Contemporary Educational Technology, 2024, **16**(2), ep496
ISSN: 1309-517X (Online)

OPEN ACCESS **Research Article**

Applying Kano's two-factor theory to prioritize learning analytics dashboard features for learning technology designers

Tobias Alexander Bang Tretow-Fish [1*]
0000-0002-2604-0153

Md Saifuddin Khalid [1]
0000-0002-3731-2564

[1] DTU Compute, Lyngby, DENMARK
[*] Corresponding author: tabtr@dtu.dk

Citation: Tretow-Fish, T. A. B., & Khalid, S. M. (2024). Applying Kano's two-factor theory to prioritize learning analytics dashboard features for learning technology designers. *Contemporary Educational Technology, 16*(2), ep496. https://doi.org/10.30935/cedtech/14286

ARTICLE INFO

Received: 29 Aug 2023
Accepted: 17 Jan 2024

ABSTRACT

Existing methods for software requirements elicitation, five-point Likert scales and voting methods for requirements prioritization, and usability and user experience evaluation methods do not enable prioritizing the learning analytics dashboard requirements. Inspired by management and product design field, this research applies Kano's two-factor theory to prioritize the features of learning analytics dashboards (LADs) of adaptive learning platform (ALP) called Rhapsode™ learner, based on students' perceived usefulness to support designers' decision-making. Comparing usability and user experience methods for evaluating LAD features, this paper contributes with the protocol and a case applying Kano method for evaluating the perceived importance of the dashboards in ALP. The paper applies Kano's two-factor questionnaire on the 13 LADs features of Rhapsode™ learner. Responses from 17 students are collected using a questionnaire, which is used to showcase the strength of the two-factor theory through five tabular and graphical techniques. Through these five tabular and graphical techniques, we demonstrate the application and usefulness of the method as designers and management are often carried away by the possibilities of insights instead of actual usefulness. The results revealed a variation in the categorization of LADs depending on the technique employed. As the complexity of the techniques increases, additional factors that indicate data uncertainty are gradually incorporated, clearly highlighting the growing requirement for data. In the case of RhapsodeTM learner platform, results based on the students responses show that 11 of 13 LADs being excluded due to low significance level in categorization (technique 1) and low response rate.

Keywords: adaptive learning platforms, Kano's two-factor theory, learning analytics dashboard, design methods

INTRODUCTION

Designers of learning technology face a difficult task, as they try to improve elements of Learning Technology to better suit teachers' and students' needs. Learning analytics dashboards (LADs) are visual displays that provide overviews and insights about learners, the learning process, and the learning context (Schwendimann et al., 2017; Sedrakyan et al., 2019). LADs are designed for teachers and students to adjust students' learning strategies, motivate students, deepen students' understanding of subjects, and identify learning goals and contents requiring further facilitation by teachers (Verbert et al., 2020). Though LADs promise to improve and enrich the learning processes of students (Bodily et al, 2018), several challenges are associated with the technology such as slow adoption of the technology, the need for users to be directly involved in the data processing, and LADs being difficult to interpret (Viera et al., 2018). Nonetheless, LADs allow users to engage with data for tasks like confirming hypotheses, exploring scenarios, categorizing, and spotting interesting features. Furthermore, LADs spans from simple calculations to advanced techniques like data mining, guided by user input (Viera et al., 2018). Adaptive learning platforms (ALPs) similarly provides a personalization of the learning process of students. ALPs accomplish this as both a technology and an approach by supporting individual learners with personalized course content, instructions, and feedback. ALPs enable learning environments to dynamically provide ways of presenting course content, instruction and feedback based on the learner's understanding of the learning content by responding to embedded assessments or preferences. When students interact with ALP core elements through LADs, the information on LAD becomes granularized and dynamically changes as learners process learning. ALPs identify students' weaknesses, recommending students to practice, while providing teachers with feedback on students gaps in knowledge through LADs.

However, designing LADs of ALPs is a not a simple task and it is defined by the design context. Designing can thus become extremely complex, encompassing not only students' and teachers' perspectives but also management's (Ahn et al., 2019, p. 74). In dealing with these contextual demands, designers can apply one of two approaches. The first approach designers can apply involves qualitative methods that use design principles to guide the data collection, the analysis of findings with abstract concepts, and the translation of findings into concrete design changes. These design principles can be either usability design principles (Benyon, p. 116-122), the four-level evaluation model (Kirkpatrick & Kirkpatrick, 2006), visualization design principles (Few, 2006) or others. This first approach is typically very time-consuming and expensive, providing a lot of complex data. Studies that apply the first approach include Park and Jo (2015) and Yoo and Jin (2020). In these two studies, they develop LAD prototypes using interviewing methods building upon design principles. They refine and test prototypes, including surveys and usability studies to improve the product. Through a thorough design and assessment process, these studies identify issues experienced by the users of LADs, but the process is time-consuming. The second approach entails quantitative methods that apply one-dimensional usability or user experience methods to identify issues experienced by LAD users (Hornbæk, 2006; Salas et al., 2019). These quantitative methods seldom give feedback on how issues related to LADs might be solved (Nakamura et al., 2017). If these issues are to be solved, designers must follow up with methods from the first approach in studies such as Iordache and Pribeanu (2009) and Lesemann et al. (2007), which again leaves us with a time-consuming process.

In the absence of resource constraints, prioritization may seem unnecessary in the design of ALP and their LADs. However, the presumption of unlimited resources is seldom, if ever, realistic.

In any case, tried and tested evaluation methods for LADs on ALPs, which provide prioritized and actionable design choices are sparse. Evaluation methods, such as heuristic evaluations receive criticism for reliability and effectiveness of measures as well as the need for subjective interpretations of results from evaluators (de Kock et al., 2009). By providing a method that prioritizes LAD features for designers to improve or remove from the platform, good design does not solely have to rely on the subjective interpretations of a good designer. Additionally, as good design depends on access to a wealth of evaluation methods, we present the case of Kano's two-factor theory with five different ways to present the prioritization of dashboards. Thereby, the contribution of this study is a quantitative method that gives suggestions to designers on how LAD issues might be solved. Furthermore, Kano's two-factor theory provides a methodological rigor, that traditional prioritization techniques such as one-dimensional Likert scales and dot voting does not provide with their proven quick and dirty approach.

Crucially, the novelty of the proposed method provides designers with a prioritization of the specific features of LADs rather than determining the effectiveness, efficiency, and user-satisfaction of LAD and subsequently requiring interpretation by designers to transform the results into actionable design interventions (de Kock et al., 2009). Additionally, the novelty of this study also includes the application of Kano's two-factor method in evaluating and prioritizing LAD features based on students' responses to a developed Kano's two-factor method questionnaire (Berger et al., 1993).

The research question addressed is how LAD features for learners can be prioritized for redesign through the application of Kano's two-factor method.

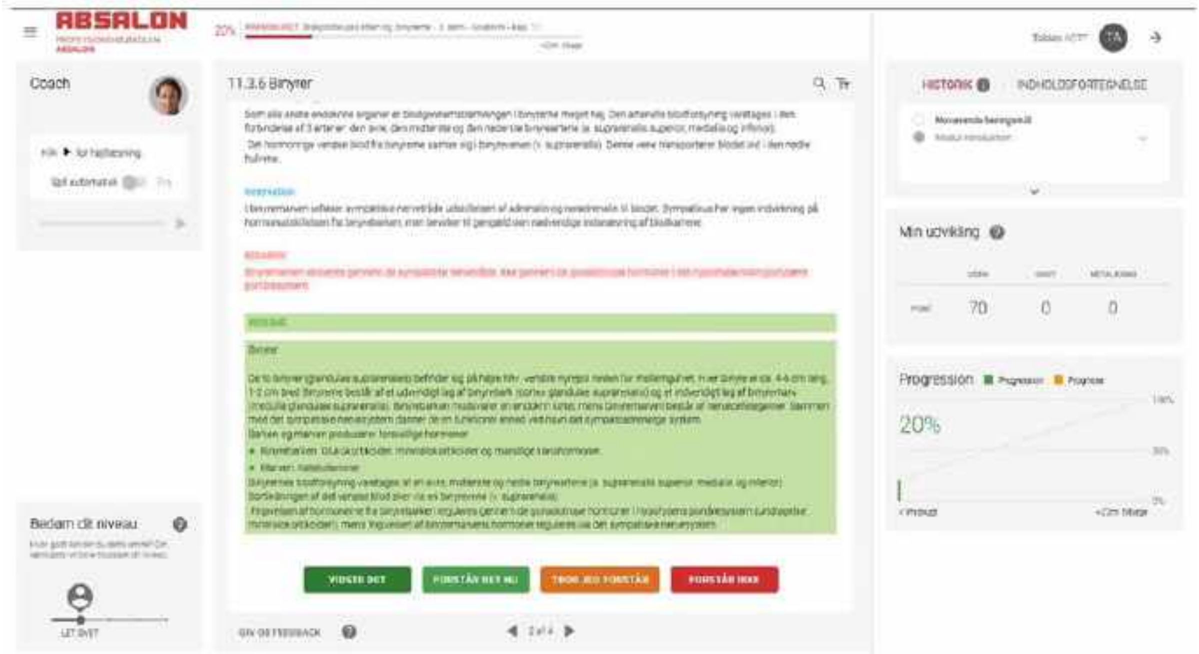

Figure 1. Screenshot of Rhapsode™ learner layer (Area9, 2024)

To answer this question, we apply this method in the context of a Danish higher educational institution using the company Area9's Rhapsode™ platform.

As well as providing further context information, this study provides a state-of-the-art account of the evaluation of ALPs and of current methods for the prioritization of LADs for designers. The methods section reports on the development of our Kano's two-factor method questionnaire, on the data collection process, and on the application of Kano's two-factor method to fulfil the aims of this study. The results section presents the users' responses to the questionnaire and demonstrates the various ways in which LAD features can be visualized for prioritization by the application of Kano's two-factor method. The discussion reflects on the findings and the method. The conclusion summarizes the study's central contributions and addresses its limitations.

THE CONTEXT

In the fall of 2020, the nursing education program at a Danish university college introduced ALP Rhapsode™ as part of a three-year research project called BLINDED. The platform was adopted in the natural sciences and related subjects in a flipped classroom setting. The initial introduction of the platform consisted of a single day of teaching, where researchers and educators presented the teaching content, the platform, and its activities to 105 first-semester nursing students.

The students were grouped into four classes taught by three educators. Rhapsode™ consists of three layers: Rhapsode™ curator, where learning activities are constructed, Rhapsode™ educator, where teachers assess students' preparation activity, and Rhapsode™ learner, where students complete their learning activities. This study focuses solely on students' experiences using the Rhapsode™ learner layer to interact with learning activities before a lecture. We used the questionnaire to examine students' perceptions of LADs features, elements, and sub-elements. Subsequently, we applied Kano's two-factor theory as a quantitative method to prioritize LAD features for improvement in a design and development process - that is to evaluate which LAD features students find least or most useful. **Figure 1** shows screenshot of Rhapsode™ learner layer.

STATE OF-THE-ART

This section identifies the scope of the contribution of this study by addressing how the design of ALPs is conducted through non-prioritizing and prioritizing methods fleshing out Kano's two-factor method.

Evaluation of Adaptive Learning Platforms

An examination of the literature on the evaluation of ALPs designs shows that this is primarily done through ontology and framework evaluations (Tretow-Fish & Khalid, 2022), which is of limited use to designers or developers in improving existing platforms. Plenty of work has been done on ALPs with learners from the health professions, as seen in literature studies by Fontaine et al. (2019) and Xie et al. (2019). Xie et al. (2019) reveals that 46% of the existing literature on ALPs for higher education focuses on the learner perspective. Additionally, the application of ALPs for health professionals' and students' education has also been tried and tested, as seen in another literature study by Fontaine et al. (2019). In other words, the novelty of evaluating ALPs in our study lies not in the data being on users' experiences or the nursing education context. Rather it lies in how this specific data can be applied in prioritizing LAD features from a design or development perspective.

Non-Prioritizing Methods for Design of Learning Analytics Dashboard

Qualitative approaches have been widely applied in the design of LADs in the field of human-computer interaction. Qualitative evaluation of LADs can give us insights into unquantifiable phenomena valuable for the design process (Earnshaw et al., 2018). Importantly, qualitative methods do not result in a prioritized list of features to be addressed. Furthermore, qualitative evaluations can be costly regarding time expenditure, the number of participants, and budget constraints. For these reasons, this study will not examine qualitative approaches extensively. However, we recognize that from methods such as interviews (Roberts et al., 2017), we can attain detailed insights about LADs and from methods such as think-aloud protocols and focus-group methods (Bodily et al., 2018), we can gain valuable information on student-facing LADs in use.

In the evaluation of technology, quantitative methods have the inherent limitation of applying one-dimensional responses regarding users' experience of satisfaction (user experience) and/or functionality (usability), as presented by Sauro and Lewis (2016). Typically, these do not provide a prioritized list to guide designers in improving or removing features of the platforms. Nevertheless, quantitative methods have successfully been applied to support the design process of student-facing LADs (Earnshaw et al., 2018; Sauro & Lewis, 2016). None of these quantitative and qualitative evaluation methods provides a prioritized or ranked list of LADs or LAD features accordingly to the user's preferences for design purposes.

Prioritization Methods for Design of Learning Analytics Dashboards

Prioritization methods for designers are available, but these methods have distinct limitations. Prioritization can be accomplished with self-developed instruments such as the satisfaction questionnaire presented by Kim et al. (2016). They developed an instrument to measure and prioritize students' satisfaction with goal orientation, information usefulness, visual effectiveness, appropriateness of the visual representation, and user-friendliness of a technology. For self-developed instruments, it is important to validate the methods' results. Therefore, reliability tests, such as Cronbach's alpha, are especially important for the internal consistency of the responses. While a self-developed instrument is a good option, implementing the instrument requires at least one iteration of testing of the instrument, thereby increasing its implementation costs.

Kano's two-factor method

A promising method for evaluation and prioritization is Kano's two-factor method. This method examines the functional and dysfunctional quality of a product features from the users' perspective. It can also include users' assessment of the importance of the features (Witell et al., 2013), thus making it a two-dimensional evaluation method. It is precisely the two-dimensionality that enables designers to prioritize features.

Prof. Nariaki Kano and colleagues developed a method for categorizing product/service features into quality categories based on customer feedback from questionnaires (Berger et al., 1993) to evaluate users' perception of products' quality. The method classifies customer features into categories, making it possible for designers to prioritize customer responses. The responses are collected through a survey with two questions on each feature or element of the product. One question is a functional question, assessing how the respondent feels about the feature being part of the product; the other question is a dysfunctional question, assessing how the respondent feels about the feature not being part of the product. An example of the questions is given from this study's questionnaire in evaluating feature no. 11:

> How do you feel about the 'metacognition element' of Rhapsode being on the platform? (functional).
>
> How do you feel about the 'metacognitions element' in Rhapsode NOT being on the platform? (dysfunctional).

Respondents are then able to answer either "I like it", "it must-be so", "I'm neutral about it", "I can live with it", or "I dislike it". Through the calculations, tabular, and graphical techniques explained below, the features can be listed in prioritization categories, which tell the designer, which are most important to improve or remove. These categories are *attractive* (A), which defines a feature as satisfactory when achieved but does not result in customer dissatisfaction when not fulfilled; *must-be* (M), which defines the requirements as users exhibiting discontent when deprived of such attributes, while their presence fails to notably enhance satisfaction; *one-dimensional* (O), which defines the requirements as satisfactory when fulfilled and results in dissatisfaction for customers when not fulfilled; *indifferent* (I), which defines the requirements as neither good nor bad; *reverse* (R), which defines the requirements as the opposite of our expectations. This concept acknowledges the inherent heterogeneity among customers, as not all share uniform preferences or aversions; and *questionable* (Q), which defines responses as being unreliable (Kano et al., 1984). Kano method is widely applied in various fields, enabling a better understanding of customer requirements and prioritizing development/design activities.

The method has been widely used and applied in multiple different areas (Löfgren & Witell, 2008) such as trade-off processes (Löfgren & Witell, 2008), product packaging (Löfgren, 2005), and information systems quality and design (Khalid et al., 2008). Since 1984 it has been successful as a product feature evaluation tool and continuously improved (Löfgren & Witell, 2008; Witell et al., 2013). It has also seen use in the general area of learning technology, which has been applied in studies such as Chen and Hsu (2019). They applied Kano in an empirical study of multimedia learning. Hereby, they obtained a prioritized list of what a designer or developer could focus on in the next iteration of the product's development. Chen and Hsu's (2019) clear-cut categorization of the features shows the number of indifferent respondents but does not quantify the disagreement with the categorization. Without this quantification, it is difficult for the designers to reflect on the prioritization. Another example of use is by Bauk et al. (2014). They applied Kano method to evaluate the University Mediterranean's blended learning system and included the many indifferent responses they received.

While studies have applied Kano method in a learning and learning technology context (Chen et al., 2022), to the best of our knowledge, none have applied it in the context of ALPs or on LADs. Thus, in assessing the application of Kano's theory to ALPs, this study devises an instrument designed to evaluate and prioritize the various features of LAD Rhapsode™ student platform. In doing so, we apply and assess five separate tabular and graphical techniques developed in previous applications of Kano's theory (Berger et al, 1993). The study then compares the results of the five alternative techniques regarding prioritization and decision-making for designers.

METHODS

The following subsections describe how the survey instrument was produced, how data was collected, and which formulas were used for producing the five tabular and graphical techniques used for prioritizing the features of Rhapsodes™ LADs.

Instrument Preparation & Data Collection

The first iteration of the survey instrument produced by the researchers contained 39-questions on the 13 features of Rhapsode™ learner, three questions per feature, one revealing whether the feature is functional, one revealing whether the feature is dysfunctional, and one on the importance of the said feature. The survey was tested with two students in 30-minute recorded concurrent think-aloud sessions.

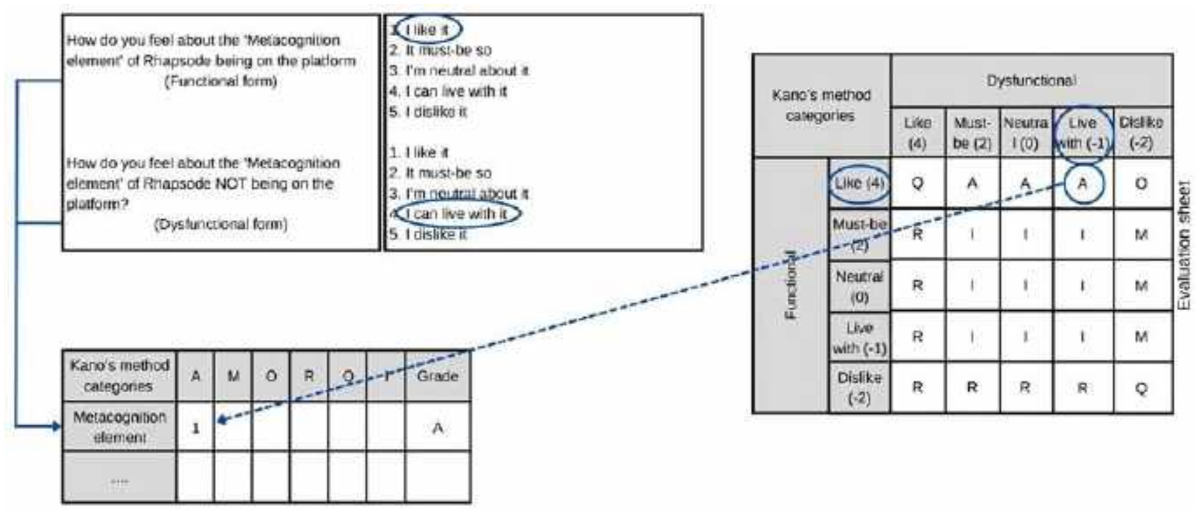

Figure 2. Computation procedure of applying Kano questionnaire & evaluation table (as described in Kano's method: A: Attractive; M: Must-be; O: One-dimensional; I: Indifferent; R: Reverse; & Q: Questionable) (inspired by Berger et al., 1993)

Subsequently, the wording was revised on the survey. Eventually, Kano's questionnaire was divided into two parts:

(1) the two-factor questionnaire for 13 LADs with 26 questions and

(2) the importance rating of 13 LADs on a nine-point Likert scale (not important to extremely important).

The survey was conducted using Microsoft Forms, and 17 (n=105) voluntary responses from second-semester nursing students were exported as CSV files to be treated in the statistic software R. The survey was distributed to students after they had familiarized themselves with Rhapsode™, spending an average of 59.8 minutes (standard deviation [SD]=39.6 minutes).

Data Analysis: Formulas & Evaluation of LAD Features

This section presents

(1) the conversion of student responses to Kano method categories and

(2) the five alternative techniques of analysis.

To effectively categorize the human perception of the features or requirements, the utilization of the Kano questionnaire is imperative. The questionnaire comprises of a set of paired questions for each of the features. The first question is framed positively, for example, "how you feel about the 'metacognition element' of Rhapsode being on the platform"–referred to as the functional question. The student respondents then choose from five options, expressing varying degrees of satisfaction. The positive formulation gauges the respondent's sentiment when the specified requirement is met. Conversely, the second question is formulated negatively, for instance, "how do you feel about the 'metacognitions element' in Rhapsode NOT being on the platform?"–referred to as the dysfunctional question. Respondents again provide feedback on a scale from like to dislike. This negative framing reveals the impact of non-compliance with the requirement on the respondent's satisfaction. The combination of positive and negative questions results in 25 possible response combinations (5×5). To assign a requirement to a specific category, an Evaluation Sheet encompassing all these combinations is employed. **Figure 2** illustrates the systematic process of categorization. Berger et al. (1993) and many other scholars applied unequal intervals of scores (i.e., 4, 2, 0, -1, and -2) and higher points for satisfaction as opposed to dissatisfaction, while some scholars (Madzík et al., 2019) applied the 1-5 Likert scale scores.

Applying the procedure of **Figure 2**, all student responses were categorized into one of the six categories for prioritization. These categories were attractive, must-be, one-dimensional, indifferent, reverse, and questionable. The responses under the "questionable" category are excluded due to the underlying interpretation that one cannot like and dislike for both the inclusion and exclusion of a feature.

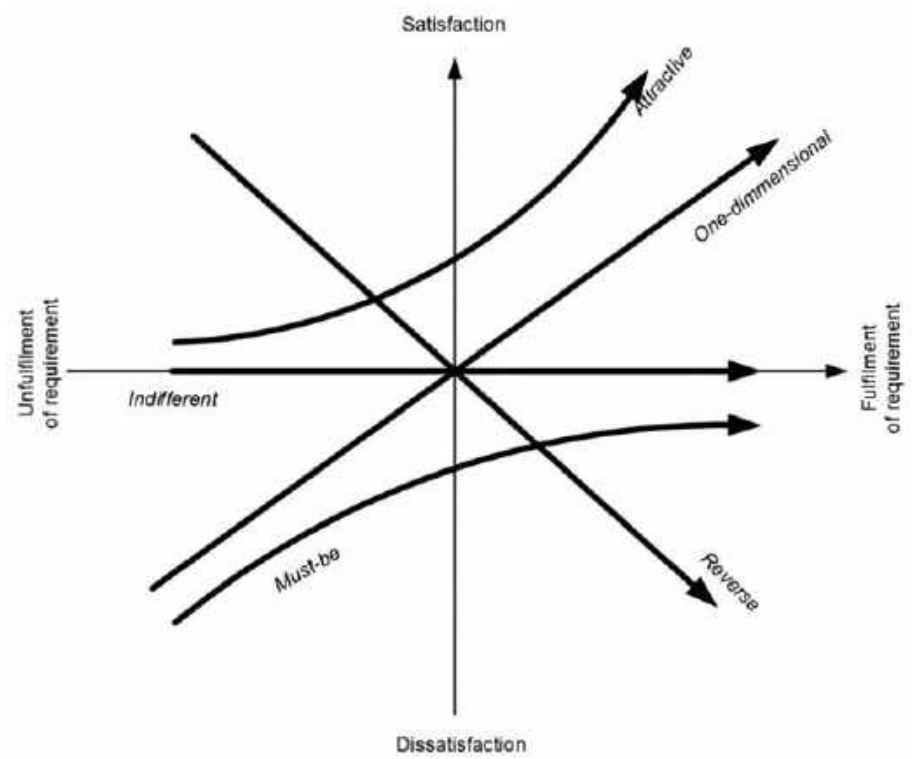

Figure 3. Kano model (Kano et al., 1984)

Subsequently, the five separate techniques of analysis were applied and assessed to compare the usefulness of their results for prioritization and design decision-making.

To enhance comprehension, it is common to represent requirements visually. The visual representation is typically done using an *x-y* coordinate system, where the *x*-axis signifies the extent of requirement fulfilment (ranging from completely unfulfilled to fully fulfilled), and the *y*-axis represents the degree of (dis)satisfaction (ranging from complete dissatisfaction to zero dissatisfaction, and from zero satisfaction to complete satisfaction). Specific categories of requirements are depicted as curves within coordinate system (**Figure 3**).

Categorization of quality attributes using frequency & evaluating significance using Fong's test

The first technique of analysis was to create a tabular summary of the frequency of quality categories (i.e., A, M, etc.) for each of LAD features. The table contains the most frequent student category (a) and the second most frequent student category (b), as well as the Fong test results (Fong, 1996) for each feature. Fong test was applied to assess whether the attributed Kano category (i.e., the category with the highest frequency of student responses) was statistically significant. How Fong test was calculated is shown in Eq. (1). For *n* responses, if the inequality was untrue, then the category was a significant statical classification (Fong, 1996).

$$a - b < 1.65 * \sqrt{\frac{a+b*(2n-a-b)}{2n}}. \qquad (1)$$

From the first technique, we acquire a categorization for designers. To assess the categorization of LADs, another technique of analysis can be applied.

Categorization of quality attributes using category strength scores & stated importance scores

The second technique of analysis was to create a tabular summary of classification agreement, category strength (CS; Eq. [2]), total strength (TS; Eq. [3]), and the average score on students' self-reported importance (scored 1-9) for each of LAD features. The classification agreement is the percent frequency of the category with the highest frequency. Eq. (2) is defined, as follows:

$$CS = Percentage\ score\ of\ most\ frequent\ category - percentage\ score\ of\ 2nd\ most\ frequent\ category. \qquad (2)$$

Eq. (3) is defined, as follows:

$$CTS = Percentage\ score\ of\ attractive + percentage\ score\ of\ one - dimension + percentage\ score\ of\ must - be. \qquad (3)$$

(Löfgren, 2005). If CS>5%, the feature unmistakably belongs to its classification. If TS>50%, we interpret LAD feature as important to the respondents. If TS<50% there was insufficient evidence for the feature to be considered useful (Löfgren 2005).

Categorization of quality attributes using "better" & "worse" scores

The third technique of analysis is a sorting of the prioritized list according to either the 'worse' scores (Eq. [4]) or 'better' scores (Eq. [5]) in combination with the stated importance of users. This ranked list makes it possible to emphasize what is most important for the specific designer's' design process (lowest 'worse' scores or highest 'better' scores).

Calculating the features' 'worse' scores (Eq. [4]) and 'better' scores (Eq. [5]) were done to present them as numerical values in the table. Eq. (4) is defined, as follows:

$$Worse = \frac{O+M}{A+O+M+I}. \tag{4}$$

Eq. (5) is defined, as follows:

$$Better = \frac{A+O}{A+O+M+I}. \tag{5}$$

The 'worse' score indicates the degree to which an LAD feature contributes to dissatisfaction, while the 'better' score indicates the degree to which the same feature contributes to satisfaction (Berger et al., 1993, p. 18). The maximum value for 'better' and 'worse' scores is 1, where 1 indicates a high impact on customer satisfaction, and 0 indicates a minor impact (Matzler et al., 1996). While CS and TS extends the options of designers to assess the categorization's strength and reliability the presentation of the better-worse scores in a numerical format is not as accessible. The following technique leaves a decision to be made by the designer as to what is most important for the design decision-making process.

Categorization of quality attributes by "better" (y-axis) & "worse" (x-axis) score plotting

In the fourth technique, a combination of both 'better' and 'worse' scores is explored.

The fourth technique of analysis displayed the 'better' and 'worse' scores in a graphical representation. In this representation, 'worse' was plotted on the *x*-axis and 'better' was plotted on the *y*-axis. This better-worse plot enables the features to be mapped visually into the four categories 'attractive', 'one-dimensional', 'indifferent', and 'must-be'. While the better-worse plot enables an alternate computation of categorizing LADs, it does not contain the assessment of uncertainty related to the categorizations. This will be included in the fifth analysis technique.

Categorization by calculating & plotting functional & dysfunctional weighted average scores

The fifth technique of analysis applied an alternative approach to the conversion of student responses into Kano categories. Looking at previously given example of the student response conversion, if a student responded 'like' to a feature's functional question, the 'like' response is converted to its numerical value of four. If the same student responded 'dislike' to the dysfunctional question, this response is converted to its numerical value of -2. For each of the features, the students' responses to the functional (Y) questions were added together, as were the students' responses for the dysfunctional (X) questions (Berger et al., 1993). The numerical values for the functional and dysfunctional questions were then used to create a tabular summary. The table contains the X_{ave} (dysfunctional attribute; Eq. [6]), Y_{ave} (functional attribute; Eq. [7]), W_{ave} (perceived importance of each feature; Eq. [8]), X_{wave} (weighted importance of dysfunctional attribute; Eq. [9]), Y_{wave} (weighted importance of functional attribute; Eq. [10]), X_{SD} (SD of the dysfunctional attribute; Eq. [11]), and Y_{SD} (SD of the functional attribute; Eq. [12]). Eq. (6) is defined, as follows:

$$X_{ave} = \frac{\sum_i X_{ij}}{N}. \tag{6}$$

Eq. (7) is defined, as follows:

$$Y_{ave} = \frac{\sum_i Y_{ij}}{N}. \tag{7}$$

Eq. (8) is defined, as follows:

$$W_{ave} = \frac{\sum_i W_{ij}}{N}. \tag{8}$$

Eq. (9) is defined, as follows:

$$X_{wave} = \frac{\sum_i W_{ij} * X_{ij}}{\sum_i W_{ij}}. \tag{9}$$

Eq. (10) is defined, as follows:

$$Y_{wave} = \frac{\sum_i W_{ij} * Y_{ij}}{\sum_i W_{ij}}. \tag{10}$$

Eq. (11) is defined, as follows:

$$X_{SD} = \frac{\sum_i |X_j - X_{ave}|^2}{N}. \tag{11}$$

Eq. (12) is defined, as follows:

$$Y_{SD} = \frac{\sum_i |Y_j - Y_{ave}|^2}{N}. \tag{12}$$

From tabular summary a plot is produced, where uncertainty of categorizations is included in visualization.

Ethical Procedures

Prior to commencement, participants were briefed on the study's objectives and requirements. Voluntary participation was emphasized, with clear communication regarding the expected time commitment and the option to withdraw at any point. Responses were collected, anonymized, and kept confidential within the research group, ensuring participant privacy and data protection.

RESULTS

The results in this section present five tabular and graphical analysis techniques by applying Kano's two-factor method for evaluating users' (i.e., students') perceived usefulness of LAD features of an ALP. This demonstration also includes prioritizing the features with decision-making dilemmas rather than presenting the features' scores. One student's response was excluded by the researchers because all features were categorized by Kano method to be "questionable" making the students' response unusable. To assess the internal consistency and reliability of the Kano questionnaire Cronbach's alpha was calculated to be 0.82 leaving us with a very good reliability of the instrument.

Categorization of Quality Attributes Using Frequency & Evaluating Significance Using Fong's Test

The first technique is to categorize each of the features by applying Table 1 and calculate the significance of the categorization, as shown in **Table 1**. Here, we show how the students' responses are grouped into the A, M, O, R, Q, and I categories. **Table 1** shows LAD features' category and Fong test results to evaluate whether features attain significance for this categorization. This is only attained by No. 11 and No. 12. Solely using the most frequent response for classification does not give the possibility of evaluating how distinct a category is compared to LADs or to what extent respondents perceive the importance of the features of LAD.

Categorization of Quality Attributes Using Category Strength Scores & Stated Importance Scores

The second technique enables the designer to apply a different approach to assess the importance and category strength. In **Table 1**, classification is derived from 'most frequent response (a)'. Newly added columns in **Table 1** are category strength (CS), total strength (TS), 'better', 'worse', and average stated importance. These columns are described in the method section. In **Table 1**, evaluating whether a LAD feature fulfils the TS and CS criteria, a LAD feature needs CS>5% and TS>50% (see Eq. [2] and Eq. [3]). The criteria are fulfilled by feature No. 2, 4, and 10, which are classified as attractive (A), indifferent (I) and one-dimensional (O), respectively. The three features fulfil the CS and TS criteria and enable classifying into reliable categories. As opposed to the first technique, the features with reliable classification are non-overlapping due the differences in the numerical methods.

Table 1. Significance of categorization

No	Dashboard name	A	M	O	R	Q	I	MFR (a)	SMFR (b)	|a-b|	1.65*sqrt[(a+b*[2n-a-b]/2n)]	Test
1	History element	6	0	2	1	1	6	A (6)	I (6)	0<	4.19	No
2	Table of contents	7	1	4	0	0	4	A (7)	O (4)	3<	4.16	No
3	Progression overview	2	1	4	1	1	7	I (7)	O (4)	3<	4.16	No
4	Assessment of own level	2	3	5	0	0	6	I (6)	O (5)	1<	4.16	No
5	My progress	5	0	2	2	1	6	I (6)	A (5)	1<	4.16	No
6	How sure am I	5	0	3	1	1	6	I (6)	A (5)	1<	4.16	No
7	Read aloud	1	1	4	3	0	7	I (7)	O (4)	3<	4.16	No
8	Learn more	3	0	4	2	1	6	I (6)	O (4)	2<	4.09	No
9	Break recommendations	3	1	2	3	1	6	I (6)	A (3)	3<	4.00	No
10	Text formatting	2	0	7	1	0	6	O (7)	I (6)	1<	4.21	No
11	Metacognition	2	0	3	2	0	9	I (9)	O (3)	6<	4.19	Yes
12	Library	3	0	3	1	0	9	I (9)	O (3)	6<	4.19	Yes
13	Progression	2	0	5	1	0	8	I (8)	O (5)	3<	4.21	No

Note. MFR: Most frequent response; SMFR: Second most frequent response; & T: Significant Kano categorization according to Fong test; A: Attractive; M: Must-be; O: One-dimensional; R: Reverse; Q: Questionable; I: Indifferent; & to determine category when scores are equally distributed, prioritization falls in order must-be>one-dimensional>attractive>indifferent (Berger et al., 1993)

Table 2. Average stated importance

No	Classification	Classification agreement	CS	TS	Better	Worse	Average stated importance
1	A	A (35.29) & I (35.29)	0.00%	47.06%	0.57	- 0.14	4.65
2	A	41.18	**17.65%**	**70.59%**	**0.69**	**- 0.31**	**5.35**
3	I	41.18	17.65%	41.18%	0.43	- 0.36	4.24
4	I	35.29	**5.88%**	**58.82%**	**0.44**	**- 0.50**	**4.65**
5	I	35.29	5.88%	1.18%	0.54	- 0.15	4.76
6	I	35.29	5.88%	47.06%	0.57	- 0.21	4.41
7	I	41.18	17.65%	35.29%	0.38	- 0.38	4.24
8	I	35.29	11.76%	41.18%	0.54	- 0.31	4.65
9	I	35.29	17.65%	35.29%	0.42	- 0.25	4.12
10	O	35.29	5.88%	52.94%	0.60	- 0.47	5.47
11	I	52.94	52,94%	29.41%	0.36	- 0.21	4.06
12	I	52.94	35,29%	35.29%	0.40	- 0.20	4.24
13	I	47.06	17.65%	41.18%	0.47	- 0.33	4.24

Prioritization of Quality Attributes Using "Better" & "Worse" Scores

The third technique includes decision-making based on 'better' & 'worse' scores, as presented in **Table 2**.

The subsequent application of Eq. (4) and Eq. (5) for calculating 'better' and 'worse', along with the 'average stated importance', are useful and reliable only for the three features. To assess 'better' and 'worse' scores, you rank it either according to the highest 'better' score or according to the lowest 'worse' score. Evaluating this ranking of each LAD feature will be a combination of the highest 'better', lowest 'worse', and highest importance scores. The prioritization here depends on the evaluator's opinion based on the better and worse scores (Berger et al., 1993). For the three features identified, No. 2, 4, and 10, the prioritization from the tabular presentation would be to first focus on 2, then 10, and lastly, 4.

The 'better' scores indicate that satisfaction will be increased by providing the (attractive and one-dimensional) elements and 'worse' numbers indicate that satisfaction will be decreased if these (one-dimensional and must-be) elements are not included. Therefore, only feature 2 (A) and 10 (O) are the distinctly identifiable, including the highest 'average stated importance', while feature 4 with classification indifferent, may cause dissatisfaction (worse score -0.50) if it is not included.

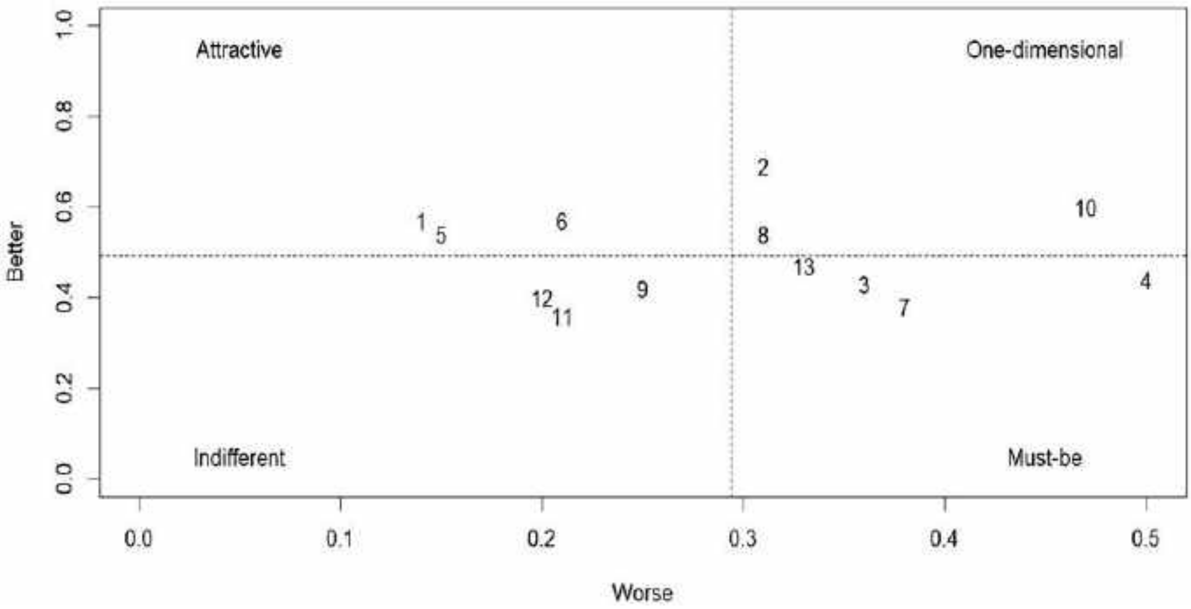

Figure 4. Kano better-worse plot (Khalid et al., 2008)

Categorization & Prioritization of Quality Attributes by "Better" (Y-Axis) & "Worse" (X-Axis) Score Plotting

Even though not all features have sufficient CS and TS, **Figure 4** demonstrates the use of plotting the 'better' and 'worse' scores as X and Y coordinates for mapping, categorizing, and prioritizing the features into attractive, one-dimensional, indifferent, and must-be.

Deciding from the plot, the order of prioritization is must-be>one-dimensional>attractive>indifferent (Berger et al., 1993). So, LADs are classified into one of four quadrants, where they can be placed in order of importance going from quadrant IV (LADs 3, 4, 7, and 13), I (LADs 2, 8, and 10), II (LADs 1, 5, and 6), and III (LADs 9, 11, and 12) (Berger, 1993). Plotting with 'better' and 'worse' gives a more nuanced classification of LADs, which can be further nuanced by adding the stated importance of each LAD. Either visually with color-coding of LADs in the plot or by adding the stated importance score in the plot. Plotting with 'better' and 'worse' as a method for visualizing results gives a visualization and a categorization of results, but not a visual way of determining whether categorizations are correct, the uncertainty of the results, or a way to evaluate how important each feature is for the students. To meet these requirements, we now apply the fifth technique.

Categorization & Prioritization by Calculating & Plotting Functional & Dysfunctional Weighted Average Scores

In the fifth technique, for the same calculation of scores in **Table 3** based on the responses, there are two different plotting techniques for decision-making with nuanced interpretation. For calculating the dysfunctional weighted average (X_{wave}), functional weighted average (Y_{wave}), and their corresponding standard deviations (Y_{SD} & X_{SD}), the total score of each of the quality attributes or LADs and the total perceived importance scores (W) are applied with the formula shown in the methods section.

Taking the values of the X_{wave} (dysfunctional weighted average) and Y_{wave} (functional weighted average), the Kano X_{wave}-Y_{wave} plot (**Figure 5**) enabled categorizing LADs into four categories, which are attractive, must-be, one-dimensional, and indifferent. Although only LADs 2, 4, and 10 are acceptable, all LADs are shows on the plot.

Berger et al. (1993) explain that the limits of the plot should be from 0 to 4 on both axes since questionable (Q) classifications are sorted out of the data and classifications should, therefore, not have a significant impact on the data. Berger et al. (1993) also explains that adding error bars to scores visualizes uncertainty related to each score. The uncertainty is visualized by adding $\pm X_{SD}$ as a horizontal line from the coordinates and $\pm Y_{SD}$ as a vertical line.

Table 3. Calculation of scores

No	Y=Functional	X=Dysfunctional	W=Importance	Y_{ave}	X_{ave}	W_{ave}	Y_{wave}	X_{wave}	Y_{SD}	X_{SD}
1	32	6	79	1.88	0.35	4.65	2.51	0.13	2.39	1.77
2	44	-9	91	2.59	-0.53	5.35	3.32	-0.65	2.09	1.33
3	22	-8	72	1.29	-0.47	4.24	2.06	-0.72	2.34	1.66
4	27	-14	79	1.59	-0.82	4.65	2.15	-0.78	2.29	1.42
5	19	4	81	1.12	0.24	4.76	1.83	0.74	2.60	1.71
6	40	8	75	2.35	0.47	4.41	3.04	0.00	2.03	1.94
7	24	0	72	1.41	0.00	4.24	2.36	-0.50	2.32	2.00
8	26	8	79	1.53	0.47	4.65	2.43	0.13	2.50	2.18
9	20	15	70	1.18	0.88	4.12	2.21	0.54	2.42	2.29
10	34	-4	93	2.00	-0.24	5.47	2.88	-0.60	2.35	1.99
11	13	4	69	0.76	0.24	4.06	1.51	0.06	2.28	1.86
12	20	-2	72	1.18	-0.12	4.24	1.89	-0.06	2.24	1.32
13	24	-6	72	1.41	-0.35	4.24	1.94	-0.31	2.40	1.46

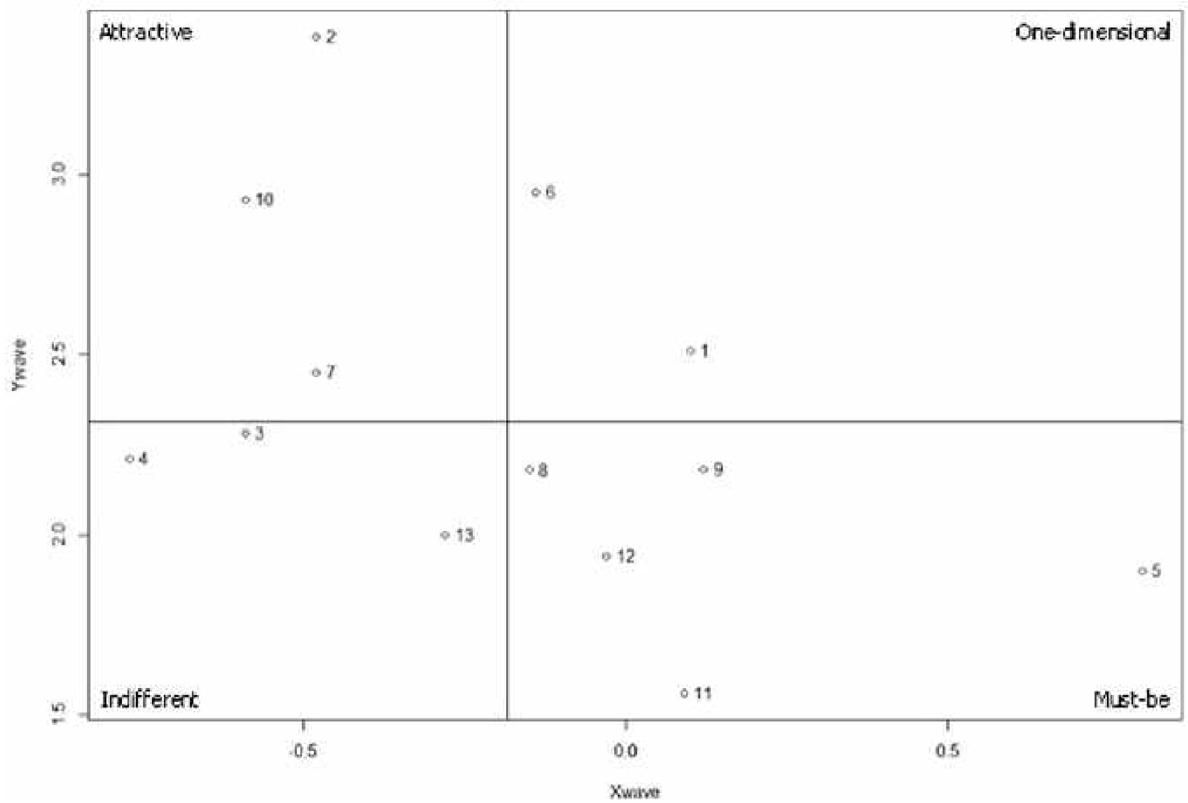

Figure 5. Kano Xwave-Ywave plot for categorization of LADs (inspired by Berger et al., 1993, p. 20-23)

Figure 6 shows the plotting of the perceived importance scores for LADs 2, 4, and 10 including error bars to the plot, which give a notion of how off-target the categorizations are. While **Figure 5** plot sets the coordinates based on the scores, **Figure 6** shows the 0-4 normalized plotting of the three LADs that fulfil the CS and TS criteria. The three LADs are attractive for the students.

Overview of Results Based on Five Techniques

An overview of the categorization and prioritization of LADs based on the different numerical, tabular and graphical methods are shown in **Table 4**. Only technique 5 takes the stated importance into consideration for each of the quality attributes or LADs. The first technique results the list of features with significant Kano categorization according to Fong's test, resulting only LADs 11 (metacognition) and 12 (library) to have reliable categorization, which is one-dimensional (indifferent). Even with higher number of respondents, the categorization may have similarly low number of features with significant Kano categorization.

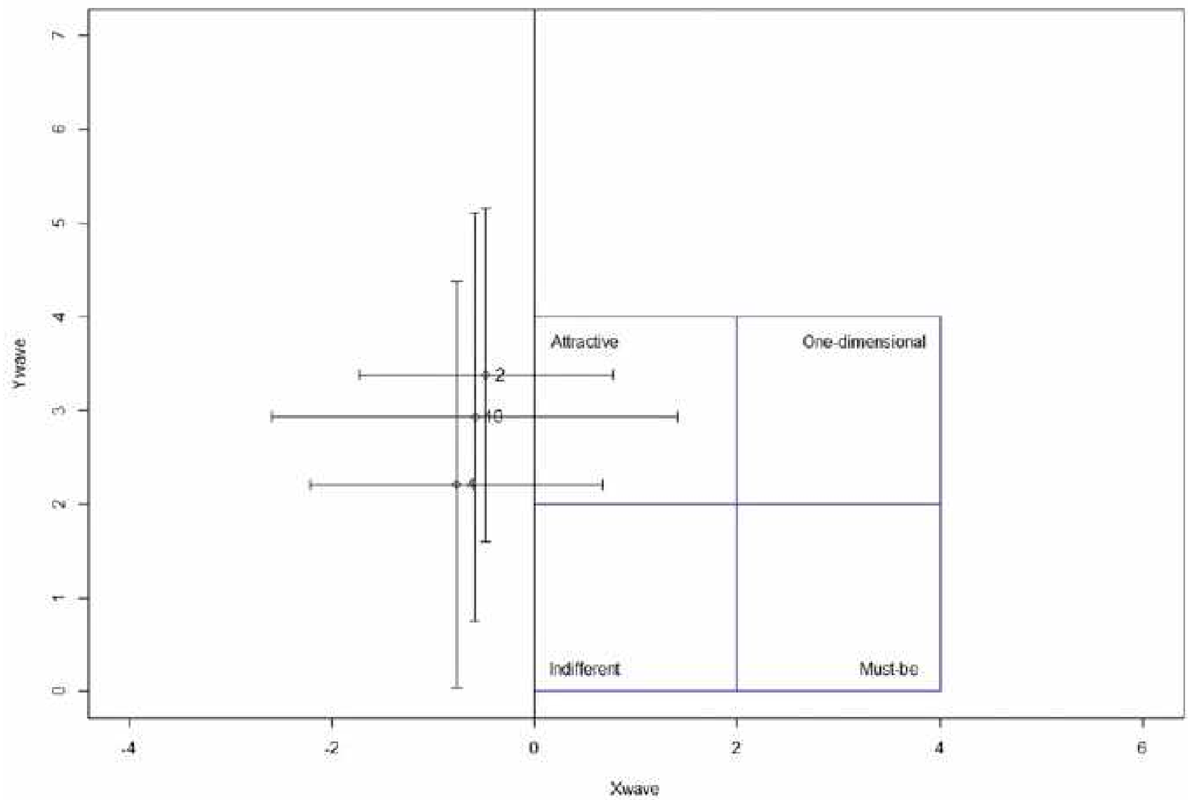

Figure 6. Kano Xwave-Ywave plot of 2, 4, & 10 (inspired by Berger et al., 1993)

According to the protocol, for the subsequent techniques, the rest of LADs or quality attributes should be excluded. However, for the demonstration of the subsequent techniques 2-5, the exclusion process was ignored. There is a development throughout the techniques towards higher nuance and complexity, going from the first to the fifth techniques and, thereby, a need for more responses. In choosing the right technique, designers must also keep in mind what their needed level of complexity is regarding the amount of data they can attain. Therefore, based on learners' perception towards to availability and unavailability of each of LADs and the stated importance (on a scale of 1 to 10), the prioritization of LADs can be calculated with some build-in validation techniques. In the case of Rhapsode ALP, with the different Kano theory techniques, it is possible to prioritize LAD features, giving feedback to designers on how to develop an ALP.

DISCUSSION & CONCLUSIONS

Theoretically, this paper demonstrates the application of Kano method as five inter-related techniques, as a way of prioritizing features of an ALP's LAD for designers. Thus, results from the five techniques can supply designers, e-learning specialists, software developers, and educational managers with more actionable feedback than usability and user experience methods. Subsequently, Kano theory techniques can also be preferable to self-developed questionnaires, which require pilot-testing and internal consistency validation of responses. In practice, management roles in educational institutions, e-learning professionals, software developers, and researchers can apply the five-step Kano method for systematically exclude features/functionalities based on reliable users' responses. This paper demonstrates the application of the method in the case of perceived usefulness of LADs of an ALP, but the protocol can be applied in the case of any product or service. Multiple reliability techniques are demonstrated for informed exclusion of features by applying Cronbach's alpha, Fong test, category strength and total strength, numerical and visual analysis of the scores. In the case of the nursing students using the Rhapsode™ ALP, only two of the 13 LADs could be reliably categorized as indifferent. Ignoring the Fong's test-based significance scores for categorization, only three of 13 LADs could be shown with normalized plotting and depicting of uncertainty of categorization, which are (#2) table of content, (#4) assessment of own level, (#10) text formatting–as attractive features.

Table 4. An overview of categorization & prioritization of LADs

Techniques	Results of each technique
1. Categorization of quality attributes using frequency & significance using Fong's test	Significant categorization through Fong test of LAD 11 (metacognition) & 12 (library)
2. Categorization of quality attributes using category strength scores & stated importance scores	Distinct categorization through category strength (CS) & total strength (TS) of LAD 2 (table of content), 4 (assessment of own level), & 10 (text formatting)
3. Prioritization of quality attributes using "better" & "worse" scores	Priority of previous LADs in following order 2 as attractive (A), then 10 as one-dimensional (O), & lastly, 4 as indifferent (I)
4. Categorization of quality attributes by "better" (y-axis) & "worse" (x-axis) score plotting	Categorization & prioritization of LADs in categorise of 4 must-be (M), 10 (O), 2 (O), without normalized plotting
5. Categorization by calculating & plotting functional & dysfunctional weighted average scores	Categorization and prioritization of LADs in categorise of 4 (A), 10 (A), 2 (A), with normalized plotting & depicting of uncertainty of categorization

Ensuring an adequate number of responses is a recognized concern in Human-Computer Interaction research. The recommended number of participants, ranging from five to sixty, depends on factors such as research method, design phase, and user information sought (Baxter et al., 2015). Failure to meet user participation requirements introduces uncertainty, but the Kano two-factor method techniques presented in this study address uncertainty by revealing uncertainties related to the number of user responses, enhancing their applicability. Witell et al. (2013) reviewed studies (n=147) on the application of Kano method, where they found that: "*almost all empirical investigations [...] simply count the number of responses in a certain quality dimension*" (Witell et al., 2013, p. 14) and thereby not using methods of testing the reliability of the categorization. This study applies Fong's test for reliable categorization and shows that only two of 13 LADs could have reliable categorization using small samples size (n=17).

This enables a nuanced assessment of users' sentiments towards the design, and in the case of LAD prioritization design, Kano method techniques provide multiple ways of prioritizing the design of LAD. Furthermore, from the 'average stated importance' we can attain that students overall did not see any of LADs as being very important as all LADs average score range was between 4.12-5.47 on a Likert scale of 9. Recommendation for future work consists of expanding the application of Kano method with perceived usefulness, and perceived ease of use from technology acceptance model in the domain of ALPs and LADs to assess both ease of use as well as the usefulness of the technologies alongside the prioritization on perceived usefulness of Kano method. This combination might further develop the understanding of theory behind Kano method, as requested by (Witell et al., 2013).

Empirical experimental research can be conducted by increasing sample size and by built-in simulation in R/Python or other statistical simulation tools for analyzing the change in the pattern. Moreover, change in the categorization of LADs analyzed by increasing the number of dashboards, and test the differences by grouping LADs in sub-categories. Future research in higher education context should consider two central limitations experienced in this research and devised strategies as an attempt to circumvent. First, this large-scale funded project was flagged as increasingly interacting with the students resulting in increased time spent that are not immediately value-adding for their learning and the project team were allowed very limited interaction. Secondly, we hypothesize that the Nordic region's higher individualism and autonomy, as suggested by Hofstede (2024), might have resulted in a low response rate. Simulation can be considered as a potential approach to address the issue.

Author contributions: TABT-F: collected data, analyzed data, & summarized findings & **SK:** analyzed data & supported summarizing findings. All authors approved the final version of the article.

Funding: This article was funded by Innovationsfonden (DK), grant number (file number) 9093-00012B–Project website information: https://innovationsfonden.dk/da/i/historier/ny-og-innovativ-uddannelsesmodel-skal

Acknowledgements: The authors would like to thank students who participated in the study in question.

Ethics declaration: The authors declared that the study did not require an ethics committee's approval. The authors further declared that data consists of 17 student responses to a survey. Students were informed of the study, that their data was to be used in a scientific publication, and that their consent could be withdrawn at any moment. Written informed consents were obtained from the participants. The privacy of the participants was ensured and no information to reveal the identity of students is provided in the responses. Data was stored on university servers and will only be stored for as long as the NurseEd project is still ongoing (ends in 2024) and personel and private data is neither collected or stored on the participants.

Declaration of interest: The authors declare no competing interest.

Data availability: Data generated or analyzed during this study are available from the authors on request.

REFERENCES

Ahn, J., Campos, F., Hays, M., & Digiacomo, D. (2019). Designing in context: Reaching beyond usability in learning analytics dashboard design. *Journal of Learning Analytics, 6*(2), 70-85. https://doi.org/10.18608/jla.2019.62.5

Area9 (2024, January 1). *"RhapsodeTM Learner": Area9 RhapsodeTM*. https://area9lyceum.com/the-platform/

Bauk, S., Šćepanović, S., & Kopp, M. (2014). Estimating students' satisfaction with web based learning system in blended learning environment. *Education Research International, 2014*, 731720. https://doi.org/10.1155/2014/731720

Baxter, K., Courage, C., & Caine, K. (2015). *Choosing a user experience research activity in understanding your users: A practical guide to user research methods*. Morgan Kaufmann. https://doi.org/10.1016/B978-0-12-800232-2.00005-5

Benyon, D. (2019). *Designing user experience*. Pearson.

Berger, C., Blauth, R., Boger, D., Bolster, C., Burchill, G., DuMouchel, W., Pouliot, F., Richter, R., Rubinoff, A., Shen, D., Timko, M., & Walden, D. (1993). A special issue on Kano's methods for understanding customer defined quality. *Center for Quality of Management Journal, 2*(4), 3-28.

Bodily, R., Ikahihifo, T. K., Mackley, B., & Graham, C. R. (2018). The design, development, and implementation of student-facing learning analytics dashboards. *Journal of Computing in Higher Education, 30*, 572-598. https://doi.org/10.1007/s12528-018-9186-0

Chen, L.-S., & Hsu, J.-Y. (2019). Discover users' needs in e-learning by Kano analysis and decision trees. In *Proceedings of the 2019 IEEE 6th International Conference on Industrial Engineering and Applications* (pp. 572-576). https://doi.org/10.1109/IEA.2019.8714873

Chen, W.-K., Chang, J.-R., Chen, L.-S., & Hsu, R.-Y. (2022). Using refined kano model and decision trees to discover learners' needs for teaching videos. *Multimedia Tools and Applications, 81*(6), 8317-8347. https://doi.org/10.1007/s11042-021-11744-9

de Kock, E., van Biljon, J., & Pretorius, M. (2009). Usability evaluation methods: Mind the gaps. In *Proceedings of the 2009 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists* (pp. 122-131). https://doi.org/10.1145/1632149.1632166

Earnshaw, Y., Tawfik, A. A., & Schmidt, M. (2018). User experience design. In R. E. West (Ed.), *Foundations of learning and instructional design technology*. EdTech Books.

Few, S. (2006). *Information dashboard design, the effective visual communication of data*. O'Reilly.

Fong, D. (1996). Using the self-stated importance questionnaire to interpret Kano questionnaire results. *Center for Quality Management Journal, 5*, 21-24.

Fontaine, G., Cossette, S., Maheu-Cadotte, M.-A., Mailhot, T., Deschênes, M.-F., Mathieu-Dupuis, G., Côté, J., Gagnon, M.-P., & Dubé, V. (2019). Efficacy of adaptive e-learning for health professionals and students: A systematic review and meta-analysis. *BMJ Open, 9*(8), e025252. https://doi.org/10.1136/bmjopen-2018-025252

Hofstede, G. (2024). Country comparison tool. *The Culture Factor Group*. https://www.hofstede-insights.com/country-comparison-tool

Hornbæk, K. (2006). Current practice in measuring usability: Challenges to usability studies and research. *International Journal of Human-Computer Studies, 64*(2), 79-102. https://doi.org/10.1016/j.ijhcs.2005.06.002

Iordache, D. D., & Pribeanu, C. (2009). A comparison of quantitative and qualitative data from a formative usability evaluation of an augmented reality learning scenario. *Informatica Enconomica* [*Economic Informatics*], *13*(3), 67-74.

Kano, N., Seraku, N., Takahashi, F., & Tsuji, S. (1984). Attractive quality and must-be quality. *Journal of the Japanese Society for Quality Control, 41*, 39-48.

Khalid, M. S., Mustafa, A., & Haque, I. (2008). Application of Kano's model for evaluating information quality of university websites. In H. R. Arabnia & A. Marsh (Eds.), *Proceedings of the 2008 International Conference on Semantic Web & Web Services* (pp. 277-280). CSREA Press.

Kim, J., Jo, I. H., & Park, Y. (2016). Effects of learning analytics dashboard: Analyzing the relations among dashboard utilization, satisfaction, and learning achievement. *Asia Pacific Education Review, 17*, 13-24. https://doi.org/10.1007/s12564-015-9403-8

Kirkpatrick, D. L., & Kirkpatrick, J. D. (2006). *Evaluating training programs: The four levels*. Berrett-Koehler.

Lesemann, E., Woletz, N., & Koerber, S. (2007). Combining methods to evaluate mobile usability. In *Proceedings of the 9th International Conference on Human Computer Interaction with Mobile Devices and Services* (pp. 444-447). ACM. https://doi.org/10.1145/1377999.1378051

Lewis, J. R. (2018). The system usability scale: Past, present, and future. *International Journal of Human-Computer Interaction, 34*(7), 577-590. https://doi.org/10.1080/10447318.2018.1455307

Löfgren, M. (2005). Kano's theory of attractive quality and packaging. *Quality Management Journal, 2*(3), 7-20. https://doi.org/10.1080/10686967.2005.11919257

Löfgren, M., & Witell, L. (2008). Two decades of using Kano's theory of attractive quality: A literature review. *Quality Management Journal, 15*(1), 59-75. https://doi.org/10.1080/10686967.2008.11918056

Madzík, P., Budaj, P., Mikuláš, D., & Zimon D. (2009). Application of the Kano model for a better understanding of customer requirements in higher education–A pilot study. *Administrative Sciences, 9*(1), 11. https://doi.org/10.3390/admsci9010011

Matzler, K., Hinterhuber, H. H., Bailom, F., & Sauerwein, E. (1996). How to delight your customer. *Journal of Product & Brand Management, 5*(2), 6-18. https://doi.org/10.1108/10610429610119469

Nakamura, W. T., de Oliveira, E. H. T., & Conte, T. (2017). Usability and user experience evaluation of learning management systems. In *Proceedings of the 19th International Conference on Enterprise Information Systems* (pp. 97-108). SciTePress. https://doi.org/10.5220/0006363100970108

Park, Y., & Jo, I. (2015). Development of the learning analytics dashboard to support students' learning performance. *Journal of Universal Computer Science, 21*(1), 110-133.

Roberts, L. D., Howell, J. A., & Seaman, K. (2017). Give me a customizable dashboard: Personalized learning analytics dashboards in higher education. *Technology, Knowledge and Learning, 22*, 317-333. https://doi.org/10.1007/s10758-017-9316-1

Salas, J., Chang, A., Montalvo, L., Nunez, A., Vilcapoma, M., Moquillaza, A., Murillo, B., & Paz, F. (2019). Guidelines to evaluate the usability and user experience of learning support platforms: A systematic review. In P. Ruiz, & V. Agredo-Delgado (Eds.), *Human-computer interaction* (pp. 238-254). Springer. https://doi.org/10.1007/978-3-030-37386-3_18

Sauro, J., & Lewis J. R. (2016). *Standardized usability questionnaires in quantifying the user experience: Practical statistics for user research*. Morgan Kaufmann. https://doi.org/10.1016/B978-0-12-802308-2.00008-4

Schwendimann, B. A., Rodríguez-Triana, M. J., Vozniuk, A., Prieto, L. P., Boroujeni, M. S., Holzer, A., Gillet, D., & Dillenbourg, P. (2017). Perceiving learning at a glance: A systematic literature review of learning dashboard research. *IEEE Transactions on Learning Technologies, 10*(1), 30-41. https://doi.org/10.1109/TLT.2016.2599522

Sedrakyan, G., Mannens, E., & Verbert, K. (2019). Guiding the choice of learning dashboard visualizations: Linking dashboard design and data visualization concepts. *Journal of Computer Languages, 50*, 19-38. https://doi.org/10.1016/j.jvlc.2018.11.002

Tretow-Fish, T. A. B., & Khalid, M. S. (2023). Methods for evaluating Learning Analytics and Learning Analytics Dashboards in Adaptive Learning Platforms: A systematic review. *Electronic Journal of E-Learning, 21*(5), 430-449. https://doi.org/10.34190/ejel.21.5.3088

Verbert, K., Ochoa, X., De Croon, R., Dourado, R. A., & De Laet, T. (2020). Learning analytics dashboards: The past, the present and the future. In *Proceedings of the 10th International Conference on Learning Analytics & Knowledge* (pp. 35-40). ACM. https://doi.org/10.1145/3375462.3375504

Witell, L., Löfgren, M., & Dahlgaard, J. J. (2013). Theory of attractive quality and the Kano methodology–The past, the present, and the future. *Total Quality Management and Business Excellence, 24*(11-12), 1241-1252. https://doi.org/10.1080/14783363.2013.791117

Xie, H., Chu, H.-C., Hwang, G.-J., & Wang, C.-C. (2019). Trends and development in technology-enhanced adaptive/personalized learning: A systematic review of journal publications from 2007 to 2017. *Computers & Education, 140*, 103599. https://doi.org/10.1016/j.compedu.2019.103599

Yoo, M., & Jin, S. (2020). Development and evaluation of learning analytics dashboards to support online discussion activities. *Educational Technology & Society, 23*(2), 1-18.