

Measuring digital transformation in higher education institutions – content validity instrument

Digital
transformation

Lina María Castro Benavides

*Facultad de Ciencias Humanas y Bellas Artes, Universidad del Quindío,
Armenia, Colombia and*

*Facultad de Ingeniería y Arquitectura, Universidad Nacional de Colombia,
Sede Manizales, Manizales, Colombia*

Johnny Alexander Tamayo Arias

*Facultad de Ingeniería y Arquitectura, Universidad Nacional de Colombia,
Sede Manizales, Manizales, Colombia*

Daniel Burgos

Universidad Internacional de La Rioja, Logroño, España and

*Facultad de Minas, Universidad Nacional de Colombia,
Sede Medellín, Medellín, Colombia, and*

Alke Martens

*Faculty of Computer Science and Electrical Engineering, University of Rostock,
Rostock, Germany*

Received 16 March 2022

Revised 1 June 2022

12 August 2022

Accepted 21 August 2022

Abstract

Purpose – This study aims to validate the content of an instrument which identifies the organizational, sociocultural and technological characteristics that foster digital transformation (DT) in higher education institutions (HEIs) through the Delphi method.

Design/methodology/approach – The methodology is quantitative, non-experimental, and descriptive in scope. First, expert judges were selected; Second, Aiken's V coefficients were obtained. Nine experts were considered for the validation.

Findings – This study's findings show that the instrument has content validity and there was strong consensus among the judges. The instrument consists of 29 questions; 13 items adjusted and 2 merged.

Originality/value – A novel instrument for measuring the DT at HEIs was designed and has content validity, evidenced by Aiken's V coefficients of 0.91 with a 0.05 significance, and consensus among judges evidenced by consensus coefficient of 0.81.

Keywords Digital transformation, Higher education institutions, Content validate, Delphi method, Aiken's V , Consensus coefficient

Paper type Research paper

Introduction

Higher education institutions (HEIs) have been impregnated by the technological advance and force them to face a digital transformation (DT) in its dimensions: organizational,

© Lina María Castro Benavides, Johnny Alexander Tamayo Arias, Daniel Burgos and Alke Martens. Published in *Applied Computing and Informatics*. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at <http://creativecommons.org/licences/by/4.0/legalcode>

The authors are especially grateful to the judges who voluntarily evaluated the instrument with their invaluable experience and knowledge.

Conflict of interest: The authors have no conflicts of interest to declare.



Applied Computing and
Informatics
Emerald Publishing Limited
e-ISSN: 2210-8327
p-ISSN: 2634-1964
DOI 10.1108/ACI-03-2022-0069

technological and socio-cultural [1]. However, it was with the coronavirus disease 2019 (COVID-19) pandemic, where the lockdown and mandatory social distancing forced them to have a rapid transition from traditional classroom teaching to online learning [2], spending as little as weeks on this transition [3]. As a consequence, today's societies are facing complex challenges and changes that can only be addressed and resolved by pooling expert knowledge and finding responsible solutions [4]. In this sense, HEIs require the incorporation of technologies as an essential means for the continuity of professional, educational and social activities, and it is required that researchers deepen and theorize on the subject [5].

In the literature there are different research studies that study the DT in companies, providing a conceptual foundation and a defined measurement model for digital maturity to evaluate the current state and progress of their DT efforts [6]. Nonetheless, this aspect has not been so fully pursued for the educational sector [7]. In their study, researchers propose the application of an integrated DT model to assess the maturity level in educational institutions even though it is not the aim of their study to validate the relationships among the components of the model, but rather to provide a general framework to aid in the qualitative interpretation of the results [7].

In this sense, this paper emphasizes the evaluation of an instrument that will facilitate the obtaining of scientifically valid and reliable results, to validate the relationships among the constructs (organizational, socio-cultural and technological perspectives) that impact the DT in HEIs. Therefore, it is considered that to analyze complex constructs, the creation of valid and reliable instruments is required [8]. To achieve this goal, confirmation of the validity of the instruments is a prerequisite that guarantees the integrity of the study findings [9].

This article focuses on the content validity of the measurement instrument of DT in HEIs. There is an array of prior studies that demonstrate the application of content validity of measurement instruments in different areas; for example, Thanapatra and Uengpaiboonkit [10] study the elements of the development model for DT in hotels in Thailand; Singhdong *et al.* [11] analyze factors that influence DT of logistics services providers in a case study in Thailand; Moradi and Keshmiri [12] identify and classify effective factors for preparing to lead the DT in schools in Zarandieh; Rodriguez-Abitia and Bribiesca-Correa [13] suggested a DT model for businesses, and later advanced their application to perform a diagnosis of technological leveraging in Latin American organizations [14].

Additionally, this article contributes to the literature regarding the evaluation of content validity, offering a detailed description of the process we carried out, which could be adapted to new studies.

Theoretical background

This study conceives the DT at HEI as an organizational transformation, which should be taken into account and aligned three different perspectives (organizational, social and technological) in order to get the DT synchronized in a holistic integration and evolutionary over the course of time. This evolution may be the result of a gradual change of the behavior of individual elements in the enterprise, or it may be the result of a deliberate and conscious action. An organizational transformation is defined as the latter type of change, in other words, a deliberate and conscious action aiming to make changes to an enterprise [15]. With the Fourth Industrial Revolution, as said by Ref. [16], the transformation in its scale, scope and complexity will be unlike anything humankind has experienced before, it is characterized by a fusion of technologies that is blurring the lines between the physical, digital and biological spheres. Such dynamism requires and rewards a coherent sociology of universities as organizations [17]. The goal of a DT is continuous optimization – a company that can sense shifts in the market and respond quickly [18].

To orchestrate all these profound changes, a transformation is required within universities, synchronized with a paradigm shift for leadership in education. The change takes place as a natural manifestation of the developmental sequence in the natural maturation transformation process. At the execution level, it is found that the operational enterprise is concerned with “normal” operations (a first order system), while at the transformation level is found the enterprise transformation (a second order system) enterprise [15].

Digital transformation at HEI

DT is the profound and accelerating transformation of business activities, processes, competencies and models to fully leverage the changes and opportunities brought by digital technologies and their impact across society in a strategic and prioritized way [19]. Resembling any paradigm shift, DT in HEIs has been approached from the social, organizational and technological perspectives [1]. In that context, changes that have occurred within universities are evident in different settings. Inside these perspectives, dimensions within a HEI that have been permeated by DT processes found in the literature are teaching, infrastructure, curriculum, administration, research, business process, human resource, extension, DT governance, information and marketing [1]. This dissimilarity makes it understandable and remarkable, as summarized by Ref. [1], that DT within HEIs has been approached from different perspectives and although a consensus on its definition has not yet been consolidated, it is necessary to address the definition in the context of this study. In this sense, author [20] defines the DT as a disruptive or incremental change process. It starts with the adoption and use of digital technologies, then evolving into an implicit holistic transformation of an organization.

Digital transformation maturity

A dynamic and evolving nature of DT takes place in the universities, while they are impacted by technology. From a content perspective, typical dimensions cover the following aspects: (strategic) transformation management, core business comprising the digital product and service offering, digitization of internal processes and operations, digital customer interaction as well as IT use and development [21]. From an organizational context refers to the metarules that shape the organization’s rules and actions and provide meaning for them. It should create new directions and options for the organization [21]. These models are particularly relevant when both meta-capabilities in terms of transformation management and technological-driven changes in products, processes or business models are considered from a holistic perspective [21]. After analyzing the literature, different digital maturity models and their levels, digital maturity levels at HEI, authors consolidated five maturity levels, see [Table 1](#), maturity DT at HEI levels.

Digital transformation perspectives

DT in HEIs requires rethinking, restructuring and reinventing, from its multi-purpose, multi-processes, multi-disciplinary, multi-state and multi-actoral character. It is a collective effort that places the person in the center of the process of development, transformation and its impact on society. That is, DT should be an integral and holistic transformation of the HEI and should be approached from technological, organizational and social perspectives [1]. In the following paragraphs, these perspectives will be described.

Organizational perspective. It is remarkable the importance of this organizational perspective in the process of the DT at HEI. Companies that would like to progress with transformation should begin at the top [22]. This dimension consolidates most of the certain

Maturity levels	Description
No grade	The university lacks defined objectives and DT strategy. The success of some digital initiatives tends to depend on individual efforts, and the benefits obtained do not extend to other departments, programs and/or faculties. It lacks information digitization processes or is an incipient process. The staff (students, teachers, administrators) has little or no digital skills. The technology available to the university does not allow the digitization of the business
Low grade	The university has identified the need to increase the execution of discrete digital initiatives, to solve isolated business problems and allocates financial resources for their execution. However, its world view continues unchanged. Likewise, some internal initiatives for digital products and services that have been successful in the past are identified, and they are beginning to be replicated in other departments, programs and/or faculties. The staff (students, teachers, administrators) has basic digital skills. There is strong resistance to change. The university has the technological equipment to execute the digital initiatives, however the budgetary availability required to acquire them is restricted
In moderate degree	The university analyzes its world view and considers making an internal change. The university sees the need to make changes in university objectives in the medium term, and incorporates digitization initiatives and digital user experiences, although it does not yet focus on the disruptive potential of transformation. Therefore, the investment and use of new technologies is done with caution. The staff (students, teachers, administrators) have moderate digital skills, and resistance to change is still evident
In high grade	The university has understood, accepted and internalized the new digital paradigm, and it has decided to transform itself. Therefore, the DT capabilities of the university are perfectly adapted and incorporated into the vision, strategy, objectives and processes of the university, and also has a strategic planning that enables the transition to the new paradigm of DT in a successful way. The university's business model has been modernized, user-centric and adapted to the digital age. The university begins to obtain competitive advantages over other universities. The staff (students, teachers, administrators) has the required digital skills, resistance to change does not persist, digital culture has been internalized. The university has the necessary technology to achieve DT
In very high degree	The university is very innovative and disruptive using technology and new business models. The university flows naturally in the new digital paradigm. The university is always evaluating new technologies and their possible application. The processes are automated and advanced data analysis is used for decision making. It is a visionary and smart university. The staff (students, teachers, administrators) is immersed in a new digital culture

Table 1.
Maturity levels of DT
at HEI

dimensions that researchers had picked in separate around this perspective, and mainly four dimensions emerged: business process, administration, information and DT governance [6, 18, 22–25].

Social perspective. The world view, culture and digital competences of students, faculty, administrative, stakeholders are vital in the DT process at every organization, including universities. Due to that organizational change is related to people, processes, strategies, structures, and competitive dynamics where most of the challenges and opportunities reside [26].

Technological and physical perspective. The Fourth Industrial Revolution often is described as the result of integration and compounding effects of multiple “exponential technologies,” such as artificial intelligence (AI), biotechnologies and nanomaterials [27]. Today, cross-boundary digital technologies (IoT devices, 3D printing and big data analytics) drive transformations that go far beyond internal process optimizations as they potentially induce drastic changes to business models, organizational, corporate culture and entire industry structures [28]. As a consequence, the university should provide flexible IT, new enterprise

platforms and a strong and scalable operational backbone as part of an agile digital infrastructure [28], to support human resource, teaching, innovation, administration, access, market openness, building process, society and research [1].

On the other side, the university should have physical infrastructure to satisfy contemporary educational standards and methods (i.e. institutes of innovations, labs for teaching, digital training units, buildings, labs for teaching according with digital and innovation scenario) [1].

HEIs are paying close attention to the development of the DT process and are trying to successfully change the paradigm. While the effects of DT and their analytics, along with platform technologies, are becoming pronounced in companies, there is still a need to examine their implications on higher education [29]. The current state of research lacks a theoretically and methodologically profound maturity model for DT, not only in IT business but also in HEIs [1, 30]. Therefore, this paper develops dimensions and corresponding criteria for a DT implementation at HEI, based on the earlier discussion. The current study contemplates three dimensions that could positively influence the DT at HEI: organizational perspective – information, business processes (administrative or management, missionary, and support), management, strategic planning; socio-cultural perspective and technological perspective. Each dimension of this model provides indicators of the DT at HEI implementation. The study captures the idea of DT at HEI by turning toward a thorough study to investigate the perspectives incorporated in the DT and its relationships, and the instrument of measure was created. See Appendix A (URL: <https://n9.cl/fulz6>).

Validity

A commonly accepted definition of an instrument's validity is the extent to which the instrument measures what it intends to measure [31]. It is also considered a vital factor in the selection or application of an instrument [8] and serves as a guide to determine suitable implications of the findings of the study [31]. Specifically, validity is determined by examining concepts related to the validities of construct, criterion and content [9].

Construct validity refers to “how well” an instrument identifies the correct operative measures for the theoretical concepts under study [32, 33]. The criterion validity of a measuring instrument is established by comparing its results with those of a certain external criterion that attempts to measure the same thing [32] and finding a statistically significant relationship between a measure and a criterion [34]; and the content validity refers to the degree to which the items of an instrument are a representative sample of the content domain being measured [8, 35–37].

Content validity

Content validity indicates a full range of the attribute under study [9] in terms of clarity, coherence, relevance of the item and sufficiency of the dimension, with respect to a content domain [38]. To develop a set of scale items, a researcher first defines the construct of interest and its dimensions by searching the literature, through expert opinions, conducting population sampling or through qualitative research [9]. On the other hand, content validity can be characterized as apparent validity or logical validity. Apparent validity indicates that the measure appears to be valid “at first glance.” Logical validity indicates a more rigorous process, focusing on an evaluation provided by expert judges [34, 39], so that information on the representativeness and clarity of the elements can be provided and concrete suggestions offered to improve the quality of the instrument [8, 34]. If experts in the field are perceived as true experts, then it is unlikely that there is a higher authority that would challenge the assumption of validity of the content of the test [40]. In summary, Escobar [41] suggests the

following categories: sufficiency, clarity, coherence, relevance; and rating scale: does not meet the criteria, low level, moderate level, high level.

Numerous analytical techniques have been proposed to quantify the degree of agreement of experts regarding the relevance of the content of an instrument. Among the most used are the Aiken's V coefficient [42], Lawshe's content validity index – CVI [40], the content validity ratio (CVR) and the content validity coefficient – CVC [43]. Pedrosa *et al.* [44] detail other techniques.

Aiken's V. Aiken's V coefficient is considered the most relevant statistic for assessing content validity [36]. It evaluates the relevance of each item with respect to its construct, but also considers not only the number of categories offered to the judges, but the number of participating experts. Additionally, the assigned valuations can be dichotomous (values of 0 or 1) or polychotomous (values from 0 to 5) [44]. The V statistic is computed using the formula:

$$V = \frac{\bar{X} - l}{k} \tag{1}$$

where \bar{X} represents the median of the judges' ratings sample, l represents the lowest possible value, and k represents the range of possible values within the measurement scale used. The V statistic provides an index ranging from 0 to 1. A value of V equals 0 is obtained when all judges select the lowest possible rating, and a value of V equals 1 is obtained when all the judges select the highest possible rating [45].

Additionally, the result can be statistically evaluated by making use of the right-tail table of associated probabilities and accepting as valid only items that are statistically significant at 0.05. As shown in Table 2, if there are nine judges and four scale levels, an Aiken's V of 0.74 must be obtained for a statistical significance level of $p < 0.036$ and for the item to be considered valid [36].

In addition, the calculation of the confidence intervals for this coefficient allows us to test whether the magnitude obtained from the coefficient is greater than one that is established as minimally acceptable (0.70) to conclude the content validity of the items [47]. The confidence interval for Aiken's V can be calculated with equations (2) and (3).

$$L = \frac{2nkV + z^2 - z\sqrt{4nkV(1 - V) + z^2}}{2(nk + z^2)} \tag{2}$$

$$U = \frac{2nkV + z^2 + z\sqrt{4nkV(1 - V) + z^2}}{2(nk + z^2)} \tag{3}$$

In equations (2) and (3), z corresponds to the standard value of normal distribution such that $C\%$ of the distribution area is between $-z$ and z . For example, for a 95% confidence interval, $z = 1.96$ [45].

Table 2.	No. items (m) or raters (n)	Number of rating categories (c)	
		V	p
Right-tail probabilities (p) for 9 raters values of the validity coefficient (V)	9	0.81	0.007
	9	0.74	0.036
Source(s): Adapted from Ref. [46]			

Lawshe's content validity index (CVI). Lawshe's CVI determines a quantitative index for an instrument's content validity. The index requires that the expert judges evaluate whether the skill or knowledge measured by the item is "essential," "useful, but not essential," or "not necessary" to the performance of their job [40]. The CVI is the average of the values of items retained by applying the content validity ratio (CVR) proposed by Lawshe [40]. The formula for calculating CVR is:

$$CVR = \frac{n_e - \frac{N}{2}}{\frac{N}{2}} \tag{4}$$

where n_e = the number of expert judges who agree on the "essential" category and N = the total number of expert judges.

Lawshe presents this expression to be interpreted as a correlation, as it takes values from -1 to +1; CVR is then negative if fewer than half of the judges are in agreement, CVR is null if exactly half of the participants agree, and CVR is positive if more than half agree [37].

Additionally, in Table 1 explained by Lawshe [40], the minimum CVR values with a single-tail test as $p > 0.05$ are presented.

Content Validity Coefficient (CVC). CVC is an index measuring content validity based on agreement between judges [43] which recommends the participation of three to five expert judges and the use of a Likert scale of five alternatives. It is calculated with the average of one of the items among the maximum score that the item could obtain. Consequently, the error assigned to each item (Pe_i) is calculated. The formula for calculating CVC is:

$$CVC = CVC_i - Pe_i \tag{5}$$

Hernandez-Nieto [43] recommends keeping only those items with a coefficient higher than 0.80, while Balbinotti in Ref. [47] suggests keeping those items with a coefficient higher than 0.70. The CVC evaluation scale is presented in Table 3.

In the instrument's content validity, the judges' agreement must fall in the scale's upper range [8, 43].

Consensus among judges

An equivalent phrase for consensus (general agreement) could be "the collective opinion of a group," considering that some degree of variation between individuals is still possible [48]. Dissent is defined as a difference of opinion, such that a conflict occurs within the group committed to making a decision [48].

Within the context of content validation, if major revisions are needed for the measure, the researcher may want to repeat the process [34]. When there is a high measure of agreement, it indicates that there is consensus in the rating process among the evaluators, also indicating the interchangeability of the measuring instruments and the measure's reproducibility [41].

Values	Interpretation
0.00 to 0.40	Not acceptable
0.41 to 0.60	Very low
0.61 to 0.70	Low
0.71 to 0.79	Moderately low
0.80 to 0.90	Good
0.91 to 1.00	Excellent

Table 3. CVC evaluation scale

There are several techniques to determine this agreement, including Kappa [41, 43, 49]; Kendall's W [41, 50] and Wierman [48, 51].

Kappa statistic. This agreement index is only used in nominal data; however, there exists a generalization to include ordinal data called weighted k-coefficient [41]. The statistic has a range between -1 and 1 , but is generally found between 0 and 1 . If the coefficient is 1 , it indicates perfect agreement between the raters, if it is 0 , it indicates that the agreement is no more than the agreement expected due to random allocation, and if the coefficient value is negative, the level of agreement is below what was expected due to randomness [41].

Kappa values above 0.76 indicate "excellent" agreement, those between 0.40 and 0.75 are defined as "acceptable" agreement, and values below 0.40 are deemed "deficient or unacceptable" agreement [43]. Conversely, Landis and Koch [52] suggest another interpretation: "poor or weak" for Kappa values below 0.40 , "moderate" for values between 0.41 and 0.60 , "good" between 0.61 and 0.80 , and "very good" for higher Kappa values up to 1 .

Kendall's W. Kendall's W is a non-parametric range test that indicates the degree of association or level of concordance between the mean ranges of the evaluations performed by expert judges according to an ordinal scale [50]. An interpretation of Kendall's W is presented in Table 4.

Tastle–Wierman consensus. Tastle–Wierman consensus is a measure of dispersion as a representation of consensus (agreement) and dissension (disagreement). The measure is applied to a Likert scale (or any ordinal scale) to determine the degrees of consensus or agreement. With this measure, data on ordinal scales can receive a dispersion value that is logically and theoretically solid [48, 51, 53]. The minimum value assumed by the coefficient is 0 and the maximum is 1 . William J. Tastle and Wierman [48] have not defined parameters for interpreting this index; therefore, they shall be interpreted in the same way as Kendall's W . The average ratings of each element will also be used to calculate the final rating of a complete round [54].

Practical applications of this index can be consulted in studies such as Villaverde & Kosheleva [53], W J Tastle and Wierman [51], and William J. Tastle and Wierman [48].

Expert judgment

The first objective of content validity is to gain the opinion of expert judges who support or reject the appropriate operational definition within the universe of conceptual content. The ability to make effective decisions in situations where there is contradictory or insufficient information has led to an increased use of consensus methods, namely brainstorming, nominal group technique and the Delphi survey technique [55]. This study is developed following the Delphi method, which is detailed below.

Delphi method

The Delphi prospective method consists of a technique for obtaining information, based on consulting experts in an area, in order to obtain the most reliable consensus opinion of the

Kendall's W	Interpretation	Confidence in the classification
$0.1 \leq W < 0.3$	Very weak consensus	None
$0.3 \leq W < 0.5$	Weak consensus	Low
$0.5 \leq W < 0.7$	Moderate consensus	Moderate
$0.7 \leq W < 0.9$	Strong consensus	High
$W \geq 0.9$	Exceptionally strong consensus	Very high

Table 4.
Interpretation of
Kendall's W

Source(s): Taken from [56]

group consulted [57]. If used systematically and rigorously, this method can contribute significantly to expanding knowledge [55]. A group of at least five individuals is recommended to have enough control over casual agreement; and the maximum number of judges has not yet been determined. Each judge performs the evaluation of the quantitative and qualitative views on the items; after analyzing their answers, the research team sends each judge the resulting median, requesting that they reconsider their judgment until a consensus is reached [8, 41].

To verify the consensus of answers, the use of the Tastle–Wierman consensus coefficient is suggested. When *Consensus* > 0.70, there is good agreement, and the round is terminated.

Expert competence coefficient (K). One way to select the experts is through the analysis of their relevance as experts by calculating the expert competence coefficient (*K*) [58], using the formula [6]:

$$K = \frac{1}{2}(K_c + K_a) \tag{6}$$

K_c is the knowledge or information coefficient representing a measure of the level of knowledge about the subject. The judge’s self-assessment is required on a discrete scale of 0 to 1, and then multiplied by 0.1.

K_a is the argumentation or substantiation coefficient of the criteria of the experts. The judge self-assesses addressing six possible sources of argumentation [59] (See Table 5).

After obtaining the final value, experts with a rating below 0.8 are not included in the study [60].

The following section provides a detailed description of the systematic process, and the results obtained are analyzed.

Methods

This is a quantitative non-experimental descriptive study. During the process, the qualitative and quantitative points of view of expert judges regarding the research topic were collected following the methodology proposed by Ref. [55].

Study design

A web-based Delphi process was developed to engage expert judges and validate the content of the measurement instrument regarding the implementation of DT in HEIs.

The steps followed were (1) preparation of instructions and spreadsheets, (2) engagement of experts- Delphi method, and (3) collection and analysis of outcomes.

Preparation of instructions and spreadsheets. Prior to the construction of the instrument, a literature review was carried out. As a result, 29 items were identified and grouped into four main dimensions, twelve of them belonging to the organizational perspective (41.37%), six belonging to the socio-cultural perspective (18.18%), six that represent the technological

Source of argumentation or substantiation	High [3]	Medium [2]	Low [1]
Theoretical analyses you have carried out	0.3	0.2	0.1
Your experience gained	0.5	0.4	0.2
Studies of works by national authors	0.05	0.05	0.05
Studies of works by international authors	0.05	0.05	0.05
Your own knowledge about the state of the problem abroad	0.05	0.05	0.05
Your intuition	0.05	0.05	0.05

Table 5.
Degree of influence of
the source on your
criteria

perspective (20.68%), and five related to the maturity of the DT process in HEIs (17.24%) (See Appendix A. URL: <https://n9.cl/fulz6>).

Subsequently, the elements were refined and organized in a suitable format and sequence so that the information was collected in a useable form [8]. Furthermore, the instructions to the judges regarding the dimensions and the indicator measuring each item or a group of items were explained in detail through videos, a cover letter and written documents [41]. Finally, ethical considerations such as acceptance, permission to use and safeguarding of information provided by the expert judges were respected within the study.

Experts selection – Delphi method. Identification and selection of judges. The sample selection was non-probability, purposive and critical [61], according to interest in participating in the study, knowledge of DT in HEIs, experience and availability. Some experts were contacted through the emails available in the different scientific articles on DT, and other recognized experts in the different countries. Initially, a list of 32 judges was drawn up, each of whom received a letter of introduction and invitation to be part of the expert trial; eleven judges accepted. The expert judges were asked to respond to the self-assessment, following the requirements described in expert competence coefficient (K) section to calculate the expert competence coefficient (K); finally, the nine experts with a K rating greater than 0.8 were selected.

Preparation. The process was created with a criterion seeking the participation of expert judges from different geographical regions and with heterogeneous experience. The Delphi method of expert judges consisted of a total of nine judges: doctors – 4 and systems engineers or electricians – 5, who work as professors – 5, HEI dean – 1, or innovation directors – 3 in HEIs. The expert group also reflected a diversity of opinions with varied geographical representation, from the countries of Colombia – 5, Spain – 1, Portugal – 1, Guatemala – 1 and Germany – 1.

The role of the expert judges was to review these indicators and establish the level of agreement regarding the relevance, coherence, clarity and sufficiency of each item.

Collection and statistical analysis of the data. Data collection. The survey was conducted through the SurveyMonkey application. The Delphi method took place between October and November 2021. It started with the formal invitation to be part of the expert group of judges, information on the design of the study, and the link to enter the website, sent by email. Judges who accepted the invitation were asked to respond within three weeks of receiving the email. The communication and evaluation by experts were asynchronous. Up to two reminders were sent to participants who did not complete the survey in the specified period, and the survey was finally closed after two months, with very significant, complete and detailed contributions from nine judges.

The survey included four sections. The first provided information about the study, the self-assessment required to calculate the knowledge coefficient K_c . The second section referred to the identification of the future respondent. The third contained the main university characteristics and dimensions with their respective items, along with the description of the characteristics of each construct, scientific relevance and indicators to qualify each item. Finally, the fourth section collected the judges' self-evaluation information to calculate the argumentation coefficient K_a . Additionally, each section offered several free text fields for the judges' observations.

In this round and for each indicator, the experts had to indicate their agreement or disagreement on a 4-level Likert scale (does not meet the criteria, low level, moderate level, high level), according to sufficiency, clarity, coherence and relevance.

Analysis was carried out using Microsoft Excel, and R studio applying *Consensus function*.

Statistical analysis of the data

Two sets of statistics were applied to the answers provided by the judges, specifically the Aiken’s V [36], and the Tastle–Wierman consensus coefficient [51].

Consensus among judges – Tastle–Wierman consensus

A Tastle–Wierman consensus value greater than or equal to 0.7 would indicate a strong consensus, and the expert round would be finalized. However, if the value is less than 0.7, the survey must be resubmitted to the expert judges. In each iteration, the following information would be returned for each item: (1) The average classification of the item for the judges; (2) The judge’s rating of the item in the previous round; (3) An indication of the current level of consensus, based on the value of the Tastle–Wierman consensus coefficient; and (4) A paragraph summarizing the comments of the other participants on why they evaluated that element as they did. Based on this, the participants review their evaluations for each item with their respective observations [54].

The modifications to the items are made based on the expert judges’ observations and the results of Aiken’s V , considering that the item is maintained if the value of Aiken’s V is greater than or equal to 0.74; otherwise, it is modified or deleted.

The desired outcome is a tendency towards 1, and new rounds can be held if the minimum value of 0.7 in judges’ agreement is not reached in the first round, nor a value greater than 0.74 in the Aiken’s V coefficient with a statistical significance of 0.036.

Findings

Judge selection

The coefficients of knowledge and argumentation obtained for each judge were calculated (see Table 6).

According to Table 5, nine judges were identified who met the selection criteria presented at the beginning of the study, i.e., a K coefficient greater than or equal to 0.8.

Coefficient of content validity and consensus among judges

Once the coefficients have been calculated by item, by dimension and by the instrument (Appendix B. URL: <https://n9.cl/k2a6g>), appropriate levels in the Aiken’s V coefficient are deduced for most of the items, validating their content, except for item 4 clarity indicator ($V = 0.63$) and the business processes sub-dimension sufficiency indicator ($V = 0.70$).

Regarding consensus among the judges, calculated with the *Consensus* coefficient, a value greater than or equal to 0.78 is observed in most items of the instrument, indicating “Strong

Judge	Kc	Ka	K	Decision	
1	0.8	1	0.9	Influential	
2	0.8	1	0.9		
3	1	1	1		
4	1	1	1		
5	0.8	0.8	0.8		
6	0.9	1	0.95		
7	0.6	0	0.3	Not influential	
8	0.8	1	0.9	Influential	
9	0.9	0.8	0.85		
10	0.7	0.8	0.75	Medium influence	
11	0.8	1	0.9	Influential	

Table 6. Knowledge coefficient (Kc), argumentation coefficient (Ka), and expert competence coefficient (K), obtained for each judge

Consensus.” The exceptions are the clarity of item 3 (*Consensus* = 0.57); coherence (*V* = 0.60), relevance (*Consensus* = 0.65) and clarity (*Consensus* = 0.54) of item 4; business processes sub-dimension (organizational dimension) sufficiency (*Consensus* = 0.66); strategic planning sub-dimension (Organizational dimension) sufficiency (*Consensus* = 0.65); Clarity of item 11 (*Consensus* = 0.65); Clarity of item 13 (*Consensus* = 0.66); Relevance (*Consensus* = 0.50) and clarity (*Consensus* = 0.57) of item 20; and item 24, all indicators (coherence, relevance, clarity) (*Consensus* = 0.57).

The analysis of content validity and consensus among judges of the entire instrument is presented in Table 7.

It is observed that, in all cases, the Aiken’s *V* coefficient is greater than 0.74 with a *p* < 0.05, indicating that the items are considered valid; likewise, for all the dimensions, there is consensus greater than 0.74 among all the judges, indicating “Strong Consensus.” Conversely, the evaluation indicator for clarity in the technological dimension indicates a moderate consensus (*Consensus* = 0.68).

In response to observations provided by the experts, qualitative changes were made in the drafting of the questions and some context was modified to finalize the design of the questionnaire.

Discussion

The analysis of the consolidated outcomes of the instrument through Aiken’s *V* and *Consensus* coefficients indicates that the instrument has content validity and that there is strong consensus among the judges; therefore, a second round is not required. Conversely, the values that did not exhibit statistical significance in their content validity and/or had values lower than 0.7 in *Consensus* coefficient were carefully reviewed by the authors and reflecting the observations provided by the expert judges, wording and context were adjusted to facilitate the understanding of the item. Items 20 and 21 under the technological dimension were likewise merged into a single item.

In content validity studies, researchers can receive invaluable information from expert judges who provide constructive feedback on the quality of newly developed measures and the objective criteria against which they should evaluate each item [34]. The expert judges’ clear observations made it effortless to correct and adjust the items which presented values below those required by the Aiken’s *V* and *Consensus* coefficients. Consequently, the DT at HEI instrument will allow to measure the TD in the HEI considering the three organizational, socio-cultural and technological dimensions.

Conclusions, limitations and implications

The paper introduced the theoretical background section that supports the creation of the instrument, and subsequently the validity section that contributed to the literature regarding the evaluation of content validity, offering a detailed description of the process that were carried out. In addition, the methodological process described in the validity section could be adapted and applied to new studies.

Table 7. Content validity (*V*), confidence interval (CI), and consensus among expert judges (*C*)

Indicators Dimensions	Coherence			Relevance			Clarity			Sufficiency			All dimensions		
	<i>V</i>	CI	<i>C</i>	<i>V</i>	CI	<i>C</i>	<i>V</i>	CI	<i>C</i>	<i>V</i>	CI	<i>C</i>	<i>V</i>	CI	<i>C</i>
Organizational	0.94	[0.77–1.02]	0.93	0.97	[0.81–1.03]	0.92	0.88	[0.69–0.98]	0.89	0.89	[0.7–0.99]	0.77	0.92	[0.74–1]	0.84
Socio-cultural	0.95	[0.78–1.02]	0.86	0.97	[0.8–1.03]	0.95	0.91	[0.73–1]	0.78	0.89	[0.7–0.99]	0.77	0.93	[0.75–1.01]	0.82
Technological	0.90	[0.72–0.99]	0.86	0.93	[0.76–1.01]	0.91	0.86	[0.67–0.97]	0.76	0.81	[0.62–0.94]	0.77	0.88	[0.69–0.98]	0.74
DT maturity	0.96	[0.8–1.02]	0.74	0.96	[0.79–1.02]	0.80	0.96	[0.79–1.02]	0.68	0.96	[0.8–1.02]	0.74	0.96	[0.79–1.02]	0.88

The study of an instrument's validity is a long process, and the critical first step should be the study of content validity, for content validity verifies that the instruments used for studies are suitable for the construct, the population being researched, and the sociocultural context in which the study is being carried out.

Validity is not an instrument's property, but the property of the scores obtained by an instrument used for a specific purpose in a special group of respondents [8]. Since content validity is a prerequisite for other types of validity, it should have the highest priority during the instrument's development.

The evaluation of the instrument described in this paper facilitates the obtaining of scientifically valid and reliable results, to validate the relationships among the constructs (organizational, socio-cultural and technological perspectives) that impact the DT in HEIs.

A novel measurement instrument was designed and has content validity, evidenced by Aiken's V coefficients of 0.91 with a 0.05 significance, and consensus among judges evidenced by consensus coefficient of 0.81.

After analyzing the results of the Aiken's V and *Consensus* coefficients, it was concluded that the measuring instrument has content validity, as well as consensus among the judges. This indicated that a second round was not necessary.

There are some limitations that need to be pointed out. First, even though we tried to enhance 32 experts, there were only 9 judges who were allowed to participate in the Delphi study. The number of judges was clearly limited, it is reasonable to be prudent with the findings that cannot be generalized. Second, the DT at HEI is a topic of recent study, and this paper introduces a novel relationship that need to be validated by empirical studies to generalize the findings and subsequently become a reference model to implement.

In addition, our study makes practical implications, particularly providing understandings and sheds new lights on the most essential constructs integrated and related with the DT at HEI to engender the organizational transformation. Finally, the content validity of the instrument enables us to carry out the survey in the different institutions of higher education, to analyse and prove the model.

Finally, authors recommend incorporate additional IES characteristics in Section 1 of the instrument (Appendix A. URL: <https://n9.cl/fulz6>), to be considered as mediator and moderator variables that could help go beyond studying a complex relationship.

References

1. Castro Benavides LM, Tamayo Arias JA, Arango Serna MD, Branch Bedoya JW, Burgos D. Digital transformation in higher education institutions: a systematic literature review. *Sensors* (Switzerland). 2020; 20(11): 1-23.
2. Sousa MJ, Marôco AL, Gonçalves SP, de Bem Machado A. Digital learning is an educational format towards sustainable education. *Sustain*. 2022; 14(3): 1-16.
3. Bygstad B, Øvrelid E, Ludvigsen S, Dæhlen M. From dual digitalization to digital learning space: exploring the digital transformation of higher education. *Comput Educ*. 2022; August 2021: 182.
4. Stolze A, Sailer K, Gillig H. Entrepreneurial mindset as a driver for digital transformation - a novel educational approach from university-industry interactions. In: Stolze A, Sailer K, Gillig H (Eds). 13th European Conference on Innovation and Entrepreneurship, ECIE 2018, Strasczeg Center for Entrepreneurship, Munich University of Applied Sciences. Academic Conferences and Publishing International; 2018. 806-13. Available from: <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85055453507&partnerID=40&md5=e7d4b0ddb7aa113aeec05f5e02aff738>.
5. García-Peñalvo FJ. Digital transformation in the universities: implications of the covid-19 pandemic. *Educ Knowl Soc*. 2021; 22: 1-6.
6. Rossmann A, Reutlingen H. Digital maturity: conceptualization and measurement model social media view project startups in cooperation with incumbent firms view project digital maturity:

- conceptualization and measurement model. Thirty Ninth International Conference on Information Systems, San Francisco; 2018. 1-9. Available from: <https://assets.kpmg.com/content/dam/kpmg/pdf/2016/04/ch-digital-readiness-assessment-en.pdf>.
7. Rodríguez-Abitia G, Bribiesca-Correa G. Assessing digital transformation in universities, *Futur Internet*. 2021; 13(2): 1-17.
 8. Zamanzadeh V, Ghahramanian A, Rassouli M, Abbaszadeh A, Alavi - H. Design and implementation content validity study: development of an instrument for measuring patient-centered communication. *J Caring Sci*. 2015; 4(5): 165-78. doi: [10.15171/jcs.2015.017](https://doi.org/10.15171/jcs.2015.017).
 9. Devon HA, Block ME, Moyle-wright P, Ernst DM, Hayden SJ, Lazzara DJ, *et al*. A psychometric toolbox for testing validity and reliability. *J Nurs Scholarsh*. 2007: 155-64.
 10. Thanapatra V, Uengpaiboonkit A. The development model for digital transformation of hotel in Thailand. *Turkish J Comput Math Educ*. 2021; 12(8): 2638-43.
 11. Singhdong P, Suthiwartnarueput K, Pornchaiwiseskul P. Factors influencing digital transformation of logistics service providers: a case study in Thailand. *J Asian Financ Econ Bus*. 2021; 8(5): 241-51.
 12. Moradi S, Keshmiri S. Preparing to lead the digital transformation in schools. *J Sch Adm*. 2021; 9(2): 358-68.
 13. Rodriguez-Abitia G, Bribiesca-Correa G. Modelo de Transformación Digital en las Empresas Modelo de Transformación Digital en las Empresas. XXXII Congreso Nacional; Puebla, México; 2019.
 14. Rodriguez-Abitia G, Bribiesca-Correa G. Diagnóstico de Aprovechamiento Tecnológico en las Organizaciones de América Latina. ISLA 2020 Proceedings 19; 2020. Available from: <https://aisel.aisnet.org/isla2020/19>.
 15. Harmsen F, Proper HA, Kok N. Informed governance of enterprise transformations. *Advances in Enterprise Engineering II*. Berlin: Springer Verlag; 2009. 155-80.
 16. Schwab K. The Fourth Industrial Revolution: what it means, how to respond; 2015. Available from: <https://www.foreignaffairs.com/articles/2015-12-12/fourth-industrial-revolution>.
 17. Eaton C, Stevens ML. Universities as peculiar organizations. *Sociol Compass*. 2020; 14(3): 1-13.
 18. Gobble MM. Digital strategy and digital transformation. *Res Manag*. 2018; 61(5): 66-71. doi: [10.1080/08956308.2018.1495969](https://doi.org/10.1080/08956308.2018.1495969).
 19. Demirkan H, Spohrer JC, Welsler JJ. Digital innovation and strategic transformation. *IT Professional*. IEEE Computer Society; 2016. 14-8.
 20. Teichert R. Digital transformation maturity: a systematic review of literature. *Acta Univ Agric Silvic Mendelianae Brun*. 2019; 67(6): 1673-87.
 21. Chanias S, Hess T. How digital are we? Maturity models for the assessment of a company's status in the digital transformation, Munich: LMU; 2016. 14.
 22. Gurbaxani V, Dunkle D. Gearing up for successful digital transformation. *MIS Q Exec*. 2019; 18(3): 209-20.
 23. Castro Benavides LM, Tamayo Arias JA, Burgos D. Behavioural analysis of digital transformation in Latin American and Colombian universities, based on a general identification of variables. *Radical Solutions for Digital Transformation in Latin American Universities*. Lecture Notes in Educational Technology. 2021; 129-55. doi: [10.1007/978-981-16-3941-8_8](https://doi.org/10.1007/978-981-16-3941-8_8).
 24. Von Rosing M, Etzel G. Introduction to the digital transformation lifecycle overview of the digital transformation lifecycle. 2020; (2018): 92-9.
 25. Matt C, Hess T, Benlian A. Digital transformation strategies. *Bus Inf Syst Eng*. 2015; 57(5): 339-43. Oct 4 [cited 2018 Sep 25]. doi: [10.1007/s12599-015-0401-5](https://doi.org/10.1007/s12599-015-0401-5).
 26. Wade M. Digital business transformation A conceptual framework. *Glob Cent Digit Bus Transform*. 2015; (June). Available from: http://www.huffingtonpost.com/vala-afshar/accenture-digital-7-digital-business-transformation-lessons_b_6622648.html.

-
27. Penprase BE. Higher education in the era of the fourth industrial revolution. Palgrave Macmillan Singapore. 2018; 207-229. doi: [10.1007/978-981-13-0194-0](https://doi.org/10.1007/978-981-13-0194-0).
 28. Nadkarni S, Prügl R. Digital transformation: a review, synthesis and opportunities for future research. *Manag Rev Q*. 2021; 71(2): 233-341.
 29. Jackson NC. Managing for competency with innovation change in higher education: examining the pitfalls and pivots of digital transformation. *Bus Horiz*. 2019; 62(6).
 30. Gollhardt T, Halsbenning S, Hermann A, Karsakova A, Becker J. Development of a digital transformation maturity model for IT companies. *Proceeding - 2020 IEEE 22nd Conference on Business Informatics, CBI 2020, Vol. 1; 2020*. 94-103.
 31. Slocumb EM, Cole FL. A practical approach to content validation. *Appl Nurs Res*. 1991; 4(4): 192-195.
 32. Hernández Sampieri R, Fernández Collado C, Baptista Lucio P. *Metodología de la Investigación*. Vol. 1. Sexta ed. México D.F. McGraw Hill; 2014. 1-634.
 33. Yin RK. Case study research: design and methods. *Applied social research methods series*. Vol. 5. xiv; 2009. 219. Available from: <http://unisa.aquabrowser.com/?itemid=%7Cunisa-bibs%7C1104221%5Cnhttp://books.google.com/books?id=FzawIAdilHkC&pgis=1>.
 34. Rubio DM, Berg-Weger M, Tebb SS, Lee ES, Rauch S. Objectifying content validity: conducting a content validity study in social work research. *Soc Work Res*. 2003; 27. Available from: <https://academic.oup.com/swr/article/27/2/94/1659075>.
 35. Devon HA, Block ME, Moyle-Wright P, Ernst DM, Hayden SJ, Lazzara DJ, *et al*. Clinical scholarship A psychometric toolbox for testing validity and reliability. *JOURNAL NURSING SCHOLARSHIP*. 2007; 39: 155-164.
 36. Escurra LM. Cuantificación de la validez de contenido por criterio de jueces. *Rev Psicol*. 1988; 6(1-2): 103-11.
 37. Magdalena L, Sierra P, Enselmina M, Vargas M. Análisis de validez de contenido de un instrumento de transferencia de Tecnología Universidad-Industria de Baja California, México. *XX Congr Int Contaduría. Adm e Informática*; 2015. Available from: <http://congreso.investiga.fca.unam.mx/docs/xx/docs/2.02.pdf>.
 38. Carbajal AB, Tamari K. Validez de contenido y confiabilidad inter-observadores de Escala Integral Calidad de Vida. *Rev Psicol*. 2017; 35(2): 641-66.
 39. Slocumb EM, Cole FL. A practical approach to content validation. *Appl Nurs Res*. 1991; 4(4): 192-5.
 40. Lawshe CH. A. Quantitative approach to content validity. *Pers Pshycology*. 1975; 28: 569-75.
 41. Escobar-Pérez J, Cuervo-Martínez Á. Validez de contenido Y juicio de expertos: una aproximación a SU utilización. *Av en Medición*. 2008; 6: 27-36.
 42. Aiken LR. Educational and psychological measurement. *Educ Psychol Meas*. 1980; 40: 955-61.
 43. Hernandez-Nieto R. Contributions to statistical analysis: the coefficients of proportional variance, content validity and Kappa. *Platform CIP*; 2002. editor.
 44. Pedrosa I, Suárez-Álvarez J, García-Cueto E. Evidencias sobre la Validez de Contenido: avances Teóricos y Métodos para su Estimación [Content validity evidences: theoretical advances and estimation methods]. *Acción Psicológica*. 2014; 10(2): 3.
 45. Penfield RD, Giacobbi PR. Measurement in physical education and exercise science applying a score confidence interval to Aiken's item content-relevance index. *Meas Phys Educ Exerc Sci*. 2004; 8(4): 213-25.
 46. Aiken LR. Three coefficients for analyzing the reliability and validity of ratings. *Educ Psychol Meas*. 2015; 45(1): 131-42.
 47. Juárez-Hernández LG, Tobón S. Análisis de los elementos implícitos en la validación de contenido de un instrumento de investigación Analysis of the elements implicit in the validation of the content of a research instrument. *Espacios*. 2018; 39(53): 23-30. Available from: <https://www.revistaespacios.com/cited2017/cited2017-23.pdf>.

-
48. Tastle WJ, Wierman MJ. Consensus and dissention: a measure of ordinal dispersion. *Int J Approx Reason.* 2007; 45(3): 531-45.
 49. Bernal-García MI, Salamanca Jiménez DR, Norton Perez G, Mónica Paola QM. Validez de contenido por juicio de expertos de un instrumento para medir percepciones físico-emocionales en la práctica de disección anatómica. *Educ Médica.* 2018; 21(6): 349-56. Available from: <https://www.elsevier.es/es-revista-educacion-medica-71-pdf-S1575181318302584>.
 50. Zamora-de-Ortiz MS, Serrano-Pastor FJ, Martínez-Segura MJ. Content validity of the A-SIWT (ask, see, interpret, walk, tell) teaching model by expert judgment. *Form Univ.* 2020; 13(3): 43-54.
 51. Tastle WJ, Wierman MJ. An information theoretic measure for the evaluation of ordinal scale data. *Behavior Research Methods.* 2006; 38(3): 487-494.
 52. Cortés-Reyes É, Rubio-Romero JA, Gaitán-Duarte H. Statistical methods for evaluating diagnostic test agreement and reproducibility. *Rev Colomb Obstet Ginecol.* 2010; 61(3): 247-55.
 53. Villaverde K, Kosheleva O. Towards a new justification of the Tastle-Wierman (TW) dissention and consensus measures (and their potential role in education). *Reun Anu 2010 la Soc Norteam Proces la Inf Difusa;* 2010. 1-7.
 54. Okoli C, Pawlowski SD. The Delphi method as a research tool: an example, design considerations and applications. *Inf Manag.* 2004; 42(1): 15-29.
 55. Hasson F, Sinead K, Hugh M. Research guidelines for the Delphi survey technique. *J Adv Nurs.* 2000; 32(4): 1008-15.
 56. Schmidt RC. Managing Delphi surveys using nonparametric statistical techniques. *Decision Sciences.* 1997; 28(3): 763-774.
 57. López Vidal JF, Africa Calvo L. Delphi method validation and design of a questionnaire to assess physical activity characteristics in elderly from residences. *Retos.* 2019; 36: 515-20. Available from: www.retos.org.
 58. Mengual-andrés S, Roig-vila R, Mira JB. Delphi study for the design and validation of a questionnaire about digital competences in higher education. *Int J Educ Technol High Educ.* 2016; 28(3). doi: [10.1186/s41239-016-0009-y](https://doi.org/10.1186/s41239-016-0009-y).
 59. Cruz Ramírez M, Martínez Cepena MC. Origen y desarrollo de un índice de competencia experta : el coeficiente k. *Rev Latinoam Metodol la Investig Soc.* 2019; 16: 40-56.
 60. Zartha Sossa JW, Montes Hincapié JM, Toro Jaramillo ID, Villada HS. Método Delphi-Propuesta para el cálculo del número de expertos en un estudio Delphi sobre empaques biodegradables al 2032. *Espacios.* 2014; 35(13): 10.
 61. Taherdoost H, Business H, Sdn S, Group C, Lumpur K. Sampling methods in research methodology. *How to Choose a Sampling Technique for.* 2016; 5(2): 18-27.

Corresponding author

Lina María Castro Benavides can be contacted at: lmcastro@uniquindio.edu.co