A Holistic UX Methodological Framework for Measuring the Dynamic, Adaptive and Intelligent Aspects of a Software Solution

Vivien Johnston[1[0000-0002-9388-7637]], Michaela Black[2[0000-0002-5140-2566]], and Jonathan Wallace[1[0000-0002-8415-4001]]

[1] School of Computing, Ulster University, Jordanstown, Co. Antrim, Northern Ireland
{✉johnston-v7, jg.wallace}@ulster.ac.uk
[2] School of Computing, Engineering & Intelligent Systems, Ulster University, Londonderry, Co. Londonderry, Northern Ireland
mm.black@ulster.ac.uk

Abstract. The aim of this paper is to present research that proposes a methodological assessment framework based upon criteria/variables which have been grouped into three core aspects namely how Dynamic, Adaptive and Intelligent the user experience is. The framework aims to enhance the user experience of an application by analyzing these aspects and providing recommendations to a developer to produce a more enhanced Dynamic Adaptive Intelligent User Interface. Research into this field has identified that not all current applications are aware of, or capable of measuring all the aspects. The framework is based upon a three phased approach: phase one will measure the variables within each aspect and produce a score that indicates to a developer the degree to which their current application fits within each aspect; phase two highlights the areas of growth within each aspect and provide recommendations that would enhance the application's user experience; and phase three validates the framework by highlighting the application's user experience progress from previous measurements. All three phases are combined to produce a robustly proven tool that aids the developer with validated advice in order to enhance their products user experience.

Keywords: User Experience, Methodological Framework, Measure, System Scoring.

1 Introduction

Previous research includes the proposal of a framework that would be able to enhance the user experience (UX) of an application [1]. UX relates to the design, usability and functionality of an application's interface [2]. This work identified three aspects that are key to the creation of the framework, namely: Dynamic, Adaptive; and Intelligent (Figure 1) [1]. Each aspect contains parameters, a parameter is defined as being a measurable piece of information that is linked to one of the aspects of the framework [3].

The Dynamic aspect of the framework contains parameters relating to the contextual information of an end-user, their device; and their physical environment. The parameters for the Dynamic aspect include: the type of device the end-user is using; and the time of day they are accessing the application [1]. The Adaptive aspect measures existing parameters about each end-user, such as: their knowledge set, capabilities; and their goal for using the application [1]. The Intelligent aspect uses data analysis to help identify patterns and trends within data. This aspect will help enhance the UX for each cohort that an end-user would belong to by working in conjunction with the previous two aspects [1].

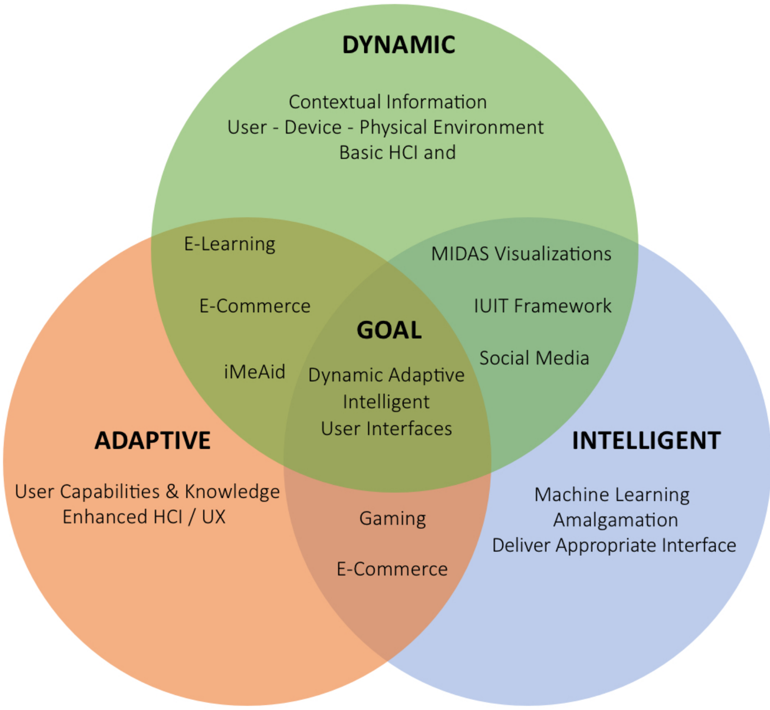

Fig. 1. The Intersections of Dynamic, Adaptive and Intelligent Aspects with Examples [1].

This paper extends research that was carried out in the work of [1]. It has been identified through previous research that not all forms of measurement have been capable of measuring all three recognized aspects however, the Hawthorne Effect has been identified during UX analysis.

The Hawthorne Effect has been identified during many forms of observational analysis by each end-user when providing feedback on the UX of an application's interface. It is the feeling of pressure whilst being observed during a task, this then leads to unusual interactions by each end-user [4]. This is a factor that the framework in question will take into consideration, it will bypass any end-user feeling pressurized. To avoid an end-user feeling pressurized when supplying feedback, it would be helpful if there was a framework that could: measure the degree to which an application fits within each aspect and illustrate this with a score, highlight the areas of growth and provide recommendations that will enhance the UX of an application; and provide validation to highlight the UX progress from previous measurements. The framework would benefit developers within multiple domains and assist with their software development process, and overall enhance the UX of their products/service solutions.

In order for this framework to be created, identification of additional parameters within each of the three aspects mentioned above is needed. A form of measurement will be highlighted as to the degree of each parameter. Once this framework is fully integrated into the developer's software development process, it will assist in producing a Dynamic Adaptive Intelligent User Interface (DAIUI).

The remainder of this paper is structured as the following: section 2 is related work, section 3 is the methodology; and section 4 contains the conclusion and future work.

2 Related Work

UX design can be described as being a narrative, by providing a story to the end-user via a network of events. A narrative could be portrayed as one of two approaches: task or experience. Task is in relation to a goal that an end-user may have or is carrying out when using an application. Experience, however, is in relation to the types of emotions and meaning behind the interactions from each end-user [5]. The main narrative that is important is the UX of an application, and feedback regarding this is normally provided by end-users' through a variety of evaluation methods.

Evaluation methods are categorized into three segments: self-reported, observational; and physiological measurements [6]. Self-reported measurements relate to an end-user documenting their thoughts and feelings via a survey or questionnaire, observational refers to observing an end-user whilst they interact with an application; and physiological relates to sensors attached to an end-user that monitor their physical movement in the form of quantifiable data [6]. Measurements including surveys can be time consuming and there could also be the issue of subjective bias, such as: are there a mix of quantitative and qualitative questions; and whether it is completed alone or alongside an observer face-to-face. This approach could sway the end result in favor of the observer or, the end-user could tell the observer what they want to hear, as opposed to what they really think themselves. The structure of how each question is presented could ultimately impact an end-user's decision. Observational methods consist of Cognitive Walkthroughs and Think-Aloud sessions, these methods help understand the thoughts and decisions an end-user makes whilst navigating an application [7]. All of these issues link back to the Hawthorne Effect, as to whether an observer is influencing their decision [4]. This is where the framework would be of benefit to the developer.

The UX industry has been using evaluation methods throughout their design process however, evaluation metrics are a developing area. Evaluation metrics measure the UX of an application to calculate a score. UserZoom created a single UX metric called qxScore [8]. qxScore benchmarks the experience of an application by evaluating two areas: behavior, relating to task success rate; and attitude, including trust, ease of use and appearance [8]. A qxScorecard generates the results by indicating the quality of experience from 0 to 100, >45 being very poor and 91-100 being great. This UX metric covers the fundamentals of UX evaluation however, it is then up to the developers and stakeholders within the company to decide on what improvements to make that will enhance the UX of their application, and there lies a gap. Alternative methods of evaluation have been used within other domains, such as education.

Within [9], they used a UX/UI evaluation framework within the cyberlearning environment to evaluate two main areas: usability; and utility. Other areas of interest included: technology, users; and context. Usability attributes consisted of problem-based learning and ease of use evaluations via Cognitive Walkthroughs and Heuristic surveys [9]. Utility attributes used pre/post-test and final scores for learning achievements, and a UX/UI survey to document the evaluation of user satisfaction [9]. These are appropriate evaluation methods and are possibly more manual and traditional compared to qxScore. Although it may be a time-consuming process, there are other factors that could assist and make the evaluation process more engaging, such as gamification.

Gamification is the incorporation of gaming elements, such as: point scoring, leader boards; and levels [10], all of which were incorporated into [9]. This kept the students engaged for longer, return often to complete assignments; and allow them to be in competition against their fellow classmates. In return, this led to honest feedback. This is due to the lack of pressure as nobody was watching, and gamification assisted in providing a sense of enjoyment whilst also keeping them engaged. In addition, gamification also allows for gaps to appear that highlight topic areas that a student might be struggling with, this has been demonstrated within M-Elo.

M-Elo incorporated gamification elements to identify the knowledge gap of each student, whilst also considering their parameters [11]. These parameters were independent and helped model each end-user's knowledge state which assisted in the recommendation process. A visualization widget allowed each end-user to track their current knowledge against their peers. In return, the application provided questions based upon their largest knowledge gap [11]. A Likert-scale survey was then used to capture end-user feedback, covering areas such as: motivation, rationality; and trust [11]. Students detailed that the incorporation of peer comparison provided a sense of trust that encouraged motivation in order to progress their education to the next level, based upon the appropriate recommendations provided [11]. This in return can reduce their cognitive load, this is the amount of cognitive effort required to understand the topic, presentation and sequence of events whilst using an application [12]. All of these factors detailed will be considered for the framework in question to assist the developer within their software development process. The framework will objectively measure and indicate where an application fits within the three aspects (this is currently not provided within other research), highlight the areas of growth and make appropriate recommendations, without an observer influencing its decisions. The framework will not only improve the user journey for each end-user, but it will provide a clear direction for the developer.

3 Methodology

The framework is based upon a three phased approach and therefore three studies will fulfil each phase, phase one is currently underway.

Phase one of this framework will input data from specific domains (which is still ongoing research) in relation to the UX of the application, for example education. A hierarchy flow weight measurement will assist in producing a score. This is the main score that the developer receives about how Dynamic, Adaptive and Intelligent the UX of their application is. In order for a score to be established, the following measurement process must take place.

The measurement takes into account that the UX of an application as a whole is marked out of 100%. This total percentage is divided between each aspect of the framework: 33.33% Dynamic, 33.33% Adaptive; and 33.33% Intelligent. Each aspect holds parameters that are specific to each aspect, these have been identified and assigned during initial research. Each parameter within its assigned aspect is assigned a weight out of 100% and is based upon how much influence and value it would contribute to the UX of an application - the heavier the weighting, the greater influence on the scoring. In addition, each parameter contains sub-parameters. These sub-parameters detail what is required within each parameter in order to achieve the full weighting listed. Each sub-parameter has its own weighting in accordance to the influence and value that is required in order to enhance the UX. Table 1 below illustrates the parameters and sub-parameters in each aspect and their weights which have been weighed out of 100%.

Table 1. Sub-Parameters and Weights within Assigned Parameters.

	Aspect	Parameter	Weighting	Sub-Parameter	Weighting
Overall Score 0-100	**Dynamic** 33.33%	Device Constraints	60%	Device Type	25%
				Screen Size	25%
				Battery Life	20%
				Processor Speeds	20%
				Memory Storage	5%
				Internet Speed	5%
		Contextual Information	20%	Home	30%
				Work	30%
				Walking	10%
				Time of Day	30%
		Location Constraints	20%	Network	80%
				Weather	20%
	Adaptive 33.33%	Capabilities	25%	Visual	40%
				Hearing	10%
				Cognitive	15%
				Physical	35%
		Knowledge Set	15%	Cohort	75%
				Technical Expertise	25%
		User Goal	60%	Searching	25%
				Browsing	25%
				Submitting Forms	25%
				Checking off a Task	25%
	Intelligent 33.33%	Knowledge Based Rules	100%	Personalized to User	100%

Figure 2 below is a hierarchy diagram providing a hypothetical example as to how the measurement and scoring would work. The circles that are highlighted in green (darker circles) will be used for demonstration purposes to showcase one set of parameters. Nodes with no colour detail the other weighting percentages that have been distributed.

As mentioned, an application as a whole is marked out of 100% (root node), and each of the three aspects (parent nodes) have been distributed a percentage of it: 33.33% Dynamic, 33.33% Adaptive; and 33.33% Intelligent. Every node from the root, to a leaf node within Figure 2 will be scored between 0 and 100, this score works in conjunction with the weighting that each parameter and sub-parameters have been allocated. For example, to understand the scoring detailed, we would start from the Network and Weather leaf nodes at the bottom of Figure 2 and work upwards.

The Network (sub-parameter) leaf node is in relation to the Location Constraints parameter. As illustrated within Figure 2, the Network leaf node has been allocated a score of 50/100, this makes converting to a percentage easier. For example, the Network leaf node has been given a weight of 80%, the score of 50 equates to 50% of the 80% weight, this means that Network has a score of 40%. The same principle is applied to the Weather leaf node, it has been given a score of 80 which equates to 80% of the 100 possible marks. 80% of the 20% weighting is equal to 16%. The percentages from each leaf node (40% and 16%) are added together to form the score for Location Constraints which is 56 out of the 100 possible marks. 56% of 20% Location Constraint allocation is equal to 11.2% and this then equates to the total score for the Dynamic parent node. As the Dynamic aspect is worth 33.33% of the 100 possible marks, the same calculation applied to the 11.2%, 11.2% of the 33.33% allocation to the Dynamic parent node is 3.73%. This means that the overall score within this example works out at 3.73% as only one aspect is being demonstrated (as an example due to the limited of space available), which is a very poor score. In order for a developer to understand if this score is poor or not, a form of gamification would be applied.

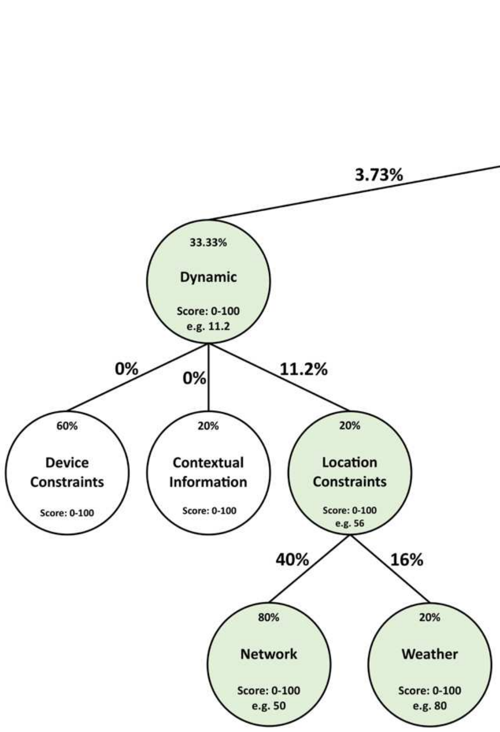

Fig. 2. An Example of Measurement and Scoring Allocation for the Dynamic Aspect via a Hierarchy Diagram.

In relation to gamification, most car racing games award bronze, silver or gold medals to those players who finish 3rd, 2nd or 1st. The same principle can be applied to the framework in question, these could be known as scoring boundaries. For example: scores between 0-39 would be bronze, 40-89 would be silver; and 90-100 would be gold. This would be the main score that a developer would see, as it is the aggregation of scores from the three core aspects.

Within education, students use a virtual learning environment known as Blackboard Learn (BBL). To produce a score for BBL, multiple forms of media have been taken into consideration: data, the application and screenshots supplied, the hierarchy diagram from Figure 2 was used to produce a final score. Figure 3 below illustrates the results: chart A indicates the percentage of each aspect that is currently being fulfilled, the highlighted border around the Adaptive aspect indicates what will be shown in chart B, chart B is drilled down from the Adaptive aspect in chart A detailing the parameters; and chart C is drilled down from chart B indicating the percentage of sub-parameters.

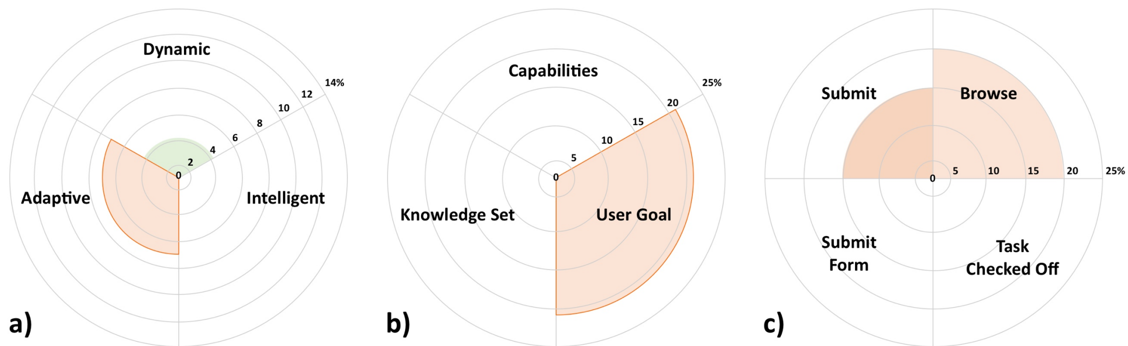

Fig. 3. BBL Results via Three Radar Visualizations.

The radar charts above show the scoring of selected parameters and sub-parameters. Based upon the scoring boundaries within this framework previously mentioned, BBL is 11.1%. As the numbers are between 0 and 39, it is categorized as being bronze. Phase two of this framework would then indicate the areas of growth and provide recommendations. The radar charts being used in Figure 3 are good to use for individual software solutions, while those provided in Figure 4 below illustrate what two software solutions would look like when compared and scored against the framework. Figure 4 is illustrating hypothetical results, and this would allow for a similarity score.

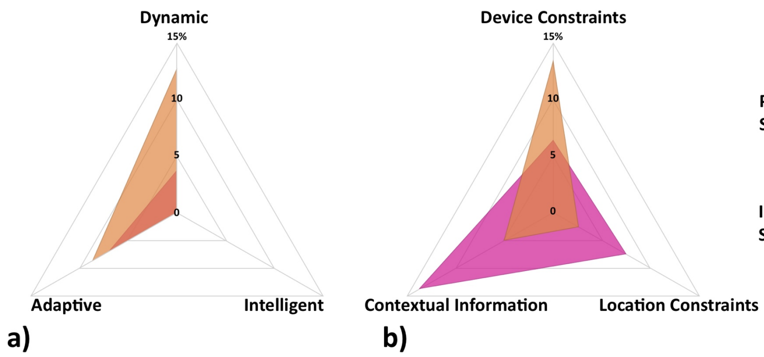

Fig. 4. Hypothetical Results Overlaying Two Software Solutions that are being Measured Against the Framework.

Phase two of the framework continues on from phase one and this will be the second study of the PhD. The main focus of phase two is to take the gaps (areas of growth) that are clearly visible from Figure 3 and provide recommendations to help the application improve. The recommendations would assist the developer by advising them on what their application needs in order to improve not only its UX, but its scoring that the framework has supplied.

As an example, based upon the results from BBL, it would be helpful if recommendations could be supplied in relation to the Intelligent aspect to boost its scoring. By boosting the Intelligent aspect, it would allow the application to provide material that is relevant to that particular student, whilst working with other parameters, such as type of device. Recommendations could be as simple as a tooltip, to draw the developer's attention to the parameter and sub-parameter from the charts illustrated above.

Phase three of the framework is the final phase. The main purpose of this phase is to validate the recommendations, and to accept that they do in fact enhance the UX of an application. In order to provide a form of validation, progress history would be an important factor for the developer. Progress history is in relation to the previous measurements and scores that the framework has produced from the same application. This allows the developer to see the progress their application is making in order to produce a more DAIUI. By viewing a progress history, it reassures the developer that the framework is having an impact on the UX of their application.

4 Conclusion and Future Work

The work presented here has detailed that not all applications are aware, or capable of measuring the three recognized aspects of: Dynamic, Adaptive; and Intelligent. In order for these aspects to be recognized, a holistic UX methodological based assessment framework has been outlined. It will assist and benefit developers within a variety of domains with their software development process. Overall, this will enhance the UX of their products/service solutions within their preferred domains.

Besides phases two and three of the framework, future work will consist of the translation of measurements that each sub-parameter contains. Translating them into a format that the framework will understand. Further work would entail how to automate the identification and measurement of each aspect in order to produce a score.

Acknowledgments

The authors wish to acknowledge the support of the European Commission on the H2020 project MIDAS (G.A. nr. 727721) and DfE NI (Department of the Economy Northern Ireland.

References

1. V. Johnston, M. Black, J. Wallace, M. Mulvenna, and R. Bond, "A Framework for the Development of a Dynamic Adaptive Intelligent User Interface to Enhance the User Experience," in *Proceedings of the 31st European Conference on Cognitive Ergonomics*, New York, NY, USA: Association for Computing Machinery, 2019, pp. 32–35.
2. "What is User Experience (UX) Design?," *The Interaction Design Foundation*. https://www.interaction-design.org/literature/topics/ux-design (accessed Dec. 20, 2018).
3. "What is parameter? - Definition from WhatIs.com," *WhatIs.com*. https://whatis.techtarget.com/definition/parameter (accessed Nov. 07, 2018).
4. R. Macefield, "Usability studies and the Hawthorne effect." Usability Professionals' Association, May 01, 2007.
5. A. P. O. S. Vermeeren, E. L.-C. Law, V. Roto, M. Obrist, J. Hoonhout, and K. Väänänen-Vainio-Mattila, "User experience evaluation methods: current state and development needs," in *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries*, Reykjavik, Iceland, Oct. 2010, pp. 521–530, doi: 10.1145/1868914.1868973.
6. J. Hussain *et al.*, "A Multimodal Deep Log-Based User Experience (UX) Platform for UX Evaluation," *Sensors*, vol. 18, no. 5, Art. no. 5, May 2018, doi: 10.3390/s18051622.
7. S. Martin, "Measuring cognitive load and cognition: metrics for technology-enhanced learning," *Educational Research and Evaluation*, vol. 20, no. 7–8, pp. 592–621, Nov. 2014, doi: 10.1080/13803611.2014.997140.
8. "Introducing UserZoom's single UX metric for experience benchmarking: the qxScore," *UserZoom*, Apr. 15, 2019. https://www.userzoom.com/using-userzoom/introducing-userzooms-single-ux-metric-for-experience-benchmarking-the-qxscore/ (accessed Jun. 28, 2020).
9. H. W. Alomari, V. Ramasamy, J. D. Kiper, and G. Potvin, "A User Interface (UI) and User eXperience (UX) evaluation framework for cyberlearning environments in computer science and software engineering education," *Heliyon*, vol. 6, no. 5, p. e03917, May 2020, doi: 10.1016/j.heliyon.2020.e03917.
10. "gamification | Definition of gamification in English by Oxford Dictionaries," *Oxford Dictionaries | English*. https://en.oxforddictionaries.com/definition/gamification (accessed Dec. 20, 2018).
11. S. Abdi, H. Khosravi, S. Sadiq, and D. Gasevic, "A Multivariate Elo-based Learner Model for Adaptive Educational Systems," *arXiv:1910.12581 [cs]*, Oct. 2019, Accessed: Mar. 13, 2020. [Online]. Available: http://arxiv.org/abs/1910.12581.
12. W. L. in R.-B. U. Experience, "Minimize Cognitive Load to Maximize Usability," *Nielsen Norman Group*. https://www.nngroup.com/articles/minimize-cognitive-load/ (accessed Mar. 13, 2020).