Vol. 5, No. 3, 2025, 1-12. DOI: 10.47909/ijsmc.236

pro metrics

ORIGINAL ARTICLE

Validating personas for better communication: A structured model for low-resource contexts

Jefferson Lewis Velasco[1,*], Julio Monteiro Teixeira[1], Adilson Luiz Pinto[1]

[1] Universidade Federal de Santa Catarina, Brazil.
* Corresponding author
Email: jeffvelasco.crm@gmail.com. **ORCID:** https://orcid.org/0000-0002-1882-1785.

ABSTRACT

Objective. This study proposed a structured and accessible model for validating data-driven personas, with a focus on precision, particularly in low-resource communication and design contexts.

Design/Methodology/Approach. The study was grounded in the design science research (DSR) methodology and introduced a seven-step validation process centered on a double weighting system. The method was applied in a real-world case involving a Brazilian fintech company. Structured interviews, predefined scoring logic, and proximity-based analysis were used to measure the alignment between personas and real individuals.

Results/Discussion. The model effectively quantified the degree of alignment between respondents and personas through a persona approximation index termed the "FIT Index." The application of a heatmap-style matrix facilitated interpretation, revealing both strong and diffuse alignments. Lower scores yielded valuable insights, suggesting partial identification with the designated personas. This identification facilitated the implementation of informed adjustments to communication strategies.

Conclusions. The proposed model demonstrated efficacy in evaluating the representativeness of personas through the utilization of rudimentary instruments such as spreadsheets and structured questionnaires. It provided a replicable and scalable framework for organizations and scientific teams that aimed to enhance communication practices with limited technical resources.

Value. This study addressed a practical gap by presenting an empirical, proximity-based model for persona validation. The originality of this approach lay in its double weighting strategy, which combined attribute relevance and response alignment. This strategy contributed to the democratization of data-driven communication and design practices in low-resource contexts

Keywords: personas; data-driven design; validation; scientific communication; quantitative methods.

1. INTRODUCTION

The employment of personas as conduits between user data and design or marketing decisions has become pervasive, due to their capacity to imbue voluminous datasets with a human dimension and to enhance intradepartmental communication (Grudin & Pruitt, 2002).

Received: 28-01-2025. **Accepted:** 29-04-2025. **Published:** 14-05-2025.

How to cite: Velasco, J. L., Teixeira, J. M., & Pinto, A. L. (2025). Validating personas for better communication: A structured model for low-resource contexts. *Iberoamerican Journal of Science Measurement and Communication; 5*(3), 1-12. DOI: 10.47909/ijsmc.236

Beyond the realms of design and marketing, personas have also gained significant traction in scientific domains such as healthcare, economics, education, and public policy. In these fields, personas function as communication artifacts, serving to represent diverse groups of individuals and facilitating evidence-based decision-making processes (Chapman & Milham, 2006; McGinn & Kotamraju, 2008; Salminen *et al.,* 2020a). By translating complex data into human-centered narratives, personas empower interdisciplinary teams to enhance their comprehension and responsiveness to diverse needs, objectives, and behaviors. Furthermore, the mounting emphasis on data-informed decision-making in scientific, design, and innovation-related fields has led to an increased value placed on tools capable of integrating complex datasets and facilitating effective communication practices (Horvitz & Mitchell, 2010). Among these tools, data-driven personas have emerged as valuable artifacts that help organizations and researchers bridge the gap between statistical metrics and accessible narratives in communication (Jansen *et al.,* 2021). Situated at the nexus of measurement and narrative, personas empower organizations to make user-centered decisions by transforming statistical insights into comprehensible profiles that direct design, policy, and strategy (Jansen *et al.,* 2020). This study is dedicated to investigating the interfaces between scientific metrics and communication processes. It explores how data-driven personas contribute to more precise, inclusive, and evidence-based practices, especially in contexts with limited technical or analytical infrastructure.

While data-driven personas are said to guarantee objectivity and accuracy (Salminen *et al.,* 2020a), the construction of these personas typically necessitates the utilization of advanced technical skills and access to sophisticated statistical tools, thereby constraining their applicability in small organizations or projects with limited resources (Jansen *et al.,* 2022). Conversely, empirical or ad hoc methods for persona creation are characterized by ease of application (Gothelf & Seiden, 2021; Pruitt & Adlin, 2006), yet frequently exhibit deficiencies in precision, consistency, and validation mechanisms (Jansen *et al.,* 2022). This presents a methodological gap: while quantitative approaches are accurate but inaccessible, other approaches are accessible but potentially imprecise (Salminen *et al.,* 2020c). Furthermore, extant literature offers few practical and validated models for measuring the precision of Personas, especially those developed outside of rigorous data environments (Jansen *et al.,* 2022). This study explores the potential of data-driven personas to enhance the precision, inclusivity, and evidence-based nature of communication practices, particularly in contexts characterized by limited resources, such as scientific and organizational settings where access to technical infrastructure and financial resources may be constrained. The development and testing of an artifact that constitutes a framework for measuring the alignment between real people and the personas that aim to represent them was conducted in accordance with the design science research (DSR) methodology (Dresch *et al.,* 2020). The model was implemented in a real-world case to assess its effectiveness and generate actionable heuristics for future use. By emphasizing validation, this study contributes to the advancement of the reliability of personas in communication practices in general, and in design in particular, thereby supporting the broader democratization of data-driven methods.

2. THEORY

2.1. Concept of personas and their narrative role

The concept of personas was initially introduced by Cooper (2004) within the domain of human-computer interaction. These personas, defined as hypothetical archetypes of actual users, serve a pivotal role in the decision-making process by representing user objectives and behaviors. Garrett (2011) subsequently refined this concept by conceptualizing personas as instruments that synthesize mental models composed of expectations, experiences, and behaviors into accessible and communicative visual formats. These representations frequently amalgamate objective (e.g., demographic) and subjective (e.g., values and expectations) data into a multimodal narrative, employing text, images, and scenarios to convey insights and cultivate empathy within multidisciplinary

teams (Jansen *et al.,* 2021). Beyond their application in design, personas have been extensively adopted as communication artifacts across scientific and professional domains, including healthcare, economics, public policy, and education (McGinn & Kotamraju, 2008; Salminen *et al.,* 2020a). In such contexts, data-driven personas assume a pivotal role in translating intricate data into coherent narratives, thereby facilitating knowledge transfer, promoting evidence-based communication, and fostering the alignment of multidisciplinary stakeholders around common objectives. From a semiotic perspective, scholars such as Manovich (2002) have conceptualized interfaces as cultural forms that mediate the relationship between humans and data through symbolic constructs. This understanding suggests that personas can be interpreted as mediating interfaces in their own right, functioning at the intersection of quantitative measurement and narrative construction. This interpretative framing, explored by Velasco *et al.* (2023), positions data-driven personas as artifacts with the potential to transform objective datasets into humanized and communicable representations of people, fostering more inclusive, empathetic, and evidence-based communication practices within scientific, design, and organizational domains. As the communication function of personas becomes increasingly employed across scientific and applied fields, it is relevant to understand how they are constructed and validated from data in order to ensure their precision and reliability. The subsequent section will address the methodological foundations of data-driven persona creation.

2.2. Data-driven personas and their construction

Conventional personas, in contrast, are predicated on qualitative insights. Conversely, data-driven personas are constructed through systematic collection and statistical analysis of user data. As Gaiser *et al.* (2006) contend, scientifically grounded personas must be based on verifiable data —preferably both qualitative and quantitative— to ensure methodological rigor. Pruitt and Grudin (2003) similarly emphasize the importance of grounding Personas in actual behavioral patterns derived from structured datasets. The development of data-driven personas typically necessitates the implementation of sophisticated methodologies, including clustering algorithms and multivariate analysis, for the purpose of identifying meaningful user segments (Salminen *et al.,* 2020a). However, these approaches frequently necessitate substantial statistical expertise and computational infrastructure, which can impede accessibility for smaller organizations or projects with limited resources. Consequently, numerous small organizations and teams with limited resources have adopted empirical or ad hoc approaches for persona creation (Norman, 2008). These approaches frequently depend on assumptions, intuition, or informal discussions rather than on rigorous data analysis. While these methods are more accessible, they can result in oversimplified or misleading representations of the target audience. This is due to the lack of precision, consistency, and validation mechanisms present in data-driven alternatives (Güneş & Ercömert, 2022). The overreliance on imprecise personas introduces strategic risks, including the misalignment of products or services with actual user needs, the reinforcement of stereotypes, and the inefficient allocation of resources in design and communication processes (Guan *et al.,* 2024). These risks underscore the necessity for more accessible yet methodologically robust approaches to persona creation and validation.

2.3. Criteria and methods for persona validation

While not an obligatory component, the validation of personas is essential for ensuring the accuracy and representativeness of the profiles developed. Validation can occur in two distinct phases: either subsequent to persona creation or as an independent process. The purpose of validation is to verify that the personas reflect the characteristics of real individuals (Nielsen, 2019). A variety of dimensions have been advanced as criteria for evaluating the quality of Personas. Salminen *et al.* (2020b) identify a comprehensive set of perceptual criteria —credibility, clarity, completeness, consistency, empathy, similarity, and willingness to use— that facilitate the assessment of the usefulness, realism, and consistency of personas. However,

beyond perception, precision has emerged in the extant literature as a pivotal criterion for validation. Precision is defined as the extent to which a Persona accurately represents behaviors, needs, and contexts of real users. This concept has been emphasized by scholars such as Galliford *et al.* (2022), Kuipers *et al.* (2023), and Grudin and Pruitt (2002). Despite its importance, few works detail practical methodologies for measuring persona precision in accessible ways.

3. METHODS

3.1. Design science research

This study employs the DSR methodology, as proposed by Dresch *et al.* (2020), which is particularly well-suited to research endeavors aimed at developing and assessing practical solutions to pressing real-world problems. This study is founded on the theoretical framework of the "sciences of the artificial," as proposed by Simon and Laird (2019). DSR offers a structured yet flexible framework for developing and evaluating artifacts in both academic and applied contexts. This approach has been particularly effective in contexts where knowledge creation is intricately linked with problem-solving and innovation. The DSR approach employed in this research is iterative and includes key stages such as problem identification, literature review, artifact design, development, and evaluation. In this study, each phase contributed directly to the construction of a persona validation model. This study was initiated with the objective of identifying the challenges encountered by small organizations in adopting quantitative persona methods. It was subsequently developed into a validation artifact, which was then applied in a real-world case. The DSR facilitated a practical and theory-informed exploration of the problem. The subsequent section delineates the steps involved in the persona validation model, which are grounded in structured evaluation and proximity-based metrics.

3.2. Steps for persona validation

The persona validation model is designed to assess the extent to which a constructed persona accurately represents actual individuals within the target group. This process enhances the reliability of personas by introducing a quantitative method to evaluate their precision, defined as the degree of proximity between the attributes of a persona and the traits of real users. The validation process is structured and comprises seven steps. This seven-step process was developed and tested in this study and is presented in Figure 1.

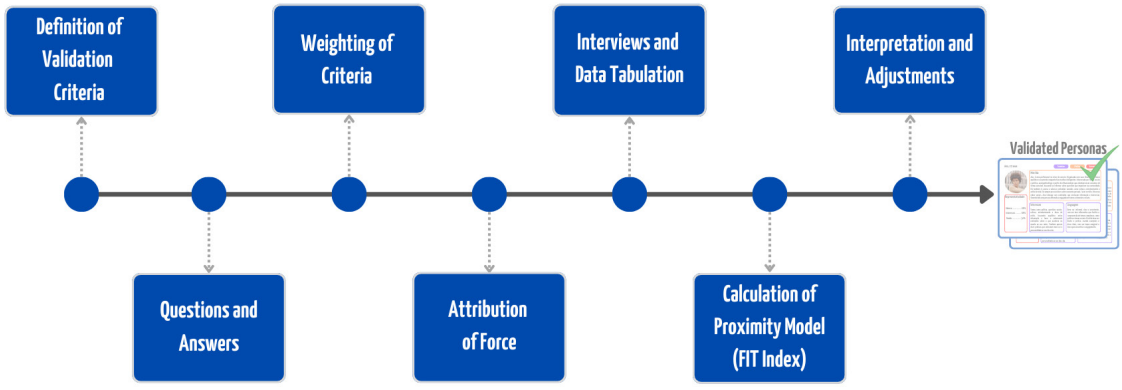

Figure 1. Steps for persona validation.

- **Step 1—Definition of validation criteria:** The process commences with the selection of validation criteria, wherein the attributes deemed essential for verifying the accuracy of the personas are identified. The aforementioned criteria are typically derived from the defining characteristics of the personas. It is imperative that these criteria are aligned with the strategic objectives of the organization and the project. The criteria may encompass demographic data, behavioral patterns, preferences, or

motivations that are most pertinent to the personas' narrative or the design project under consideration.

- **Step 2—Questions and answers:** For each validation criterion, a corresponding question is created. In the event that the criterion is associated with an attribute derived from numeric data (e.g., "age" or "number of children"), the inquiry may be posed as an open-ended question, thereby yielding a numeric variable. In instances where the question is linked to an attribute derived from qualitative data (e.g., "marital status"), it is imperative that the question is closed with a single answer selected from a predetermined list of options. The result is a questionnaire composed of questions formulated in a clear and objective manner to facilitate structured interviews or survey deployment.
- **Step 3—Weighting of criteria:** To each criterion, a weight is assigned to reflect its importance for the persona's strategic goals. This weightage also serves to indicate the relevance of each criterion to the precision of the personas. The scale of weights ranges from 1 to 4, with the maximum possible score (*Sm*) in the validation process being the summation of the highest score possible for each criterion. Weighting enables the method to account for the relative importance of certain attributes over others, prioritize what is strategically important for the persona's use, and allow for differentiated impact in the final assessment.
- **Step 4—Attribution of force:** Subsequently, the magnitude of each response is calculated, reflecting the extent to which a respondent's answer aligns with the intended persona's attribute. The range of force values is from 0 to 1 (e.g., 0, 0.5, 1), and the response score (*Sr*) is computed as the product of weight × force for each criterion. This facilitates the calibration of the criterion weight according to specific answers and their relation to particular personas.
- **Step 5—Interviews and data tabulation:** Upon completion of the questionnaire, interviews are conducted with individuals from the same general audience on which the personas were based. Their responses are systematically documented and compiled into tabulated data using a structured spreadsheet or form. The tabulation calculates the response score (Sr) for each respondent and each criterion, as demonstrated in Table 1.
- **Step 6—Calculation of proximity model (FIT Index):** The proximity model is the core metric of the model. For each interview, the total score (*St*) is calculated by summing the scores of all responses (*Sr*). This total is then divided by the maximum score (*Sm*) to yield the proximity index, also called the FIT Index (*Fp*), expressed as a percentage. This percentage indicates the extent to which the respondent's characteristics align with those of a designated persona. The final equation reads as $Fp = St / Sm \times 100$ where $St = \Sigma_{Sr}$.
- **Step 7—Interpretation and adjustments:** The results of the study are presented in a comparative matrix. In this matrix, rows correspond to interviewees and columns correspond to personas. The intersection values, also known as FIT indices, are color-coded to facilitate analysis, resulting in a heat map. The calculation of the average FIT value provides a quantitative indication of the degree to which the data align with the strategic use of a Persona. When the average FIT value is high, it suggests a strong alignment with the strategic use of a Persona. Conversely, when the average FIT value is low, it indicates a need for revision, additional segmentation, or data enrichment. In the event that interview results yield elevated FIT indices for specific personas and diminished indices for other groups, this signifies a heightened degree of approximation between the subject and particular personas as compared to others. However, this finding also suggests that the probability of this individual identifying with other personas is reduced, as is the likelihood of their aligning with products, services, or features designed for them.

This approach was developed around a double weighting system (weights and forces) to provide a mathematical representation of both personas and individuals within the selected validation criteria. In the subsequent section, the study delineates the implementation of this validation process in a real-world case study.

4. RESULTS

4.1. Case study context

To evaluate the practical application of the proposed persona validation model, the methodology was implemented in a real-world case study involving the Brazilian startup Na Sua Conta. This nascent fintech enterprise operates within the domain of payroll loan services, with the objective of providing transparent and accessible credit solutions, primarily for retirees and public servants. At the time of this research, the company had not yet launched its services commercially and was actively refining its business model, marketing strategies, and product development and offerings. Operating in a highly competitive market with 988 fintechs active in Brazil at the time —14% of which offered credit services— Na Sua Conta sought differentiation through trust, accessibility, and simplicity in the customer journey. The company's objective was twofold: first, to establish an intuitive hybrid user experience (combining digital and physical channels) and, second, to position itself as a trustworthy intermediary in the credit market. In this context, the application of the validation model was conducted as part of a consultancy provided by the Lemme Lab team. The objective of the consultancy was to build and validate personas based on market research data. The activities were meticulously structured, encompassing briefings, document analysis, persona construction, and validation. The present document contains the results of the validation process that was applied to a Portuguese-language dataset, which is available to the public at the following Uniform Resource Locator: https://tinyurl.com/Complete-Dataset. To elucidate the salient findings, the text incorporates illustrative extracts from the dataset.

4.2. Model application

4.2.1. Step 1—Definition of validation criteria

The initial phase in implementing the proposed persona validation model entailed the identification of data sources and the selection of criteria to be utilized in evaluating the representativeness of each persona. To support this, a dataset was collected through a structured online questionnaire distributed to individuals within the organization's target audience. A total of 13 participants were recruited for this study, and they all fit the general demographic and behavioral profile that was used to construct the personas. The participants in this study included retirees, pensioners, and public employees, who collectively represented the primary user segments of Na Sua Conta's payroll loan service. The selection of respondents was based on their availability and willingness to participate in a research initiative that was aligned with the company's strategic goals. The data collected were anonymous, ensuring the confidentiality and privacy of the study participants in accordance with established research ethics and legal requirements. The questionnaire was meticulously designed to mirror the pivotal attributes initially employed in the creation of the personas. These attributes were subsequently reviewed by the design team in collaboration with company stakeholders to determine which attributes would serve as validation criteria. In the final analysis, 27 validation criteria were selected for further consideration, a subset of which is depicted in Table 1. These criteria functioned as the foundation for the questionnaire employed in subsequent validation steps.

4.2.2. Step 2—Questions and answers

Subsequent to the establishment of the validation criteria, the subsequent phase entailed the operationalization of each criterion into a structured interview question. Each question was meticulously designed to elicit precise, pertinent responses that could be directly mapped to the traits delineated in the personas, employing clear and unambiguous language. Each qualitative inquiry was accompanied by a list of potential responses, meticulously designed to mirror a spectrum of respondent attitudes or behaviors associated with the criterion delineated in Table 1.

4.2.3. Step 3—Weighting of criteria

Subsequent to the establishment of validation criteria, the subsequent step entailed the allocation of weights to each criterion in accordance

Criterion	Attribute	Question
Cr1	Gender	What gender do you identify with?
Cr2	Age	What is your age?
Cr3	Occupation	Which of the two profiles below best describes your current occupation?
Cr4	Income	What income bracket do you fall into?
Cr5	Education	What is your educational background?
Cr6	Computer use	Do you use a computer in your daily life?
Cr7	Computer use function	If so, do you primarily use this computer for (see options)
Cr8	Computer use frequency	How often do you use the computer?
Cr9	Property	Is the property where you live rented, financed, or already paid off?
Cr10	Financial self-evaluation	Considering your financial situation, how do you see yourself? (see options)

Table 1. Excerpt validation criteria and questionnaire.

with its strategic pertinence to the personas and the objectives of Na Sua Conta. This weighting process was conducted in collaboration with the company's stakeholders and the design team to ensure alignment with the business context. The weights were assigned using a four-point scale (1-4), where 1 indicated low relevance, 2 indicated moderate relevance, 3 indicated high relevance, and 4 indicated very high relevance. The result was a weighted validation matrix, as presented in Table 2. This matrix enabled the calculation of the total response score (*Sr*).

Criterion	Weight	Answer	Force				
			Persona 1	Persona 2	Persona 3	Persona 4	Persona 5
C1	1	Female	1	0	1	0	1
		Male	0	1	0	1	0
		Other	0.5	0.5	0.5	0.5	0.5
C2	1	0 difference	1	1	1	1	1
		up to 3 years difference	0.5	0.5	0.5	0.5	0.5
		More than 3 years difference	0	0	0	0	0
C3	4	Public servant	1	1	0	0	0
		Formal employment contract	0	0	1	1	0
		Retired	0	0	0	0	1
C4	2	Up to R$ 3,000	1	0.5	0.5	0	1
		Between R$3,000 and R$5,000	0.5	1	1	0	0.5
		Between R$5,000 and R$7,000	0	0.5	0.5	0	0
		Between R$7,000 and R$10,000	0	0	0	0.5	0
		Between R$10,000 and R$15,000	0	0	0	1	0
		More than R$ 15,000	0	0	0	0.5	0

Table 2. Excerpt weighted validation matrix.

4.2.4. Step 4—Attribution of force

In the fourth step of the methodology, a force value between 0 and 1 was assigned to each possible answer, thereby indicating the extent to which it aligned with the Persona's trait for that attribute (Table 2). The aforementioned procedure enabled the calculation of the response scores (*Sr*) and the total scores (*St*) for each respondent. This established the foundation for computing the final FIT Index (*Fp*) in step 6.

4.2.5. Step 5—Interviews and data tabulation

The fifth step of the validation process entailed conducting interviews with users from the general target audience and organizing the collected data for subsequent analysis. The primary objective of this phase was to obtain consistent, structured responses from actual individuals and to apply the scoring system that had been defined in the preceding stages. To qualify for inclusion in the study, respondents were required to be affiliated with one

of the three distinct occupational categories delineated in the original persona segmentation framework: public servants, workers with formal employment contracts, or retirees. A total of 13 interviews were conducted, distributed as follows: 5 public servants, 4 individuals with formal employment contracts, and 4 retirees. The interviews were conducted individually and followed a standardized script based on the previously developed questionnaire. Each session was approximately 20 minutes in duration and was designed to ensure consistency in data collection across participants. The responses were meticulously documented and systematically arranged in a structured spreadsheet, with each answer assigned a force value according to its alignment with the corresponding persona attribute as defined in Step 4. Subsequently, these values were multiplied by the criterion weights that had been preassigned and determined on Step 3. This calculation resulted in individual response scores (*Sr*) for each question presented in Table 3.

Interview	*Sm*	Persona 1		Persona 2		Persona 3		Persona 4		Persona 5	
		Sr	*Fp* (%)	*Sr*	*Fp* (%)	*Sr*	*Fp* (%)	*Sr*	*Fp* (%)	*Sr*	*Fp* (%)
I1		14	53.85	15	57.69	16.5	63.46	20.5	78.85	13.5	51.92
I2		12.5	48.08	15.5	59.62	18	69.23	18.5	71.15	13.5	51.92
I3		14	53.85	13.5	51.92	19	73.08	18	69.23	14	53.85
I4		15	57.69	12.5	48.08	9.5	36.54	7	26.92	11.5	44.23
I5		20	76.92	18.5	71.15	14.5	55.77	13.5	51.92	15.5	59.62
I6		12	46.15	15	57.69	17.5	67.31	16	61.54	11	42.31
I7	26	14.5	55.77	15	57.69	13.5	51.92	11.5	44.23	16.5	63.46
I8		18.5	71.15	20	76.92	14.5	55.77	9.5	36.54	15.5	59.62
I9		12	46.15	13.75	52.88	15.5	59.62	18	69.23	12.5	48.08
I10		20.5	78.85	20.5	78.85	17	65.38	13	50.00	14	53.85
I11		14.5	55.77	14	53.85	13	50.00	14.5	55.77	19	73.08
I12		16	61.54	17.5	67.31	18	69.23	16.5	63.46	20.5	78.85
I13		8.5	32.69	8.5	32.69	8.5	32.69	7.5	28.85	11.5	44.23

Table 3. Tabulation results.

4.2.6. Step 6—Calculation of proximity model (FIT index)

Subsequent to the tabulation of responses, the subsequent step entailed the calculation of the overall alignment between each respondent and the personas being evaluated. The total score (*St*) for each respondent was computed by summing all individual response scores (*Sr*). This score serves as an indicator of the extent to which a respondent's responses align with the characteristics associated with a specific Persona. To facilitate a comparative analysis across respondents and ensure the normalization of the data, each total score was divided by the maximum possible score (*Sm*). This yielded the FIT Index (*Fp*), a proximity percentage that indicates how closely a respondent matches a specific persona, with values approaching 100% signaling strong alignment. Lower FIT scores, while indicating weaker alignment, still reflect a meaningful probability that the individual may partially identify with content, services, or strategies designed for other personas. In the context of this study, these percentages were organized in a matrix where rows represented respondents and columns represented personas. This matrix provided a comprehensive view of how well each respondent aligned with the five personas constructed during the project. Through a meticulous examination of the distribution of proximity scores, it was possible to discern patterns of convergence and divergence between profiles and real individuals. To facilitate interpretation, the results were represented in a heatmap-style table, where color gradients highlighted higher and lower alignment percentages (Table 4). This approach enabled the immediate identification of which Personas best matched each individual, facilitating the detection of strong associations as well as inconsistencies or ambiguous alignments.

Interview	Persona 1 *Fp* (%)	Persona 2 *Fp* (%)	Persona 3 *Fp* (%)	Persona 4 *Fp* (%)	Persona 5 *Fp* (%)
I1	53.85	57.69	63.46	78.85	51.92
I2	48.08	59.62	69.23	71.15	51.92
I3	53.85	51.92	73.08	69.23	53.85
I4	57.69	48.08	36.54	26.92	44.23
I5	76.92	71.15	55.77	51.92	59.62
I6	46.15	57.69	67.31	61.54	42.31
I7	55.77	57.69	51.92	44.23	63.46
I8	71.15	76.92	55.77	36.54	59.62
I9	46.15	52.88	59.62	69.23	48.08
I10	78.85	78.85	65.38	50.00	53.85
I11	55.77	53.85	50.00	55.77	73.08
I12	61.54	67.31	69.23	63.46	78.85
I13	32.69	32.69	32.69	28.85	44.23
Average (*Fp*)	**56.62**	**59.29**	**59.19**	**56.62**	**52.78**

Table 4. FIT score matrix.

4.2.7. Step 7—Interpretation and adjustments

The final step of the validation process entailed the interpretation of the FIT Index results and the extraction of actionable insights to refine the personas and guide future strategies. Utilizing the matrix of proximity scores as a foundation, the research team examined the alignment patterns between respondents and personas to evaluate the effectiveness and distinctiveness of each profile. The utilization of a heatmap visualization facilitated the expeditious identification of the personas that exhibited the highest and lowest levels of representation among the participants. In several cases, one persona emerged with a notably higher FIT score, signaling strong alignment and validating its core traits. For instance, in interviews I1, I5, I8, and I10, FIT values above 70% indicated a clear and confident match between the respondent and the target persona. However, certain participants exhibited moderate alignment across multiple personas, as opposed to a pronounced connection to a singular one. These cases prompted the team to reflect on potential overlaps between personas, leading to the identification of shared traits that may indicate the need to adjust boundaries between segments. In other scenarios, such as interviews I4 and I13, the consistent occurrence of low FIT scores indicated potential deficiencies, possibly attributable to underrepresentation in the initial construction or inconsistencies in the attribute selection process. It is noteworthy that even lower scores were not disregarded as irrelevant; rather, they were interpreted as suggesting a nondominant yet significant probability of identification with those personas. In light of the findings, minor adjustments were proposed for two personas whose alignment scores were lower and more diffuse.

5. DISCUSSION

The findings of this study underscore the significance of precision as a crucial component in persona validation. They also address a notable lacuna in the extant literature: the paucity of practical, accessible methods for evaluating the accuracy with which a persona represents its intended audience. While prior work has explored perceptual dimensions (Salminen *et al.,* 2020c), few studies have proposed objective models for measuring representativeness based on real user alignment, particularly in low-resource contexts. The model developed in this research demonstrates that persona validation can be methodologically rigorous without being technically complex. The process is characterized by its reliance on structured questionnaires, predefined scoring systems, and straightforward calculations. This approach enables individuals lacking expertise in the field to assess persona accuracy by employing only spreadsheet software and basic interview protocols. This addresses a common

barrier in small organizations and design teams with limited access to data science expertise or computational infrastructure. A salient feature of the model under scrutiny is its double weighting strategy, which evaluates two distinct components of the data. The first component is the relative importance of each persona attribute (weight), and the second component is the degree of alignment between respondents and expected traits (force). This layered logic enhances analytical sensitivity and enables more nuanced conclusions than binary matching. It is important to note that this weighting system is adaptable and could be tested in other contexts where multiple factors may influence the results.

The method's strengths lie in its scalability, transparency, and ease of implementation. The utilization of the matrix and the FIT Index offers a visual and quantifiable representation of alignment, thereby facilitating communication between technical and nontechnical stakeholders. Moreover, the implementation of heatmaps facilitates the iterative refinement of personas, which are informed by empirical feedback. However, the model is not without its limitations. The validity of the model is contingent upon the subjective decisions regarding which attributes to validate and how to assign weight and force values. If these decisions are not made with the utmost care, the model may be subject to bias. Additionally, while the model is effective for assessing precision, it does not evaluate other persona qualities such as emotional resonance, internal coherence, or storytelling strength. These additional qualities are important to consider in narrative-based design processes. Notwithstanding these limitations, the model contributes to the democratization of persona practices by offering a structured yet accessible framework that empowers small organizations to move beyond intuition-driven representations. The model thus stands as a novel, practical contribution to the field, offering a replicable method for increasing the reliability of personas in low-resource or applied contexts. The model's objective is to cultivate a culture of evidence-based persona development, with the aim of assisting design teams in aligning their strategies with the realities and diversity of their users.

6. CONCLUSIONS

This study proposes and tests a structured model for validating personas with a specific focus on precision —a dimension often neglected in extant literature and practice. The artifact was developed according to the DSR methodology, with the objective of ensuring accessibility for low-resource organizations while maintaining methodological rigor. The application of this approach in a real-world context demonstrated the feasibility of achieving accurate and evidence-based Persona validation using simplified tools, such as structured questionnaires, basic spreadsheets, and guided interpretation heuristics. The model introduces an original double weighting strategy, allowing for nuanced evaluations of user-persona alignment. The incorporation of attribute relevance and response alignment into the approach serves to enhance the analytical depth of traditional validation processes. Furthermore, the approach's extensibility to other applications in user research and segmentation, as well as prescriptive actions based on user or customer classification, is a notable strength. The research makes a significant contribution to the field by offering a validation framework that is not only effective but also replicable and scalable. By doing so, it supports the democratization of data-driven persona practices, enabling design teams in diverse organizational settings to build user representations that are both strategic and empirically grounded. Subsequent research endeavors may build upon this foundation by implementing the model in various sectors, integrating perceptual dimensions with precision scoring, and exploring automated or AI-supported adaptations of the framework.

Conflict of interest

The authors declare that there are no conflicts of interest.

Contribution statement

Conceptualization, methodology, formal analysis, investigation, and writing original draft: Jefferson Lewis Velasco.

Supervision: Julio Monteiro Teixeira.

Writing review: Adilson Luiz Pinto and Julio Monteiro Teixeira.

Statement of data consent

Data generated during the development of this study have been deposited in Google Drive, and it is accessible at https://tinyurl.com/Complete-Dataset. ●

REFERENCES

Chapman, C. N., & Milham, R. P. (2006). The personas' new clothes: Methodological and practical arguments against a popular method. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 50*(5), 634-36. https://doi.org/10.1177/154193120605000503

Cooper, A. (2004). *The inmates are running the asylum: Why high tech products drive us crazy and how to restore the sanity* (2nd Revised ed.). Sams Publishing.

Dresch, A., Lacerda, D. P., & Antunes Júnior, J. A. V. (2020). *Design science research: Método de Pesquisa para Avanço da Ciência e Tecnologia* (1ª Edição). Bookman.

Gaiser, B., Panke, S., & Arnold, P. (2006). *Community design — The personas approach* (pp. 520-25). Association for the Advancement of Computing in Education (AACE).

Galliford, F., Counsell, S., & Swift, S. (2022). Assessing persona accuracy: A user-centered model for persona validation. *Behavior & Information Technology, 41*(8), 1623-1637. https://doi.org/10.1080/0144929X.2021.1892824

Garrett, J. J. (2011). *The elements of user experience: User-centered design for the web and beyond* (2nd ed.). New Riders.

Gothelf, J., & Seiden, J. (2021). *Lean UX: Designing great products with agile teams* (3rd ed.). O'Reilly Media.

Grudin, J., & Pruitt, J. (2002). *Personas, participatory design and product development: An infrastructure for engagement* (pp. 144-152). PDC.

Guan, K. W., Salminen, J., Jung, S.-G., & Jansen, B. J. (2024). Leveraging personas for social impact: A review of their applications to social good in design. *International Journal of Human-Computer Interaction, 40*(19), 5569-84. https://doi.org/10.1080/10447318.2023.2247568

Güneş, S., & Ercömert, C. (2022). The ethical dimension of the persona concept. *Journal of Science, 10,* 147-158.

Horvitz, E., & Mitchell, T. (2010). *From data to knowledge to action: A global enabler for the 21st century: A white paper prepared for the Computing Community Consortium committee of the Computing Research Association*. http://cra.org/ccc/resources/ccc-led-whitepapers/

Jansen, B. J., Jung, S. G., Nielsen, L., Guan, K. W., & Salminen, J. (2022). How to create personas: Three persona creation methodologies with implications for practical employment. *Pacific Asia Journal of the Association for Information Systems, 14*(3), 1-28. https://doi.org/10.17705/1pais.14301

Jansen, B. J., Jung, S. G., Salminen, J., Guan, K. W., & Nielsen, L. (2021). Strengths and weaknesses of persona creation methods: Guidelines and opportunities for digital innovations. In *Proceedings of the 54th Hawaii International Conference on System Sciences* (p. 10). Hawaii USA.

Jansen, B. J., Salminen, J. O., & Jung, S.-G. (2020). Data-driven personas for enhanced user understanding: Combining empathy with rationality for better insights to analytics. *Data and Information Management, 4*(1), 1-17. https://doi.org/10.2478/dim-2020-0005

Kuipers, R., Salminen, J., & Jansen, B. J. (2023). A framework for evaluating persona effectiveness in digital service design. *Journal of Documentation, 79*(5), 1163-1185. https://doi.org/10.1108/JD-01-2023-0012

Manovich, L. (2002). *The Language of New Media* (Revised ed. edição). MIT Press.

McGinn, J. J., & Kotamraju, N. (2008). *Data-driven persona development. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '08* (pp. 1521-1524). Association for Computing Machinery.

Nielsen, L. (2019). *Personas — User focused design*. Springer London.

Norman, D. (2008). Ad-hoc personas & empathetic focus. *Don Norman's JND.Org*. https://jnd.org/ad-hoc-personas-empathetic-focus/

Pruitt, J., & Adlin, T. (2006). *The persona lifecycle: keeping people in mind throughout product design* (1st ed.). Morgan Kaufmann.

Pruitt, J., & Grudin, J. (2003). Personas: Practice and theory. In *Proceedings of the 2003 conference on Designing for user experiences, DUX '03* (pp. 1-15). Association for Computing Machinery.

Salminen, J., Guan, K., Jung, S.-G., Chowdhury, S. A., & Jansen, B. J. (2020a). A literature review of quantitative persona creation. In *Proceedings of the 2020 CHI conference on human factors in computing systems* (pp. 1-14). ACM.

Salminen, J., Kinnunen, A., Lähteenmäki, S., Jung, S., & Jansen, B. J. (2020b). Perception gap analysis: Data-driven persona generation and persona perception. *International Journal of Human-Computer Studies, 138,* Article 102409. https://doi.org/10.1016/j.ijhcs.2019.102409

Salminen, J., Santos, J. M., Kwak, H., An, J., Jung, S.-G., & Jansen, B. J. (2020c). Persona perception scale: Development and exploratory validation of an instrument for evaluating individuals' perceptions of personas. *International Journal of Human-Computer Studies, 141,* Article 102437. https://doi.org/10.1016/10.1016/j.ijhcs.2020.102437

Simon, H. A., & Laird, J. E. (2019). *The sciences of the artificial, reissue of the third edition with a new introduction by John Laird*. The MIT Press.

Velasco, J., Picalho, A., Fonseca, R., Viola, H., Teixeira, J., Fadel, L., & Pinto, M. (2023). Data driven personas: Uma Interface de Mediação Entre Dados e Usuários. In *Perspectivas em Engenharia, Mídias e Gestão do Conhecimento* (p. 398, Vol. 4).