller] will deliver this product according to the posted delivery terms. [Seller] is likely to be credible during this transaction.	I think this seller is honest. I believe this seller will deliver to me the product I purchase according to the posted delivery terms and conditions. I believe this seller will deliver to me a product that matches the posted description	This Web retailer is trustworthy. This Web retailer is one that keeps promises and commitments. I trust this Web retailer because they keep my best interests in mind.	Integrity Promises made by Amazon.com are likely to be reliable I do not doubt the honesty of Amazon.com I expect that Amazon.com will keep promises they make Ability Amazon.com understands the market they work in Amazon.com knows about books Amazon.com knows how to provide excellent service	

Table 4: Scale mixology

All things together, we consider that Dimoka (2010)'s trust scale could benefit from passing through the validities of a new scale development process. As we observed these kinds of practices and changes in the other scales adapted from multiple sources, we recommend, that authors involved in scale mixology, test their mixed scales as if they were newly developed scales.

4.3 Single source adaptations

Looking at scales adapted from a single source, the text similarity index of 0,57 and the standard deviation of 0,32 highlight a great diversity of approaches. To an extreme, Awad and Rogowsky (2008) adapted the trust scale by Gefen et al. (2003), indeed, the number of items and dimensions of the scale remained the same, both scales are employed in the online shopping context, the rater is a consumer and the object of trust is for both scales a online vendor. At the end the similarity index is equal to 1 (Table 5).

Adapted scale' article	Adapted scale' items	Sourced scale's article	Sourced scale's items
Awad and Rogowsky	Based on my experience with the online vendor in the past, I know it is honest.	Gefen et al. 2003	Based on my experience with the online vendor in the past. I know it is honest .

2008	Based on my experience with the online vendor in the past, I know it cares about customers. Based on my experience with the online vendor in the past, I know it is not opportunistic. Based on my experience with the online vendor in the past, I know it provides good service. Based on my experience with the online vendor in the past, I know it is predictable. Based on my experience with the online vendor in the past, I know it is trustworthy. Based on my experience with the online vendor in the past, I know it knows its market.		Based on my experience with the online vendor in the past. I know it cares about customers . Based on my experience with the online vendor in the past. I know it is not opportunistic Based on my experience with the online vendor in the past. I know it provides good service . Based on my experience with the online vendor in the past. I know it is predictable . Based on my experience with the online vendor in the past. I know it is trustworthy Based on my experience with the online vendor in the past. I know it knows its market
------	--	--	--

Table 5: Single sourcing adaptation with the highest similarity

On the opposite, we have the adaptation by Wakefield et al. (2004) of the trust scale by Larzelere and Houston (1980). The similarity scale is 0,18 and the context, rater and object changed. The scale moved from measuring a close relationship between partners to measuring E-commerce context where a consumer is assumed to trust a web site (Table 6).

Adapted scale' article	Adapted scale' items	Sourced scale's article	Sourced scale's items
Wakefield et al. 2004	CameraShop's web site shows that it : Is primarily interested in the customer's welfare Can always be trusted Seems to be perfectly honest and truthful with me Can be trusted completely Is truly sincere in its promises to customers	Larzelere and Houston 1980	My partner is primarily interested in his (her) own welfare. There are times when my partner cannot be trusted. My partner is perfectly honest and truthful with me. I feel that I can trust my partner completely. My partner is truly sincere in his (her) promises. I feel that my partner does not show me enough consideration. My partner treats me fairly and justly. I feel that my partner can be counted on to help me

Table 6: Single sourcing adaptation with the lowest similarity

These two extremes show well the diversity of adaptations, which would imply performing different kind of validities. We would recommend to consider that extensive adaptations, like the one by Wakefield (2004), should be considered whole new scales which would benefit from passing through the full process of scale development. On the opposite, we wonder to what extent it is useful to mention that the Awad and Rogowsky (2008)'s scale is adapted from the Gefen et al. (2003)'s scale, rather than simply stating that the scale has been taken as it was, without any adaptation. This precaution of mentioning an adaptation while the scale remains the same is less risky than the opposite i.e. mentioning an adaptation while the scale is completely renewed.

Between these two extremes, we explored to what extent we could advance some sorts of thresholds or rules of thumb to discriminate scale adaptations, basing on the extension of their adaptation to recommend appropriate validation processes. Unfortunately, our sample revealed to be too small to advance any precise threshold or rule of thumb. Beyond the extreme cases cited before, we were able to lever only three additional adaptations: Pavlou (2003)'s adaptation of Jarvenpaa (2000) scale, Sun (2010)'s adaptation of Gefen et al. (2003)'s scale, Qiu (2009)'s adaptation of Wang and Benbasat (2007)'s scale. We processed the scale adaptations from the highest similarity index downward (Table 7).

Similarity index	Adaptation	Adapted scale	Sourced scale	Context, rater, object of trust and dimensions of the scale
0,86	Pavlou 2003's adaptation of	This Web retailer is trustworthy. This Web retailer is one that keeps promises and commitments.	This store is trustworthy. This store wants to be known as one who keeps promises and commitments.	The same contexts, raters objects of trust and dimensions

	Jarvenpaa et al. 2000's scale	I trust this Web retailer because they keep my best interests in mind.	I trust this store keeps my best interests in mind.	
0,81	Sun 2010's adaptation of Gefen et al. 2003's scale	I know buyers at uBid are honest. I know buyers at uBid usually care about sellers. I know buyers at uBid are not opportunistic. I know uBid buyers' behaviors are predictable. I know buyers at uBid are capable of doing business.	Based on my experience with the online vendor in the past. I know it is honest . Based on my experience with the online vendor in the past. I know it cares about customers . Based on my experience with the online vendor in the past. I know it is not opportunistic Based on my experience with the online vendor in the past. I know it provides good service . Based on my experience with the online vendor in the past. I know it is predictable . Based on my experience with the online vendor in the past. I know it is trustworthy Based on my experience with the online vendor in the past. I know it knows its market	The same contexts, raters objects of trust and dimensions
0,54	Qiu 2009's adaptation of Wang and Benbasat 2007's scale	The agent was competent in recommending digital cameras. The agent performed its role of recommending digital cameras very effectively. Overall, the agent was capable of providing suitable digital camera recommendations. In general, the agent was very knowledgeable about digital cameras. I believe that the agent's dealings with me were in my best interest. The agent's dealings with me felt like that it would do its best to help me. The agent's dealings with me felt like that it was interested in my well-being, not someone else's. I believe the agent's recommendations to me were truthful. I would characterize the agent's dealings with me as honest. The agent appeared to be unbiased	Trust–Competence 1. This virtual advisor is like a real expert in assessing digital cameras. 2. This virtual advisor has the expertise to understand my needs and preferences about digital cameras. 3. This virtual advisor has the ability to understand my needs and preferences about digital cameras. 4. This virtual advisor has good knowledge about digital cameras. 5. This virtual advisor considers my needs and all important attributes of digital cameras. Trust–Benevolence 1. This virtual advisor puts my interest first. 2. This virtual advisor keeps my interests in its mind. 3. This virtual advisor wants to understand my needs and preferences. Trust–Integrity 1. This virtual advisor provides unbiased product recommendations. 2. This virtual advisor is honest. 3. I consider this virtual advisor to be of integrity.	The same contexts and raters, but the object of trust changed from a non anthropomorphic to an anthropomorphic product recommendation agent. Dimensions are reduced from 3 to 1.

Table 7: Decreasing similarity of single sourcing scales

Two adapted scales have respectively 0,86 and 0,81 similarity indexes on their respective sourced scale. Given the limited differences in terms of wording, context, rater, object of trust and dimensions, we considered that the term ‘adapted’ is well suited for these kinds of changes. We do not recommend and specific validities before employing these adapted scales. The last scale has a similarity index of only 0,54 and it shows significant differences with its sourced one. The object of trust, the number of items and the number of dimensions changed. We think that these kinds of changes cannot simply considered adaptations and we recommend for scales having so many differences to consider them as new scales and validating them through the new scale development process. Between 0,81 and 0,54, we do not have examples in our sample, hence we cannot precise more the threshold between what we could consider still an adaptation or what should be considered as a new scale.

4.4 Cascading adaptations

As well as adaptations were sourced from one article or multiple articles, also cascading adaptation practices involve a single source or multiple sources. Concerning the 10 scales adapted from a single source, we identified only a cascading adaptation from another single source. All the other nine cascading adaptations involve multiple sources when moving backward. This single case concerns the trust scale by Qiu et al. (2009) sourced in Wang and

Benbasat (2007) who sourced their scale in McKnight et al. (2002). The three scales keep the same context of online shopping, the rater remains the consumer, while the object of trust evolved. Qiu et al. (2009) targeted an anthropomorphic product recommendation agent, Wang and Benbasat (2007) targeted a (non anthropomorphic) product recommendation agent, while McKnight et al. (2002) targeted a web-based vendor. The number of items and dimensions changed too. If Wang and Benbasat (2007) did not change the 11 items and 3 dimensions of McKnight et al. (2002)'s scale, Qiu et al. (2009) reduced to 1 the number of dimensions and to 10 the number of items. The similarity index is equal to 0,54 between Qiu et al. (2009)'s scale and Wang and Benbasat (2007)'s scale. The similarity decreases down to 0,32 between Qiu et al. (2009)'s scale and McKnight et al. (2002)'s scale (Table 8).

Scale version	Qiu et al. 2009	Wang and Benbasat 2007	McKnight et al. 2002
Number of dimensions	1	3	3
Number of items in the final version	10	11	11
Object of trust	Anthropomorphic product recommendation agent	Product recommendation agent	Web-based vendor
Similarity		0,54 with Qiu et al. 2009	0,28 with Wang and Benbasat 2007 0,32 with Qiu et al. 2009

Table 8: The case of cascading adaptation from a single source

As far as we considered, in the previous section, that a scale with a similarity index of 0,54 should pass through the whole process of a new scale development, the similarity indexes of the only example of cascading adaptation from a single source, invite to validate the scale as a new one at each step of this cascading adaptation. As far as we have only one example of cascading adaptation from a single source at each adaptation, we can only image the existence of cascading adaptations which keep adaptations close enough to the original source to still be confident that the changes are simple adaptations not requiring to pass through the validities of a new measurement scale.

Apart this unique case of cascading adaptation from a single source, all the other cascading adaptations were from multiple sources. We unpack one example of such cascading adaptation as illustration of the cascading adaptation practices. Our example is the Trust scale used in Zahedi et al. (2016), which we sourced back to 1991. This lineage of adaptations is the longest in our sample with 5 successive adaptations over 25 years. The latest version of the scale, which measures patients' trust in augmented virtual doctor office in the context of online medical care (Zahedi et al., 2016) is adapted from two scales: consumers' trust in health infomediaries in online environments (F. "Mariam" Zahedi & Song, 2008) and buyers' trust in community of online auction sellers in the context of online B2C marketplaces (Pavlou & Gefen, 2004). If the scale for consumers' trust in health infomediaries in online environments (F. "Mariam" Zahedi & Song, 2008) is self-developed, the other scale is adapted from two other scales: the consumers' trustworthiness of celebrities in the context of television advertising published in the Journal of Advertising Research (Ohanian, 1991), and the buyers' trust in sellers in the context of online B2B marketplaces (Pavlou, 2002). It follows that the scale for buyers' trust in sellers in the context of online B2B (Pavlou, 2002) is adapted from a scale measuring the retail buyers' trust in sellers in the context of brick-and-

mortar buyer-seller relationships published in the Journal of Marketing (Ganesan, 1994) (Table 9).

Scale version	Zahedi et al. 2016	Zahedi and Song 2008	Pavlou et al. 2004	Pavlou et al. 2002	Ganesan 1994	Ohanian 1991
Number of dimensions	1	Not available	1	2	4	1
Number of items in the final version	3	Not available	3	7	19	5
Context	Online healthcare	Online healthcare	Online B2C marketplace	Online B2B marketplace	Brick and mortar B2C retail	Television advertising
Rater	Patients	Consumers	Buyers	Buyers	Retail buyers	Consumers
Object of trust	Augmented virtual doctor office	Health infomediaries (e.g., websites, forums)	Community of online auction sellers	Online retailers (e.g., Amazon)	Physical Retailers	Celebrities as spokespeople
Similarity		Not available	0,36 with Zahedi et al. 2016	0,31 with Pavlou et al. 2004 0,40 with Zahedi et al. 2016	0,30 with Pavlou 2002 0,13 with Zahedi et al. 2016	Not available

Table 9: Example of cascading adaptation from multiple sources

A visual depiction of this cascading adaptation is provided in Figure 1.

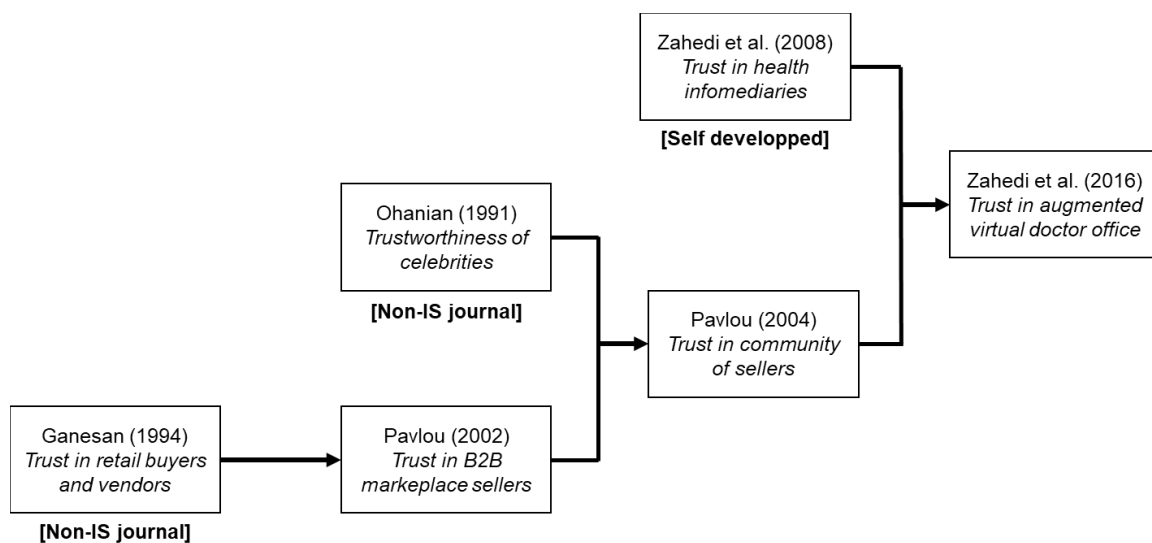


Figure 1 : Cascading Adaptation Resulting in the Trust in Augmented Virtual Doctor Office Scale

Moreover, these cascading adaptations involve changes in context, rater or object, which entails significant item wording alterations. Looking at the four successive adaptations undergone by the scale initially introduced by Ganesan (1994), we note that with the exception of the word “honest”, no other term is consistently found across the four adaptations of the scale. Even when the term “honest” is used across the four scales, it takes a very different meaning from one adaptation to the other:

- “If problems such as shipment delays arise, the resource's representative is honest about the problems” (Ganesan 1994)

- “Sellers in this B2B marketplace are likely to be honest in dealing with buyers” (Pavlou 2002)
- “Sellers in Amazon’s auctions are in general honest”. (Pavlou, 2004)
- “In receiving medical services at [physicians' office], I believe that [physicians' office] is very honest” (Zahedi et al 2016).

As can be seen in this example of a cascading adaptation, the wording in the final version of the item (Zahedi et al 2016) has little grammatical, lexical, and semantic overlap with the wording of the initial version of the item (Ganesan 1994). Between Zahedi et al (2016)’s scale and Ganesan (1994)’s scale there is at the end only 0,13 of similarity.

This evolution highlights the profound impacts of cascading adaptation practices in MIS. As far as we recommend for scale mixology, to go through a whole scale development process, consistently cascading adaptations from multiple sources should be avoided and prefer cascading adaptation only from single sources.

5. Conclusions

Scale adaptation is a widespread practice in MIS research to maintain a close alignment between a construct’s definition, its measurement, and the rapidly evolving contemporary IT or behavioural contexts. We challenged the established methodological assumption that scale adaptation is a harmless practice that ultimately leads to the betterment of existing scales. We explored the practices of scale adaptation in the MIS discipline, looking at the Trust scales. Based on a sample of 30 articles we highlighted the current scale adaptation practices, we classified them in scale mixology, single source adaptations, and cascading adaptations. We provided some recommendations for each kind of practices.

We encourage the fellow scholars to push this exploration further by extending the sample to the other most important constructs, beyond Trust, in order to have a better coverage of the MIS practices and to be able to better quantify the approximative thresholds we advanced.

References

- Alvesson, M., & Sandberg, J. (2011). Generating Research Questions Through Problematization. *Academy of Management Review*, 36(2), 247–271.
- Anastasi, A. (1954). *Psychological testing*. Macmillan.
- Awad, N. F., & Ragowsky, A. (2008). Establishing Trust in Electronic Commerce Through Online Word of Mouth: An Examination Across Genders. *Journal of Management Information Systems*, 24(4), 101–121. <https://doi.org/10.2753/MIS0742-1222240404>
- Ba, S., & Pavlou, P. A. (2002). Evidence of the Effect of Trust Building Technology in Electronic Markets: Price Premiums and Buyer Behavior. *MIS Q.*, 26(3), 243–268. <https://doi.org/10.2307/4132332>
- Boudreau, M.-C., Gefen, D., & Straub, D. W. (2001). Validation in Information Systems Research: A State-of-the-Art Assessment. *Management Information System Quarterly*, 25(1), 1–16.
- Burton-Jones, A., & Lee, A. S. (2017). Thinking About Measures and Measurement in Positivist Research: A Proposal for Refocusing on Fundamentals. *Information Systems Research*, 28(3), 451–467.

- Chandra, S., Srivastava, S. C., & Yin-Leng Theng. (2012). Cognitive Absorption and Trust for Workplace Collaboration in Virtual Worlds: An Information Processing Decision Making Perspective. *Journal of the Association for Information Systems*, 13(10), 797–835.
- Compeau, D., Correia, J., & Thatcher, J. (2022). When Constructs Become Obsolete: A Systematic Approach to Evaluating and Updating Constructs for Information Systems Research. *Management Information Systems Quarterly*, 46(2), 679–712.
- Daneshvar Kakhki, M., Mousavi, R., & Palvia, P. (2021). Evidence Quality, Transparency, and Translucency for Replication in Information Systems Survey Research. *Communications of the Association for Information Systems*, 49(1). <https://doi.org/10.17705/1CAIS.04903>
- Davis, J. M., Kettinger, W. J., & Kunev, D. G. (2009). When users are IT experts too: The effects of joint IT competence and partnership on satisfaction with enterprise-level systems implementation. *European Journal of Information Systems; Basingstoke*, 18(1), 26–37. <http://dx.doi.org/10.1057/ejis.2009.4>
- DeVellis, R. F. (2017). Scale Development: Theory and Applications. In *Applied Social Research Methods Series* (Vol. 4). Sage Publications.
- Dimoka, A. (2010). What Does the Brain Tell Us About Trust and Distrust? Evidence from a Functional Neuroimaging Study. *MIS Quarterly*, 34(2), 373–A7. <https://doi.org/10.2307/20721433>
- Ganesan, S. (1994). Determinants of Long-Term Orientation in Buyer-Seller Relationships. *Journal of Marketing*, 58(2), 1–19. <https://doi.org/10.1177/002224299405800201>
- Gefen, D. (2000). E-commerce: The role of familiarity and trust. *Omega*, 28(6), 725–737. [https://doi.org/10.1016/S0305-0483\(00\)00021-9](https://doi.org/10.1016/S0305-0483(00)00021-9)
- Gefen, D. (2002). Reflections on the dimensions of trust and trustworthiness among online consumers. *ACM SIGMIS Database: The DATABASE for Advances in Information Systems*, 33(3), 38–53. <https://doi.org/10.1145/569905.569910>
- Gefen, D., Karahanna, E., & Straub, D. W. (2003). Trust and TAM in Online Shopping: An Integrated Model. *Management Information System Quarterly*, 27(1), 51–90.
- Heggestad, E. D., Scheaf, D. J., Banks, G. C., Monroe Hausfeld, M., Tonidandel, S., & Williams, E. B. (2019). Scale Adaptation in Organizational Science Research: A Review and Best-Practice Recommendations. *Journal of Management*, 45(6), 2596–2627. <https://doi.org/10.1177/0149206319850280>
- Im, G., & Wang, J. (2007). A TETRAD-based Approach for Theory Development in Information Systems Research. *Communications of the Association for Information Systems*, 20(1). <https://doi.org/10.17705/1CAIS.02022>
- Jarvenpaa, S. L., Tractinsky, N., & Vitale, M. (2000). Consumer trust in an Internet store. *Information Technology and Management*, 1(1), 45–71. <https://doi.org/10.1023/A:1019104520776>
- Larsen, K., & Bong, C. H. (2016). A Tool for Addressing Construct Identity in Literature Reviews and Meta-Analyses. *MIS Quarterly*, 40, 529–551; A1. <https://doi.org/10.25300/MISQ/2016/40.3.01>
- Larzelere, R. E., & Huston, T. L. (1980). The Dyadic Trust Scale: Toward Understanding Interpersonal Trust in Close Relationships. *Journal of Marriage and Family*, 42(3), 595–604. <https://doi.org/10.2307/351903>
- Lee, A. S. (1999). Rigor and Relevance in MIS Research: Beyond the Approach of Positivism Alone. *MIS Quarterly*, 23(1), 29–33. <https://doi.org/10.2307/249407>

- McKnight, D. H., Choudhury, V., & Kacmar, C. (2002). Developing and Validating Trust Measures for e-Commerce: An Integrative Typology. *Information Systems Research*, 13(3), 334–359. <https://doi.org/10.1287/isre.13.3.334.81>
- Newman, D. A., Harrison, D. A., Carpenter, N. C., & Rariden, S. M. (2016). Construct Mixology: Forming New Management Constructs by Combining Old Ones. *Academy of Management Annals*, 10(1), 943–995. <https://doi.org/10.5465/19416520.2016.1161965>
- Ohanian, R. (1991). The impact of celebrity spokespersons' perceived image on consumers' intention to purchase. *Journal of Advertising Research*, 31(1), 46–54.
- Pavlou, P. A. (2002). Institution-based trust in interorganizational exchange relationships: The role of online B2B marketplaces on trust formation. *The Journal of Strategic Information Systems*, 11(3), 215–243. [https://doi.org/10.1016/S0963-8687\(02\)00017-3](https://doi.org/10.1016/S0963-8687(02)00017-3)
- Pavlou, P. A. (2003). Consumer acceptance of electronic commerce: Integrating trust and risk with the technology acceptance model. *International Journal of Electronic Commerce*, 7(3), 101–134.
- Pavlou, P. A., & Gefen, D. (2004). Building Effective Online Marketplaces with Institution-Based Trust. *Information Systems Research*, 15(1), 37–59. <https://doi.org/10.1287/isre.1040.0015>
- Pillet, J.-C., Vitari, C., Pigni, F., & Carillo, K. (2018). Detecting Biased Items When Developing a Scale: A Quantitative Approach. *AMCIS 2018 Proceedings*. <https://aisel.aisnet.org/amcis2018/MetaResearch/Presentations/8>
- Qiu, L., & Benbasat, I. (2009). Evaluating Anthropomorphic Product Recommendation Agents: A Social Relationship Perspective to Designing Information Systems. *Journal of Management Information Systems*, 25(4), 145–182. <https://doi.org/10.2753/MIS0742-1222250405>
- Recker, J. (2021). *Scientific Research in Information Systems: A Beginner's Guide*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-85436-2>
- Schmitz, K., & Storey, V. (2020). Empirical Test Guidelines for Content Validity: Wash, Rinse, and Repeat until Clean. *Communications of the Association for Information Systems*. <https://doi.org/10.17705/1CAIS.0xxXX>
- Schmitz, K., Teng, J. T., & Webb, K. (2016). Capturing the Complexity of Malleable IT Use: Adaptive Structuration Theory for Individuals. *MIS Quarterly*, 40(3), 663–686.
- Smith, G. T., McCarthy, D. M., & Anderson, K. G. (2000). On the sins of short-form development. *Psychological Assessment*, 12(1), 102.
- Straub, D. W. (1989). Validating instruments in MIS Research. *MIS Quarterly*, 13(2), 147–169.
- Straub, D. W., Gefen, D., & Boudreau, M.-C. (2005). Quantitative Research. In *Research in Information Systems: A Handbook for Research Supervisors and Their Students* (pp. 221–238). Elsevier.
- Sun, H., & Fang, Y. (2010). Toward a model of mindful acceptance of technology. *Proceedings of the 31st International Conference on Information Systems*, 12–15.
- TEO, T. S. H., SRIVASTAVA, S. C., & JIANG, L. (2008). Trust and Electronic Government Success: An Empirical Study. *Journal of Management Information Systems*, 25(3), 99–131.
- Wakefield, R. L., Stocks, M. H., & Wilder, W. M. (2004). The Role of Web Site Characteristics in Initial Trust Formation. *Journal of Computer Information Systems*, 45(1), 94–103. <https://doi.org/10.1080/08874417.2004.11645821>

- Wang, W., & Benbasat, I. (2007). Recommendation Agents for Electronic Commerce: Effects of Explanation Facilities on Trusting Beliefs. *Journal of Management Information Systems*, 23(4), 217–246. JSTOR.
- Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS Q.*, 26(2), xiii–xxiii.
- Weijters, B., Cabooter, E., & Schillewaert, N. (2010). The effect of rating scale format on response styles: The number of response categories and response category labels. *International Journal of Research in Marketing*, 27(3), 236–247. <https://doi.org/10.1016/j.ijresmar.2010.02.004>
- Zahedi, F. M., Walia, N., & Jain, H. (2016). Augmented Virtual Doctor Office: Theory-based Design and Assessment. *Journal of Management Information Systems*, 33(3), 776–808. <https://doi.org/10.1080/07421222.2016.1243952>
- Zahedi, F. “Mariam,” & Song, J. (2008). Dynamics of Trust Revision: Using Health Infomediaries. *Journal of Management Information Systems*, 24(4), 225–248.