REVISTA ECUATORIANA DE DESARROLLO SOCIAL Y AMBIENTAL
Octubre - Diciembre, 2026 | Vol. 1 | Núm. 4
ISSN: 3151-8303

Validez de contenido de veinte reactivos sobre desigualdad digital universitaria mediante juicio de tres expertos

Content validity evaluation by three expert judges of twenty items on digital inequality in higher education

Martín Romero Cruz
1 Universidad de Sonora; México;
martin.cruz@unison.mx
https://orcid.org/0009-0008-5199-083X

Enviado: 07/11/2026 **Aceptado:** 12/11/2026 **Publicado:** 16/11/2026

Tipo de Investigación: Artículo Original **Páginas:** 169 - 190

DOI: https://doi.org/10.66646/redsa-2026/11nov55

Conflicto de Intereses Ninguno que declarar.

Correspondencia: **Martín Romero Cruz**
Universidad de Sonora; México;
martin.cruz@unison.mx

Esta obra se distribuye bajo la licencia Creative Commons Atribución-NoComercial-CompartirIgual 4.0 Internacional (CC BY-NC-SA 4.0).
https://creativecommons.org/licenses/by-nc-sa/4.0/

RESUMEN

Este artículo presenta los resultados de una evaluación de validez de contenido mediante juicio de tres expertos aplicada a veinte reactivos del cuestionario «Percepción de Justicia en la Educación Digital». Los jueces, identificados mediante los códigos J01, J02 y J03 para preservar su confidencialidad, calificaron la relevancia, claridad y adecuación de cada reactivo en una escala de 1 a 4. Se analizaron 180 puntuaciones mediante el índice de validez de contenido por reactivo (I-CVI), indicadores de validez de contenido de la escala y V de Aiken. Las observaciones cualitativas agrupadas como J04/J05 se examinaron por separado y no se contabilizaron como puntuaciones adicionales. La relevancia obtuvo un I-CVI promedio de 0.800, frente a 0.533 para claridad y 0.767 para adecuación. Al integrar las puntuaciones con las observaciones sobre la formulación y correspondencia conceptual, se identificaron nueve reactivos con prioridad alta de revisión y once con prioridad media. Los hallazgos precisan qué aspectos del banco de reactivos requieren conservación, reformulación, separación o reasignación antes de su aplicación a estudiantes. La evaluación por expertos constituye una fase de examen del contenido del instrumento y no equivale a una validación psicométrica basada en respuestas estudiantiles. La lectura comparada de los criterios permitió diferenciar problemas de redacción de desplazamientos conceptuales, especialmente en equidad percibida, confianza digital y condiciones de estudio.

Palabras clave:
desigualdad digital; educación superior; validez de contenido; juicio de expertos; justicia transformativa.

ABSTRACT

This article reports the findings of a content validity evaluation conducted through the judgment of three experts on twenty items of the Perception of Justice in Digital Education questionnaire. The experts, coded J01, J02, and J03 to protect their confidentiality, rated each item for relevance, clarity, and adequacy on a four-point scale. Their 180 ratings were examined using the item-level content validity index (I-CVI), scale-level content validity indices, and Aiken's V. Qualitative observations grouped as J04/J05 were analyzed separately and were not treated as additional numerical ratings. The mean I-CVI was 0.800 for relevance, 0.533 for clarity, and 0.767 for adequacy. Integrating the expert ratings with the qualitative observations and item-domain correspondence led to nine items being assigned high revision priority and eleven medium priority. These findings identify which statements require retention, clarification, separation, or conceptual reassignment before administration to students. The expert-judgment phase evaluates the questionnaire's content and is distinct from subsequent psychometric analyses of student responses. The ratings were contrasted with item-domain correspondence and the recorded observations. Several difficulties involved absolute wording, combined events, and questions about other students' circumstances. The proposed revision path distinguishes connectivity incidents, institutional assistance, assessment practices, self-efficacy, and technological anxiety without drawing conclusions about students' responses.

Keywords:
digital inequality; higher education; content validity; expert judgment; transformative justice.

INTRODUCTION

Three expert judges evaluated the content of twenty statements in the Perception of Justice in Digital Education questionnaire. The immediate research problem was not whether digital inequality exists in universities, but whether the items written to examine it represented their intended attributes with sufficient relevance, clarity, and adequacy. The three judges' rating matrices showed a marked difference: relevance received more favorable scores than clarity. This discrepancy matters because an item may address a relevant problem yet combine different experiences, request information the respondent cannot know, or confuse material conditions with personal capabilities. The ratings, read alongside the qualitative observations, made it possible to identify which statements needed clearer wording and which required a more substantive conceptual adjustment.

University digital inequality is not explained by the mere existence of a connection. Van Deursen and van Dijk (2019) distinguish the possession of resources from effective material access; Michikyan et al. (2025) address differences between access, connectivity, skills, and participation. These distinctions were relevant when reviewing the items: having a computer does not guarantee that it is compatible with mandatory software, and experiencing an interruption in an exam does not by itself identify whether the network or institutional platform failed. The content evaluation had to verify that the questionnaire did not treat these circumstances as equivalent.

The questionnaire version submitted for expert judgment contained twenty ordinal statements, distributed initially across five domains with four items each. Three expert judges assessed the relevance, clarity, and adequacy of each statement; their coded quantitative evaluations are identified as J01–J03. Qualitative observations grouped as J04/J05 were examined separately and were not counted as votes in the quantitative indices. This phase establishes evidence about the content of the twenty statements and guides their revision. Administration of the questionnaire to students belongs to a subsequent phase and addresses different empirical questions.

Content validity requires examining the relationship between conceptual definition, content universe, and items (Haynes et al., 1995). Measurement standards further specify that validity concerns the interpretations and uses of scores (American Educational Research Association et al., 2014). In the present expert-judgment assessment, these requirements were operationalized as three explicit rating criteria—relevance, clarity, and adequacy—alongside qualitative examination of the wording and conceptual scope of the twenty items. The numerical indices summarize how the expert judges rated the statements; the qualitative observations help explain why individual items require revision.

The objective of this study was to evaluate the content validity of the twenty questionnaire items through the ratings of three expert judges, calculating I-CVI, S-CVI/AVE, S-CVI/UA, Aiken's V, and κ*; to contrast those numerical results with the separately recorded qualitative observations; and to determine which items should be retained, clarified, divided, or reassigned to a different domain. The article documents the conceptual domains, the expert-judgment procedure, the item-level evidence, and the resulting revision decisions.

It does not report the subsequent administration of the questionnaire to students.

1.1. Conceptual basis for the expert judgment of questionnaire content

1.1.1. Access: Having a device does not necessarily mean having adequate equipment

The expert-judgment review of Material Access included four items on device availability, software compatibility, access to applications, and equipment sharing (R12, R16, R02, and R09). Interpreting these items requires distinguishing the possession of a computer from the resources needed to perform university tasks, as discussed by van Deursen and van Dijk (2019) and Ali and Katz (2024). A shared device may be sufficiently available for academic work, whereas an individually owned computer may be unable to run required software. This distinction guided the examination of whether the four statements represented access conditions precisely rather than merely recording device ownership.

In the evaluation of R12 and R02, formulations such as "all my activities" and "without limitations" posed a specific content problem: a single exceptional failure could affect a response intended to describe the student's ordinary access. R16 required distinguishing device capability from software problems, and R09 required separating shared ownership from actual time available for use. The proposed revisions thus specify sufficiency and use conditions instead of leaving broad expressions to be resolved by the ordinal response scale.

1.1.2. Connectivity: recording an outage does not identify its origin

The four Connectivity items evaluated by the experts addressed connection stability, file transfers, interruptions during assessments, and internet availability (R19, R03, R07, and R14). These issues are consistent with the distinctions developed by Lythreatis et al. (2022) and Michikyan et al. (2025). R07 posed the most consequential problem: its wording combined a disconnection on the student's network with an interruption of the university assessment platform. A student can report an interruption and its consequences without necessarily knowing which system failed. This is why the content review recommended separating the two events rather than treating them as a single manifestation of connectivity.

1.1.3. Study conditions: a related set does not equate to a single dimension

In the four Study Conditions items (R05, R15, R20, and R10), the expert-judgment assessment concerned content from two different environments: students' conditions at home and support provided by the university. Study space, digital resources, platform operation, and opportunities to obtain assistance affect participation, as discussed by Amjad et al. (2024) and Timmis and Valladares-Celis (2026). Their relevance does not, however, make them interchangeable indicators. The review of R20 and R10, in particular, raised the need to distinguish forms of institutional assistance and platform accessibility; otherwise, a combined score could obscure which condition requires attention.

1.1.4. Perceived equity: listening to an assessment without confusing it with a population diagnosis

In the Perceived Equity domain, the review of R18 exposed a problem of what the respondent can know: the statement asks for a judgment about access opportunities for "all students," rather than

about experiences the individual has observed. By contrast, R08 and R13 concern more specific practices in assessment and teaching, while R04 compares face-to-face and virtual education. Luo and Liu (2026) distinguish conceptual approaches to digital equality; in this evaluation, that discussion supports separating personal experiences of institutional treatment from general claims about a university population. The recommended revisions therefore seek an observable referent for each item.

1.1.5. Confidence, ability, and anxiety: related phenomena that may vary separately

Within Digital Trust, the expert review of R06, R11, R01, and R17 required distinguishing perceived competence, adaptation to new tools, technical difficulties, and anxiety. This distinction is supported by Compeau and Higgins (1995) on computer self-efficacy and by the DigComp 2.2 framework (Vuorikari et al., 2022). R06 asks whether the student "feels confident" using technological platforms; that expression can indicate perceived ability or familiarity, but it does not directly establish trust in data protection. Thus, even a favorable quantitative assessment of R06 does not by itself resolve the conceptual interpretation of the term.

R01 made the boundary especially visible in the expert-judgment results: its wording expresses concern about technology-dependent exams and assignments, not necessarily low technological competence. Dikmen (2023) addresses anxiety as a specific phenomenon. A capable student may fear an external disconnection, while another student with fewer technical skills may not report the same worry. The recommendation emerging from the content analysis is therefore to assess anxiety distinctly rather than reverse its score and equate it with low self-efficacy.

1.1.6. The role of transformative justice within the model

Transformative justice served as the interpretive perspective for the expert-judgment findings, not as a sixth questionnaire domain. Fraser (1995), Young (2011), and Gready and Robins (2014) situate disadvantage and responsibility in institutional conditions rather than exclusively in individual attributes. Applied to the evaluated items, this perspective is particularly relevant to R07, R08, R18, and R20: their revision should make it possible to distinguish technological barriers from institutional decisions, available alternatives, and support responses. These implications arise from assessing item content; they do not establish the prevalence of barriers among students.

2. Methodology: operationalization and expert-judgment assessment

2.1. Domain operationalization and item correspondence

Table 1 identifies the twenty questionnaire items submitted for expert judgment and links each code to its intended content, bibliographic foundation, and the problem identified during the content review. This correspondence provides a basis for interpreting the experts' scores and the qualitative observations: revising an absolute term is different from separating two phenomena combined in one statement. The complete original wording appears in Annex A2, and the three judges' quantitative ratings appear in Annex A3.

TABLE 1.

Conceptual correspondence of the twenty items evaluated through expert judgment

Item	Specific content	Theoretical or empirical foundation	Content observation and revision decision
R12	Personal device adequacy	van Deursen and van Dijk (2019); Ali and Katz (2024)	The expression "all" can cause a single exception to determine the answer.
R16	How programs and platforms work on a student's computer	van Deursen and van Dijk (2019); Amjad et al. (2024)	It could confuse equipment capacity and software failures.
R02	Availability of required programs	van Deursen and van Dijk (2019); Ali and Katz (2024)	"All" and "without limitations" are broad formulations; distinguish licenses and actual access.
R09	Availability of a non-shared device	Michikyan et al. (2025); van Deursen and van Dijk (2019)	Sharing equipment does not necessarily imply insufficient access; assess time available for use.
R19	Connection stability for synchronous classes	Michikyan et al. (2025)	It may be irrelevant for students who do not attend synchronous classes.
R03	Connection speed for transferring academic files	Michikyan et al. (2025); Lythreatis et al. (2022)	Two actions are assessed: downloading and uploading files.
R07	Interruptions during assessments	Michikyan et al. (2025)	Conflates network outages with platform failures; consider separating them.
R14	Availability of internet service when needed	Michikyan et al. (2025)	"Always" may be excessively absolute.
R05	Space and distractions in the study environment	Timmis and Valladares-Celis (2026)	Adequate space and the absence of distractions do not always coincide.
R15	Access to digital academic resources	Amjad et al. (2024); Timmis and Valladares-Celis (2026)	This may belong to institutional resources rather than domestic conditions.
R20	Institutional support when difficulties arise	Amjad et al. (2024); Timmis and Valladares-Celis (2026)	Distinguish technological from academic support and assess students' own experience of institutional assistance.
R10	Platform accessibility and ease of use	Amjad et al. (2024); Luo and Liu (2026)	The item asks students to judge the experience of "everyone," which they may not be able to observe.
R18	Overall perception of equal access	Luo and Liu (2026)	Judgment about "all students" without access to all their circumstances.
R08	Perceived fairness of digital assessments	Luo and Liu (2026); Amjad et al. (2024)	Combines assessment fairness and technological advantage; check comprehensibility.
R13	Teachers' consideration of technological differences	Timmis and Valladares-Celis (2026); Luo and Liu (2026)	This may reflect experience of teaching practices rather than a general judgment about equity.
R04	Comparison of opportunities with face-to-face education	Luo and Liu (2026)	Comparing face-to-face and virtual modalities does not establish equity within digital education.
R06	Confidence in using platforms and tools	Compeau and Higgins (1995); Vuorikari et al. (2022)	"Confident" may refer to perceived ability or familiarity; it does not necessarily assess data privacy.
R11	Self-efficacy in solving common difficulties	Compeau and Higgins (1995)	"Without affecting my performance" also depends on factors unrelated to ability.
R01	Concern about technology-dependent exams and assignments	Dikmen (2023); Compeau and Higgins (1995)	Assesses technological worry or anxiety, not simply the inverse of self-efficacy.
R17	Confidence in adapting to new tools	Compeau and Higgins (1995); Vuorikari et al. (2022)	"Without difficulty" may be too absolute.

Source. Author's item–domain mapping based on the evaluated questionnaire and the cited literature.

2.2. Expert-judgment procedure and content-validity analysis

2.2.1. Questionnaire items and expert-rating criteria

The study evaluated twenty ordinal items from the Perception of Justice in Digital Education questionnaire through the judgment of three experts. Each item, coded R01–R20, was assessed for relevance, clarity, and adequacy on a four-point ordinal scale. The basic unit of the quantitative analysis was the rating assigned by one expert judge to one item on one criterion. The procedure corresponds to the content-validity assessment phase of instrument development (Rubio et al., 2003; Boateng et al., 2018).

The twenty items were grouped into five initial domains of four questions each: material access (A), connectivity (C), study conditions and institutional resources (SC), perceived equity (E), and digital trust (DC). The judges' ratings and the qualitative analysis were used to examine the appropriateness of the statements within these domains. Sociodemographic and technological characterization variables are part of the student questionnaire, but the expert-judgment evaluation reported here was restricted to the twenty content items.

2.2.2. Three expert judges, evaluation records, and analytical scope

The quantitative evidence comprised ratings issued by three expert judges, coded J01, J02, and J03. Each judge scored the twenty items on relevance, clarity, and adequacy using a 1–4 scale, yielding 60 ratings per criterion and 180 ratings in total (20 items × 3 criteria × 3 judges). The material examined also included qualitative observations grouped as J04/J05. Those observations informed the interpretation and revision proposals but were not counted as additional judges' numerical ratings; all content-validity indices were calculated using J01–J03 only.

The codes J01, J02, and J03 identify the three expert judges whose assessments form the quantitative basis of this study; their names were withheld to preserve confidentiality. The reported ratings were checked against the original assessment matrices. Observations recorded under J04/J05 were organized and analyzed separately from the three judges' ordinal scores.

Analysis of the expert judgments followed a defined sequence. First, the 180 ratings were ordered by item, evaluation criterion, and judge code. Second, item-level and scale-level content-validity indices were calculated. Third, differences among the three criteria were interpreted together with the qualitative observations and the conceptual definition of each domain. This procedure makes clear, for example, why an item that experts considered relevant may still need changes in clarity or conceptual placement.

Three outputs were developed: (a) complete matrices by criterion, with I-CVI, Aiken's V and κ* for each item (Annex A3); (b) a synthesis of content problems, preserving the code R01–R20 (Table 1 and Annex A1); and (c) a set of reasoned decisions of high or medium-priority revision (Tables 2 and 3). This sequence allows us to verify how a rating of favorable relevance can be accompanied by low clarity and why the decision on an item requires a review of both sources of evidence.

2.2.3. Qualitative observations and conceptual correspondence of items

The qualitative review organized each statement

according to its domain, observable referent, supporting literature, and interpretive problem. The relationship between what the question affirms and what a recipient could answer was read. For example, R07 links internet failures and platform failures; R20 brings together advice, technical support, and tutorials; R10 and R18 require "all" students to be assessed. In these cases, the observation is not limited to the length of the sentence, but indicates that an answer may represent different events or populations.

Four categories were applied to systematize the comments: (1) absolute formulations or without temporal reference; (2) combination of contents that admit different answers; (3) mismatch between question and assigned domain; and (4) knowledge requested beyond the experience of the respondent. Each item maintained its original code to relate comment, qualification, and adjustment proposal. The thematic organization procedure was supported by Gale et al. (2013).

The observations grouped under J04/J05 were linked to identifiable examples: "all" and "always" in R02 and R14; combined services or events in R07 and R20; domain problems in R01 and R04; and universal statements in R10 and R18. When objections coexisted, the one with the greatest effect on interpretation was recorded in the matrix. The priority assigned to each question is reported separately from its I-CVI so as not to confuse conceptual judgment and favorable coincidence.

2.2.4. Review priority rules

The review classification was constructed from the problems observed and their effect on what could be inferred from the item. Objections were considered to be absolute expressions ("always", "all" or "without limitations"), questions that bring together two different events and statements that require knowledge of the situation of other people. The review also examined items that mix anxiety, security and technological ability, or that compare educational modalities when the proposed domain is digital equity.

High priority was assigned when the content could change the interpreted attribute, mix non-equivalent events, or require information beyond the informant's reach. Medium priority was assigned when the primary referent was still identifiable and the adjustment consisted of delimiting frequency, activity, period, or sufficiency. The nine high-priority items are not simply the nine with the lowest rating: for example, R06 was unanimously favorably rated in relevance, clarity, and adequacy, but was singled out to clarify what "feeling confident" means in the digital trust domain.

2.2.5. Expert-judgment rating matrices and content-validity indicators

For every item and evaluation criterion, the matrices preserve the ratings of the three expert judges J01–J03 ($20 \times 3 \times 3 = 180$ ratings). Each item has three relevance ratings, three clarity ratings, and three adequacy ratings. The J04/J05 block was used only as a separate source of qualitative observations and did not increase the number of expert ratings used in the indices. Tables A3a–A3c document the individual scores and permit recalculation of the results.

Scores 3 and 4 on the ordinal scale from 1 to 4 were considered favorable. For each item and criterion, the I-CVI was calculated by dividing the number of favorable scores by three; its possible values were 0.000, 0.333, 0.667, and 1.000. Aiken's V

was obtained as $V = \Sigma(r - 1) / [3(4 - 1)] = \Sigma(r - 1) / 9$, where r represents each score (Aiken, 1985; Penfield & Giacobbi, 2004). The S-CVI/AVE corresponds to the mean of the twenty I-CVI and the S-CVI/UA to the proportion of items with three favorable scores. $\kappa^* = (I\text{-}CVI - Pc) / (1 - Pc)$ was also calculated, with $Pc = C(3, A) / 2^3$ and A equal to the number of favorable scores (Polit et al., 2007).

The content-validity indicators were calculated from the three expert judges' scores, first for each item and criterion and then for the set of twenty items. Table 4 reports the aggregated results and Tables A3a–A3c preserve every numerical rating. Among the three criteria, the experts evaluated clarity less favorably: its mean I-CVI was 0.533, compared with 0.800 for relevance. Revision priorities were assigned by integrating these results with the qualitative observations and conceptual correspondence; no single index was used as an automatic revision rule.

The contrast of both sources is essential. R08 received 4/4/4 in relevance (I-CVI = 1.000), but 2/2/2 in clarity (I-CVI = 0.000): its topic was considered pertinent and its formulation insufficiently clear. R18 had 3/3/3 in relevance and 2/2/2 in clarity, in addition to 2/2/1 in adequacy; the difficulty was linked to requesting a universal assessment of the students' opportunities. The matrix thus shows differences that a general average of the instrument would hide.

2.2.6. Reading and inference criteria

With three expert judges, the I-CVI for an item and criterion can take only four values: 0.000, 0.333, 0.667, or 1.000. One rating below 3 reduces full favorable agreement from 1.000 to 0.667. For that reason, the analysis presents both I-CVI and Aiken's V and interprets them alongside the qualitative observations. These indices provide evidence concerning the content judged by experts; they do not establish cognitive comprehension, internal consistency, or factor structure in students, which must be studied separately.

3. Results of the expert-judgment evaluation

3.1. Content-review priorities derived from expert judgments

The analysis of the expert judges' scores and the separately considered qualitative observations assigned nine items (45%) to high revision priority and eleven (55%) to medium priority. Table 2 presents this distribution by domain: all four Material Access items received medium priority; Connectivity had one high and three medium; Study Conditions had two of each; and both Perceived Equity and Digital Trust had three high and one medium. These are specific decisions concerning the twenty evaluated items, not general claims about the prevalence of digital inequality among university students.

TABLE 2.

Distribution of review priorities by domain

Domain	Items (n)	High priority (n)	Medium priority (n)	High (%)	Medium (%)
Material access (A)	4	0	4	0%	100%
Connectivity (C)	4	1	3	25%	75%

Domain	Items (n)	High priority (n)	Medium priority (n)	High (%)	Medium (%)
Study conditions (SC)	4	2	2	50%	50%
Perceived equity (E)	4	3	1	75%	25%
Digital trust (DC)	4	3	1	75%	25%
Total	20	9	11	45%	55%

Source. Review priorities derived from the item-level evidence shown in Appendix A1.

3.2. Two forms of ambiguity in access and connectivity

In Material Access, the records of relevance were favorable for R02, R09, R12 and R16 (I-CVI = 1.000 in all four). However, R02 received clarity ratings of 2/2/2 (I-CVI = 0.000), while R12 received 3/3/2 (I-CVI = 0.667). The difference is consistent with the comments on "all programs", "without limitations" and the scope of activities that a device must perform. The proposed adjustment does not eliminate the material content: it requires technical sufficiency, actual availability and conditions of shared use.

In Connectivity, the main objection was concentrated in R07. Its relevance scores were 2/2/2 and clarity scores were 2/2/2 (I-CVI = 0.000 in both criteria); adequacy reached 3/3/2 (I-CVI = 0.667). The qualitative reading explained the result: the item combines a disconnection attributable to the service used by the student with an interruption of the assessment platform. The recommendation was to distinguish these events. In contrast, R03 received 4/4/4 in relevance and 3/4/3 in clarity, although it was suggested to specify the meaning of downloading and uploading files.

3.3. The limits of a common score in study conditions

Two items from the Study Conditions domain concentrated the high-priority revision. R10 obtained 3/3/2 in relevance (I-CVI = 0.667), 1/2/1 in clarity (0.000) and 2/2/2 in adequacy (0.000). The content mixes accessibility and ease of use and, in addition, requires judging the platform "for everyone". R20, on the other hand, registered 4/4/3 in relevance (1.000), 2/2/2 in clarity (0.000) and 3/3/3 in adequacy (1.000). It was proposed to separate technical support, advice and tutoring, because a favorable response on one of these services does not allow us to conclude that the others were available.

3.4. Fairness: differences between an observed practice and a general opinion

Perceived equity brought together three high-review items: R04, R08, and R18. The numerical contrast is especially clear: R04 obtained I-CVI = 0.000 in relevance and adequacy; R08 obtained 1.000 in relevance, but 0.000 in clarity; R18, also 1.000 in relevance, dropped to 0.000 in clarity and adequacy. These are not three identical errors. R04 compares modalities and shifts the focus from digital equity; R08 mixes fairness in assessment with technological advantage; R18 assumes universal information that a person cannot know. The proposed reformulations distinguish teaching practices, real alternatives, and procedural justice.

3.5. Digital trust: the mistake of turning anxiety into low ability

Digital Trust ratings showed that the disagreement is not concentrated in a single writing flaw. R01 scored 1/2/1 in relevance (I-CVI = 0.000), although

it received 3/3/3 in clarity and 3/4/3 in adequacy (I-CVI = 1.000 in both criteria). The item can therefore be understood without difficulty, yet it measures concern about technological failures rather than confidence in one's own ability. The difference is not minor: a person may master the tool and still fear that an interruption will prevent them from finishing an exam or turning in an assignment.

The comparison with R06 and R11 prevents treating all domain problems in the same way. R06 received 4/4/3 in relevance, 3/3/3 in clarity, and 4/4/3 in adequacy, with I-CVI = 1.000 in all three criteria; however, the comment on the meaning of "feeling confident" advises distinguishing perceived competence or comfort with technology, without necessarily establishing confidence in data protection. In R11, the scores were 2/3/2 in relevance (I-CVI = 0.333), 2/2/2 in clarity (0.000), and 3/3/2 in adequacy (0.667), because solving technical problems and avoiding academic consequences are different results. In contrast, R17 achieved favorable unanimity in all three criteria. These differences justify keeping and refining statements with sufficient support, rephrasing those that mix attributes, and treating anxiety as specific content, without automatically converting their scores into the absence of self-efficacy.

TABLE 3

Priority conceptual review items

Item / domain	Central problem	Suggested revision or reformulation
R01 / DC	Anxiety is not the inverse of self-efficacy; a distinct affective domain may be needed.	Consider placing the item under technological anxiety: "I worry that a technical failure will prevent me from completing an online assessment."
R04 / E	Comparing educational modalities does not measure equity among students.	Replace the modality comparison: "When a technological limitation affects my participation, I have alternatives for completing the activity."
R06 / DC	"Feeling confident" may refer to perceived competence or familiarity, without establishing confidence in privacy protection.	Specify self-efficacy: "I am confident in my ability to use the platforms required for my courses."
R07 / C	The student's internet connection and the institutional platform can fail independently.	Separate disruptions to the student's internet connection from institutional platform failures during assessments.
R08 / E	General judgments about fairness and technological advantage are combined.	Specify procedural fairness: "If a technological problem arises during an assessment, I know how to request an alternative."
R10 / SC	Accessibility, ease of use, and judgments about the entire population are conflated.	Separate functional accessibility from ease of use; ask about the respondent's experience, not "all students."
R11 / DC	Ability to resolve technical problems is combined with the absence of academic consequences.	Remove the external outcome: "I am confident in my ability to solve common problems when using educational platforms."

Item / domain	Central problem	Suggested revision or reformulation
R18 / E	An individual respondent cannot assess fairness for all students.	Replace the universal judgment with an observable practice: “In my courses, justified connectivity problems are considered when setting deadlines.”
R20 / SC	Tutoring, academic advice, and various forms of technical support are conflated; availability differs from effectiveness.	Distinguish technical from academic support through separate questions about availability and assistance actually received.

Source. Synthesis of item-level ratings and recorded qualitative observations.

3.6. Quantitative content-validity results from three expert judges

The 180 ratings issued by the three expert judges (J01–J03) were processed by item and criterion. I-CVI expresses the proportion of judges assigning 3 or 4; Aiken’s V incorporates the intensity of their scores relative to the scale minimum; and κ* provides a descriptive chance-corrected agreement estimate under the binary assumption defined in the method. Table 4 summarizes the judges’ results across the three criteria, while Annex A3 preserves their scores for each item, allowing the differences among individual statements to remain visible.

TABLE 4

Content-validity indices based on the ratings of three expert judges (20 items)

Criterion	S-CVI/AVE (mean I-CVI)	S-CVI/UA	Mean Aiken’s V	Items with I-CVI = 1.000
Relevance	0.800	0.750	0.761	15/20
Clarity	0.533	0.450	0.506	9/20
Adequacy	0.767	0.650	0.639	13/20

Source. Computations using J01–J03 ratings reported in Appendix A3.

Note. S-CVI/AVE is the average of the twenty I-CVI values; S-CVI/UA is the proportion of items with three favorable ratings; Aiken’s V = Σ(r − 1)/9. There are 60 ratings per criterion and 180 ratings overall (J01–J03).

Relevance reached S-CVI/AVE = 0.800 and mean Aiken’s V = 0.761, with fifteen items of favorable unanimity (15/20). Clarity registered S-CVI/AVE = 0.533 and mean V = 0.506, with only nine (9/20). Adequacy obtained 0.767 and 0.639, respectively, with thirteen (13/20). The gap of 0.267 points between the average I-CVIs of relevance and clarity explains an important part of the drafting decisions. R08, R18 and R20 illustrate why thematic relevance is not enough; R01 and R04 also show that some questions require reconsideration of their domain rather than wording changes alone. Across all twenty items, the threshold of scores 3 or 4 therefore yielded 48 favorable relevance ratings, 32 favorable clarity ratings, and 46 favorable adequacy ratings out of sixty per criterion. These aggregate counts do not replace the separate item-level matrices: the items with unanimous favorable ratings differ sharply among criteria, and a favorable relevance rating does not imply that wording or conceptual placement is adequate.

4. Discussion of the expert-judgment findings and implications for the instrument

4.1. Expert-rated relevance versus clarity: an item-level discrepancy

The ratings of the three expert judges revealed a recurring discrepancy between identifying an item's content as relevant and considering its wording sufficiently clear. For R08, all three judges assigned 4 for relevance but 2 for clarity; R20 showed a similar contrast, with ratings of 4/4/3 for relevance and 2/2/2 for clarity. These are findings from the twenty evaluated statements, not evidence that the broader domain is unimportant. They indicate that some items introduce competing meanings into questions intended to represent a single attribute, consistent with the concerns raised by Jarvis et al. (2003) about the relationship between indicators and constructs. Revision should therefore retain content supported by the judges while removing ambiguity that undermines its interpretation.

4.2. Expert-judgment findings on item-domain correspondence

The observations went beyond grammatical errors. In Access, asking to have "all" resources without limitations produces an absolute criterion; in Connectivity, a platform interruption does not equate to a network failure. In Study Conditions, R20 adds different services; in Equity, R04 compares modalities rather than opportunities between students; in Trust, R01 introduces anxiety into a set aimed at self-efficacy. These cases do not require uniform correction: some may remain in their assigned domain with a delimited wording, while others require separation or relocation.

The interpretive scope of the expert-judgment findings is specific. The ratings evaluate the relevance, clarity, and adequacy of the twenty statements; they cannot establish how frequently students experience any particular barrier or how students will respond to revised items. A later student administration will permit investigation of response patterns, consistency, and internal structure. The expert assessment of questionnaire content and the empirical study of student experiences are different evidentiary stages.

4.3. Transformative justice and institutional accountability in the definition of items

Transformative justice fulfilled a specific function in the reading of content: to examine whether the questions distinguish personal disadvantages from the requirements, decisions, and responses of the institution. Fraser (1995), Young (2011), and Gready and Robins (2014) discuss why identifying a technological difficulty is different from explaining the conditions that produce it. In the instrument, this distinction especially affects R07, R08, R18, and R20. The analysis of the observations recorded revealed that some formulations transfer to the respondent the task of judging technical failures or the circumstances of an entire population; in these cases, the content must be oriented towards observable experiences and verifiable institutional procedures.

Separating connection failures from platform issues in R07, or reformulating R08 to ask about alternatives to technological problems, is not just a linguistic correction. It changes the type of information that the survey will be able to produce later. Instead of recording an undifferentiated gap and automatically attributing it to the student, the revised items could distinguish support needs, assessment design problems, and institutional responsiveness. This is a conceptual implication

derived from content assessment, not a student application result.

4.4. Item-revision decisions derived from expert ratings and qualitative observations

Combining the three judges' quantitative assessments with the qualitative observations led to differentiated item-level recommendations. For Material Access, technical sufficiency and effective availability should be specified. For Connectivity, R07 should distinguish network failures from platform incidents. For Study Conditions, academic support, technical assistance, and usability should be considered separately. For Perceived Equity, the statements should focus on observable institutional practices and alternatives. For Digital Trust, perceived competence must be distinguished from anxiety and data-protection concerns. Table 5 links each decision to the verification needed for a revised version of the questionnaire.

TABLE 5

Content decisions and relevant contrasts

Original domain	Proposed conceptual decision	Evidence needed before adoption
Material access (A)	Retain the content while distinguishing ownership, availability, and technical sufficiency; review R12, R16, R02, and R09.	Item-level relevance and cognitive testing of access and shared use.
Connectivity (C)	Retain stability and availability; separate network outages from platform failures in R07.	Free classification of items; specific experiences during assessments and activities.
Study conditions (SC)	Provisionally treat environment, resources, support, and platforms as a profile; do not aggregate automatically.	Assess content coverage, subdomains, and the relevance of R10 and R20.
Perceived equity (E)	Reconsider the placement of modality comparisons; distinguish institutional practice, personal experience, and general assessment.	Review R04, R08, and R18; compare with institutional information.
Digital trust (DC)	Separate anxiety from self-efficacy and specify context-dependent skills; review R01, R06, and R11.	Conceptual review and cognitive interviews; avoid reverse-coding anxiety as self-efficacy.

Source. Interpretation of the expert-judgment findings and item-level review decisions.

The evaluation also shows that a favourable numerical result does not automatically authorise the literal retention of an item. R06 is the clearest example: all three assessments exceeded the favourable threshold in all criteria, but the expression "feel confident" still allowed more than one interpretation in the observations. Conversely, R01 was understandable and suitable for expressing technological concern, although its relevance within Digital Trust was low. This led to the need to document what is retained, what is split, and what is reassigned to another domain before the student response phase.

4.5. Interpretive limits and directions for subsequent instrument development

The meaning of these findings is bounded by the evaluation design. Three judges supplied the quantitative scores; their assessments offer

a focused appraisal of item content, not a broad estimate of consensus across all potential disciplinary perspectives. This matters particularly for the I-CVI: with three judges, one rating that falls below the favorable threshold changes an item's proportion from 1.000 to 0.667. Such a shift has a clear arithmetical meaning, but it does not establish that the content of an item is either wholly suitable or wholly unsuitable. For this reason, reading the I-CVI together with Aiken's V and the wording observations is more informative than assigning an automatic accept-or-reject verdict to each statement.

A second limitation arises from the kinds of judgments that a student can reasonably make. Several original statements ask about every student, every situation, or several kinds of institutional assistance at once. Even if respondents understand those statements grammatically, the resulting answers would remain difficult to interpret. Item revision must therefore specify the experiential referent: the course attended, the type of connection failure, the institutional response received, or the activity completed. This is not a merely stylistic issue. It affects whether a later score describes technological resources, an institutional practice, or the respondent's perceived capacity to act under particular conditions.

A third distinction concerns the status of qualitative observations grouped as J04/J05. The study used these comments as interpretive material; they did not constitute additional quantitative ratings. Maintaining that separation prevents the scope of the numerical panel from being overstated and helps make the revision recommendations understandable. Nevertheless, the proposed wording has not been tested in this article with students completing the revised questionnaire. Cognitive interviews could examine how students interpret words such as "confident," "equitable," and "accessible," while later empirical administration could test response patterns and the relationships among domains. Until then, the evaluation supports decisions about content representation but not claims about measurement reliability or the prevalence of digital inequality.

Finally, the transformative-justice perspective directs attention to how questionnaire wording allocates responsibility. If a failed online assessment is framed only as a student's connectivity problem, the instrument may overlook a platform malfunction or the absence of an institutional alternative. The revisions proposed for R07, R08, R18, and R20 aim to make those institutional conditions visible without presuming the outcome of future data collection. That distinction is essential if the questionnaire is to contribute to the diagnosis of unequal digital educational opportunities rather than reproduce an individualizing explanation of their causes.

5. Conclusions from the expert-judgment content evaluation

The evaluation by three expert judges produced 180 ratings of the twenty questionnaire items across relevance, clarity, and adequacy. The mean I-CVI was 0.800 for relevance, 0.533 for clarity, and 0.767 for adequacy; the corresponding mean values of Aiken's V were 0.761, 0.506, and 0.639. The judges' ratings showed favorable unanimity for fifteen items in relevance, nine in clarity, and thirteen in adequacy. When these findings were examined alongside the separate qualitative observations, nine items were classified as high

priority for revision and eleven as medium priority. The central result is not an undifferentiated judgment about the questionnaire: it is an item-specific account of where the experts' ratings supported the content and where the wording or conceptual correspondence required revision.

The contribution of this expert-judgment study is a traceable set of item-level content decisions. R07 should distinguish student internet interruptions from institutional platform failures; R10 and R20 should separate conditions and services; R04, R08, and R18 should specify the referent of perceived equity; and R01, R06, and R11 should distinguish anxiety, perceived ability, and technological security. Transformative justice frames the examination of institutional conditions that the questionnaire is intended to capture. The outcome reported here is a content-validity evaluation by experts and a set of proposed revisions. Application to students and subsequent psychometric analyses remain separate stages of the research.

REFERENCES

Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. Educational and Psychological Measurement, 45(1), 131–142. https://doi.org/10.1177/0013164485451012

Ali, C., & Katz, V. S. (2024). Evidence-based pathways for leveling digital access in higher education during and beyond the COVID-19 pandemic. Journal of Information Policy, 14, 655–685. https://doi.org/10.5325/jinfopoli.14.2024.0019

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). Standards for educational and psychological testing. American Educational Research Association. https://ncme.org/resources/books/testing-standards/

Amjad, A. I., Aslam, S., Tabassum, U., Sial, Z. A., & Shafqat, F. (2024). Digital equity and accessibility in higher education: Reaching the unreached. European Journal of Education, 59(4), e12795. https://doi.org/10.1111/ejed.12795

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. Frontiers in Public Health, 6, Article 149. https://doi.org/10.3389/fpubh.2018.00149

Compeau, D. R., & Higgins, C. A. (1995). Computer self-efficacy: Development of a measure and initial test. MIS Quarterly, 19(2), 189–211. https://doi.org/10.2307/249688

Dikmen, M. (2023). Test anxiety in online exams: Scale development and validity. Current Psychology, 42, 30210–30222. https://doi.org/10.1007/s12144-022-04072-0

Fraser, N. (1995). From redistribution to recognition? Dilemmas of justice in a 'post-socialist' age. New Left Review, (212), 68–93. https://doi.org/10.64590/4rl

Gale, N. K., Heath, G., Cameron, E., Rashid, S., & Redwood, S. (2013). Using the framework method for the analysis of qualitative data in multidisciplinary health research. BMC Medical Research Methodology, 13, Article 117. https://doi.org/10.1186/1471-2288-13-117

Gready, P., & Robins, S. (2014). From transitional to transformative justice: A new agenda for practice. International Journal of Transitional Justice, 8(3), 339–361. https://doi.org/10.1093/ijtj/iju013

Haynes, S. N., Richard, D. C. S., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. Psychological Assessment, 7(3), 238–247. https://doi.org/10.1037/1040-3590.7.3.238

Jarvis, C. B., MacKenzie, S. B., & Podsakoff, P. M. (2003). A critical review of construct indicators and measurement model misspecification in marketing and consumer research. Journal of Consumer Research, 30(2), 199–218. https://doi.org/10.1086/376806

Luo, J., & Liu, X. (2026). What do we mean by digital equality in education? Toward five conceptual lenses based on a systematic review. Journal of Research on Technology in Education, 58(5), 1165–1185. https://doi.org/10.1080/15391523.2025.2487279

Lythreatis, S., Singh, S. K., & El-Kassar, A.-N. (2022). The digital divide: A review and future research agenda. Technological Forecasting and Social Change, 175, Article 121359. https://doi.org/10.1016/j.techfore.2021.121359

Michikyan, M., Subrahmanyam, K., Regan, P., Cano, M. Á., Castillo, L. G., Harkness, A., & Schwartz, S. J. (2025). Differences in and relationship between digital access, use, connectivity, skill, engagement, and motivation: Implications for digital inequality in online learning among first-generation and continuing-generation undergraduate students. Discover Education, 4, Article 9. https://doi.org/10.1007/s44217-024-00371-9

Penfield, R. D., & Giacobbi, P. R., Jr. (2004). Applying a score confidence interval to Aiken's item content-relevance index. Measurement in Physical Education and Exercise Science, 8(4), 213–225. https://doi.org/10.1207/s15327841mpee0804_3

Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. Research in Nursing & Health, 30(4), 459–467. https://doi.org/10.1002/nur.20199

Rubio, D. M., Berg-Weger, M., Tebb, S. S., Lee, E. S., & Rauch, S. (2003). Objectifying content validity: Conducting a content validity study in social work research. Social Work Research, 27(2), 94–104. https://doi.org/10.1093/swr/27.2.94

Timmis, S., & Valladares-Celis, M. C. (2026). Digital inequalities and the COVID legacy in higher education in the global South and North: Intersecting inaccessibilities and institutional assumptions. Compare: A Journal of Comparative and International Education, 56(3), 429–447. https://doi.org/10.1080/03057925.2025.2483691

van Deursen, A. J. A. M., & van Dijk, J. A. G. M. (2019). The first-level digital divide shifts from inequalities in physical access to inequalities in material access. New Media & Society, 21(2), 354–375. https://doi.org/10.1177/1461444818797082

Vuorikari, R., Kluzer, S., & Punie, Y. (2022). DigComp 2.2: The digital competence framework for citizens—With new examples of knowledge, skills and attitudes. Publications Office of the European Union. https://doi.org/10.2760/115376

Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content

analysis and recommendations for best practices. The Counseling Psychologist, 34(6), 806–838. https://doi.org/10.1177/0011000006288127

Young, I. M. (2011). Responsibility for justice. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195392388.001.0001

APPENDICES

Traceability of the expert-judgment evaluation and questionnaire statements

Table A1 traces the priority and proposed content revision of each item; Table A2 reproduces the original statements; and Tables A3a–A3c report all ratings issued by the three expert judges J01–J03 for relevance, clarity, and adequacy. Item codes R01–R20 and domain assignments are consistent across the tables. The qualitative observations grouped as J04/J05 are considered separately and inform the interpretation of the expert ratings.

TABLE A1.

Conceptual review matrix by item

Item	Domain	Priority	Main difficulty	Decision to examine
R12	A	M	Absolute scope of "all"	Specify activities and sufficiency
R16	A	M	Equipment limitations versus software failures	Specify resource compatibility
R02	A	M	"All programs" and "without limitations"	Specify which programs are required
R09	A	M	Sharing is not equivalent to lack of access	Ask about effective availability
R19	C	M	Not everyone attends synchronous classes	Provide a not-applicable option or context
R03	C	M	Downloading and uploading are different actions	Specify sufficiency or separate activities
R07	C	H	Student internet outage versus platform failure	Separate incident types
R14	C	M	Absolute expression "always"	Specify frequency and period
R05	SC	M	Study space and distractions combined	Separate referents if appropriate
R15	SC	M	Institutional resources versus home study conditions	Confirm the subdomain
R20	SC	H	Different types of advice, mentoring, and support	Distinguish resources and purposes
R10	SC	H	Accessibility, ease of use, and "all" students	Focus on respondents' own experience
R18	E	H	Universal judgment about other students	Restrict the claim to observable experience
R08	E	H	Fairness and technological advantage combined	Separate procedural fairness from inequality
R13	E	M	Generalized statement about teaching practices	Specify courses or personal experience
R04	E	H	Comparing modalities rather than equity	Reassign or reformulate
R06	DC	H	"Confident" may refer to competence or familiarity	Specify perceived ability
R11	DC	H	Personal ability and external academic outcome combined	Separate ability from consequences
R01	DC	H	Anxiety differs from self-efficacy	Treat as a distinct affective component
R17	DC	M	"Without difficulty" is absolute	Qualify and contextualize

Source. Item-level content-review decisions documented in this study.

Note. Domain codes: A = material access; C = connectivity; SC = study conditions and institutional resources; E = perceived equity; DC = digital trust. Priority codes: H = high; M = medium.

TABLE A2

Original statements of the twenty items

Item	Domain	Original statement
R12	A	I have a personal device (computer, tablet or cell phone) with the appropriate characteristics to carry out all my academic activities online.
R16	A	The device I use for my virtual classes allows me to smoothly run the required educational platforms and software.
R02	A	I have access to all the programs and applications necessary for my academic training without limitations.
R09	A	I can do my academic assignments without sharing my device with others at home.
R19	C	I have an internet connection stable enough to attend online classes without interruptions.
R03	C	The speed of my internet connection allows me to download and upload files without significant difficulties.
R07	C	During online assessments, I don't have recurring disconnection issues or platform interruptions.
R14	C	I always have access to the Internet at the time I need to carry out virtual academic activities.
R05	SC	I have a suitable and distraction-free study space to participate in online classes and do my homework.
R15	SC	I have access to digital libraries, academic databases, and supplementary materials to enhance my online learning.
R20	SC	I have institutional support (such as consulting, technical support or tutorials) to solve technological or academic problems.

Item	Domain	Original statement
R10	SC	The educational platforms I use in my classes are accessible and easy to use for all students.
R18	E	I feel that access to online education is equitable for all students, no matter their socioeconomic or technological background.
R08	E	Online assessments are fair and do not favor students with better access to devices or the internet.
R13	E	I believe that teachers take into account differences in access to technology when designing their activities and assessments.
R04	E	The digital education model gives me the same learning opportunities as face-to-face education.
R06	DC	I feel confident using the technological platforms and tools required in my online studies.
R11	DC	When I face technology issues in online classes or assessments, I can resolve them without affecting my academic performance.
R01	DC	I feel anxiety or worry when I rely on technology to take exams or submit assignments online.
R17	DC	I rely on my digital skills to adapt to new platforms or educational tools without difficulty.

Source. Original questionnaire submitted for content evaluation.

Note. These statements reproduce the version submitted to the expert judges for content evaluation (with a minor grammatical correction to the English rendering of R01). Item codes correspond to the traceability matrix in Table A1.

APPENDIX A3.

Rating matrices from three expert judges, by item and criterion

Columns J01, J02, and J03 identify the ratings of the three expert judges included in the quantitative content-validity evaluation. For each of the twenty items, the tables report the judges' 1–4 scores, I-CVI, Aiken's V, κ*, and the conceptual review priority. Qualitative observations grouped as J04/J05 were analyzed separately and are not additional numerical ratings.

TABLE A3A.

Relevance: scores and indicators by item

Item	Domain	J01	J02	J03	I-CVI	Aiken's V	K*	Priority
R01	DC	1	2	1	0.000	0.111	-0.143	H
R02	A	4	4	4	1.000	1.000	1.000	M
R03	C	4	4	4	1.000	1.000	1.000	M
R04	E	1	2	1	0.000	0.111	-0.143	H
R05	SC	3	4	3	1.000	0.778	1.000	M
R06	DC	4	4	3	1.000	0.889	1.000	H
R07	C	2	2	2	0.000	0.333	-0.143	H
R08	E	4	4	4	1.000	1.000	1.000	H
R09	A	3	3	3	1.000	0.667	1.000	M
R10	SC	3	3	2	0.667	0.556	0.467	H
R11	DC	2	3	2	0.333	0.444	-0.067	H
R12	A	4	4	3	1.000	0.889	1.000	M
R13	E	4	4	4	1.000	1.000	1.000	M
R14	C	4	4	4	1.000	1.000	1.000	M
R15	SC	4	4	3	1.000	0.889	1.000	M
R16	A	4	4	4	1.000	1.000	1.000	M
R17	DC	4	4	4	1.000	1.000	1.000	M
R18	E	3	3	3	1.000	0.667	1.000	H
R19	C	4	4	4	1.000	1.000	1.000	M
R20	SC	4	4	3	1.000	0.889	1.000	H

Source. Coded rating matrix J01–J03 for relevance.

Note. J01–J03 rated items from 1 to 4. I-CVI is the proportion of scores ≥ 3; Aiken's V = Σ(r − 1)/9. κ is the adjusted chance-corrected agreement index described in the method. Priority: H = high; M = medium.*

TABLE A3B.

Clarity: scores and indicators by item

Item	Domain	J01	J02	J03	I-CVI	Aiken's V	K*	Priority
R01	DC	3	3	3	1.000	0.667	1.000	H
R02	A	2	2	2	0.000	0.333	-0.143	M
R03	C	3	4	3	1.000	0.778	1.000	M
R04	E	2	3	2	0.333	0.444	-0.067	H
R05	SC	2	3	2	0.333	0.444	-0.067	M
R06	DC	3	3	3	1.000	0.667	1.000	H
R07	C	2	2	2	0.000	0.333	-0.143	H
R08	E	2	2	2	0.000	0.333	-0.143	H
R09	A	3	3	3	1.000	0.667	1.000	M
R10	SC	1	2	1	0.000	0.111	-0.143	H
R11	DC	2	2	2	0.000	0.333	-0.143	H
R12	A	3	3	2	0.667	0.556	0.467	M
R13	E	3	3	3	1.000	0.667	1.000	M
R14	C	3	3	3	1.000	0.667	1.000	M
R15	SC	2	3	2	0.333	0.444	-0.067	M
R16	A	3	3	3	1.000	0.667	1.000	M
R17	DC	3	3	3	1.000	0.667	1.000	M
R18	E	2	2	2	0.000	0.333	-0.143	H
R19	C	3	3	3	1.000	0.667	1.000	M
R20	SC	2	2	2	0.000	0.333	-0.143	H

Source. Coded rating matrix J01–J03 for clarity.

Note. J01–J03 rated items from 1 to 4. I-CVI is the proportion of scores ≥ 3; Aiken's V = Σ(r − 1)/9. κ is the adjusted chance-corrected agreement index described in the method. Priority: H = high; M = medium.*

TABLE A3C.

Adequacy: scores and indicators by item

Item	Domain	J01	J02	J03	I-CVI	Aiken's V	K*	Priority
R01	DC	3	4	3	1.000	0.778	1.000	H
R02	A	3	3	3	1.000	0.667	1.000	M
R03	C	4	4	3	1.000	0.889	1.000	M
R04	E	2	2	2	0.000	0.333	-0.143	H
R05	SC	3	3	3	1.000	0.667	1.000	M
R06	DC	4	4	3	1.000	0.889	1.000	H
R07	C	3	3	2	0.667	0.556	0.467	H
R08	E	2	3	2	0.333	0.444	-0.067	H
R09	A	3	3	2	0.667	0.556	0.467	M
R10	SC	2	2	2	0.000	0.333	-0.143	H
R11	DC	3	3	2	0.667	0.556	0.467	H
R12	A	3	3	3	1.000	0.667	1.000	M
R13	E	3	4	3	1.000	0.778	1.000	M
R14	C	3	3	3	1.000	0.667	1.000	M
R15	SC	3	3	3	1.000	0.667	1.000	M
R16	A	4	4	3	1.000	0.889	1.000	M
R17	DC	4	4	3	1.000	0.889	1.000	M
R18	E	2	2	1	0.000	0.222	-0.143	H
R19	C	3	3	3	1.000	0.667	1.000	M
R20	SC	3	3	3	1.000	0.667	1.000	H

Source. Coded rating matrix J01–J03 for adequacy.

Note. J01–J03 rated items from 1 to 4. I-CVI is the proportion of scores ≥ 3; Aiken's V = Σ(r − 1)/9. κ is the adjusted chance-corrected agreement index described in the method. Priority: H = high; M = medium.*

DERECHOS DE AUTOR

Romero Cruz, M. (2026).

Este es un artículo de acceso abierto distribuido bajo la licencia Creative Commons Atribución-NoComercial-CompartirIgual 4.0 Internacional (CC BY-NC-SA 4.0). Esta licencia permite compartir y adaptar el material para fines no comerciales, siempre que se reconozcan adecuadamente la autoría y la publicación original, se indique si se realizaron cambios y las obras derivadas se distribuyan bajo la misma licencia.

COMO CITAR:

Romero Cruz, M. (2026). Validez de contenido de veinte reactivos sobre desigualdad digital universitaria mediante juicio de tres expertos. REDSA Revista Ecuatoriana de Desarrollo Social y Ambiental, 1(4) 169 - 190, https://doi.org/10.66646/redsa-2026/11nov55