Cost to develop persuasion in health behavior change support systems: A weight management app scenario

Sharon Nabwire and Harri Oinas-Kukkonen

University of Oulu, Finland

Abstract

Health Behavior Change Support Systems (HBCSS) have garnered popularity due to their intentional design to influence lifestyles and foster lasting behavioral changes. The Persuasive Systems Design (PSD) model highlights the capacity of persuasive software features to enhance the systems' ability to influence people's behavior, which holds significant promise, for instance, in reducing the prevalence of non-communicable diseases. In the medical field, HBCSS have been recognized as efficient, cost-effective, and scalable with minimal costs compared to traditional face-to-face interventions for preventing such diseases. However, every new technology comes with significant development and maintenance costs, which can either facilitate or hinder its wider adoption. The development cost may even be neglected altogether. Even if the cost was addressed somehow, evaluation methods often focus on the overall cost rather than carefully addressing the development cost of specific software functionalities and features. It is critical to make well-informed design choices rather than develop all the features that come into the designers' minds. This study conducted semi-structured expert interviews and applied the Weight Sum Model (WSM) to investigate the perceived cost implications for developing persuasive features in a weight management app. The results highlight that social and primary support features may require more financial resources to be developed than dialogue and credibility support features. Personalization and tailoring were perceived as the most expensive features due to their complex development nature. Furthermore, the results provide insights for developing HBCSS and cost-saving strategies that are important for healthcare providers, policymakers, and stakeholders in making informed decisions.

Keywords

cost estimate, resource evaluation, health behavior change support systems, persuasive systems

1. Introduction

Health Behavior Change Support Systems (HBCSS) are practical tools for altering lifestyles and encouraging sustained behavioral changes [17]. These systems leverage user experience design and behavioral psychology strategies to increase user engagement, adherence, and intervention success [8]. Using mobile apps and web-based systems, HBCSS can disseminate information, provide guidance, and help establish healthier habits [31]. At the forefront of HBCSS efficacy lies Persuasive Systems Design (PSD) [20], a framework that highlights the influential potential of persuasive features. It is used to develop, design, and evaluate these features and their impact on user attitudes and behaviors [33]. The application of this model, particularly in health, establishes that persuasive features can enhance the capability of persuasive technology to

BCSS 2024: The 12th International Workshop on Behavior Change Support Systems, April 10, 2024, Wollongong, Australia.
Sharon.nabwire@oulu.fi (S. Nabwire); harri.oinas-kukkonen@oulu.fi (H. Oinas-Kukkoken)
0009-0008-5960-6222 (S. Nabwire); 0000-0002-1442-0562 (H. Oinas-Kukkonen)

influence people's behavior. This holds considerable promise for reducing the prevalence of non-communicable diseases such as obesity and metabolic syndrome [29] [32], for instance.

With healthcare budgets facing increasing pressure, economic evaluation is becoming increasingly prevalent among organizations seeking to make informed decisions regarding developing competing health technologies. Research shows that web and mobile behavioral applications are cost-effective and more efficient than alternative traditional face-to-face interventions for preventing non-communicable diseases. For example, a systematic review determined that mobile health interventions for type 2 diabetes are cost-effective in reducing the annual patient medical costs [28]. Additionally, a study found that web-based cognitive-behavioral therapy for depression is more cost-effective than traditional treatments per quality-adjusted life year [6].

Although HBCSS are recognized for their efficiency, cost-effectiveness, and rapid scalability at small costs, it is crucial to realize that every new technology comes with significant development, implementation, adoption, and maintenance costs. These costs can potentially facilitate or hinder their wider application [27]. Technology costs are often overshadowed by the excitement encompassing technological advancements, and development costs may even be overlooked or treated as secondary concerns. Moreover, existing cost evaluation methods focus on the overall cost, and the granular evaluation of digital development costs is still lacking. A thorough economic assessment of development costs must be fully included in the economic calculations to make informed decisions [16]. Therefore, it is necessary to acquire additional evidence regarding the costs incurred in developing various features and characteristics of HBCSS.

This study explored the perceived costs of creating persuasive elements for Health Behavior Change Support Systems. Through expert interviews, we analyzed the financial implications associated with these persuasive features and identified strategies to reduce costs. The findings provide valuable information for healthcare providers, policymakers, and other stakeholders to make well-informed decisions when developing HBCSS.

2. Background

2.1. Persuasive System Design (PSD) model

The PSD model is an advanced framework for designing, implementing, and assessing persuasive technology [17, 20] Behavior Change Support Systems [17]. It defines seven key postulates that are essential for any HBCSS. These postulates include two (IT is never neutral, and people like organized and consistent information) related to user interaction and technology, and the remaining five highlight key persuasion approaches [20]. The model categorizes software features of the HBCSS into four categories: primary task support (PRIM), computer-human dialogue support (DIAL), system credibility (CRED), and social influence (SOCI). The PRIM features are designed to assist users in completing core tasks. DIAL features enhance the interaction between users and the system. CRED features make the system more believable and trustworthy, while SOCI features leverage social factors, such as comparison and competition with others, to influence user behavior. The PSD model also offers a thorough understanding of persuasive systems by examining the context of persuasion, strategies, and tactics used, as well as technical issues that must be considered [20].

2.2. Health Behavior Change Support Systems

According to Oinas-Kukkonen [17], Health Behavior Change Support Systems are designed to help people change their behavioral patterns without coercion or deceit. Using digital tools, these systems assist users in altering, forming, and maintaining healthy habits [31]. Additionally, HBCSS often use validated techniques such as cognitive-behavioral therapy to enhance behavioral outcomes [29]. Persuasive systems design plays a central role in understanding the effectiveness of HBCSS in various contexts, such as obesity [13], smoking cessation [10], type 2 diabetes [28] and coronary heart disease [1]. According to the PSD model, persuasive features can increase the system's effectiveness. Studies have tested how these components are used and aid in successfully altering an individual's behavior [1].

A systematic review [14] found that some features from the PRIM, DIAL, and SOCI categories in physical training mobile apps were utilized, whereas few features from the CRED category were used. Lehto et al. [10] found that PRIM features were more commonly used in studies that focused on alcohol and smoking interventions. A recent study that evaluated 80 mobile apps across four health domains found that personalization was the most frequently used feature (n=77), followed by surface credibility (n=69), trust (n=66), and self-monitoring (n=64) [22]. Self-monitoring is a prevalent feature of the HBCSS, as demonstrated in previous studies [1,9,10]. Research suggests that other features, such as tailoring, tunneling, reminders, trustworthiness, and expertise, are commonly used when persuading people to engage in physical activity [36].

2.3. Estimating the cost of developing persuasive features

Developing the HBCSS often involves finding cost-effective ways to integrate persuasive elements and estimating development costs [30]. Traditionally, software development costs can be estimated using expert judgment or algorithmic model estimations [23], such as the Constructive Cost Model (COCOMO) [5] or Function Point Analysis [11]. These estimation models act as a guide in software estimation of actual projects and can be used to understand the cost of developing HBCSS. COCOMO uses parameters such as code size, experience of the developing team, and project difficulty to estimate the development effort, schedule, and costs [15]. To estimate the overall cost of developing persuasive elements three parameters (cost, expertise, and effort) are used: *Cost* is described as the financial resources needed for app development; *expertise* is the skill level and experience needed to develop the app; and *effort* is the amount of time (working hours) required to develop the app. In general, these parameters offer a good understanding of the development of persuasion in the HBCSS. Naturally, to accurately assess the actual cost of persuasion, it would be crucial to consider the broader context and the potential cost synergies associated with various persuasive elements [4]. This, however, is beyond the scope of this paper.

3. Study setting

In this study, we conducted semi-structured interviews to seek to assess the perceived cost of developing persuasive features in a weight management app. This interview style allowed us to explore all relevant topics thoroughly [12]. In addition, we used the Weight Sum Model (WSM) to evaluate the cost of developing persuasive features. WSM is a multi-criteria decision-making technique that allows using multiple weighted factors in the evaluation process [34]. To estimate the perceived cost, we assigned equal weights to each evaluating parameter, namely, expertise, effort, and cost. Using the WSM helps to recognize the perceived cost of each feature systematically.

3.1. Recruitment of interview participants

A convenience sampling technique was applied to select the study's experts, which involved identifying individuals with extensive knowledge and experience in persuasive systems design and digital health interventions. This selection method was used because it is time-efficient and cost-effective. The selection process concentrated on experts involved in designing, developing, implementing, and evaluating persuasive systems for behavior changes in weight management. Seven experts were chosen, including two manager-level persons, three system designers, and two software programmers, all with more than four years of experience in digital health interventions and persuasive design. These experts were knowledgeable about PSD features and had prior experience in creating web and mobile weight management applications of a similar nature. This information was crucial in establishing the number of participants for this study.

3.2. Application scenario

An application management scenario, “HealthHabor,” targeting people aged 18 and above, was presented to the interviewees at the beginning of the interview. HealthHabor is a hypothetical mobile app that demonstrates how technology can provide personalized support to users as they navigate their weight management journey. It offers a comprehensive and customized experience tailored to users' unique needs. By using HealthHabor, individuals can transform their approach to weight loss, making it an exciting and enjoyable journey. Although HealthHabor is not an actual app, it shows the potential of combining Persuasive Systems Design (PSD) and Cognitive Behavioral Therapy principles to motivate users towards healthy behaviors.

3.3. Interviews and analysis

Following the informed consent process through email correspondence, selected participants had one-on-one online interviews with the first author, allowing for an in-depth assessment of participants' perspectives. At the beginning of each Zoom session, verbal consent was obtained to record the meeting. The interviews lasted between 58 and 80 minutes and comprised questions derived from the three stages of the PSD model, as shown in Table 1. The first five questions pertained to the functional (PSD features), non-functional (user interface and user experience design), and system (backend development, testing, and maintenance) requirements, while questions 6 to 10 addressed the PSD model’s postulates. The remaining questions, 11 to 14, focused on the context of persuasion. This questioning style enabled a targeted approach to the relevant cost of developing persuasive elements in the HealthHabor application scenario.

For this study, we focused on rating questions related to expertise, effort, and development cost (questions 1-3). As shown in Figure 1, a scoring sheet was used during the interview for questions 1-3 to assess the expertise, effort, and cost requirements for development. Before the interviews, the first author conducted two pilot interviews to test the questions and the scoring sheet. The recorded interviews were transcribed for analysis.

Table 1
Interview questions guide

PSD Model	Questions
Persuasive Features	1. Based on your experience, what level of expertise (low, medium, high) would be required when developing each of the features below in the app? 2. Based on your experience, how would you rate the overall effort (Low, medium, high) of developing each of the features below into the app? 3. How would you rate the cost (Low, medium, high) of developing each of the features below in the app? 4. Are there any budget constraints or limitations that should be considered when estimating the cost of these features? 5. Are there any specific resources, whether in terms of personnel, technology, or tools, that would be crucial for successfully developing these features?
Persuasion Postulates	6. Many devices and apps are competing for the user's attention to influence behaviors. Can you share your insights on how it might impact the cost of software/ application development? 7. How does the incremental nature of persuasion affect the cost of developing systems/ applications? 8. How much does adding nudges and cues influence overall system development cost? 9. In your opinion, to what extent does the requirement for unobtrusiveness increase initial development cost? 10. Are there any additional costs in ensuring the system is open and transparent for users?
Persuasion Context	11. a) To what extent does the application domain influence systems development costs? 11. b) b) How does it influence systems development 12. a) In your opinion, how much does personalization impact the development cost? 12. b) How does it influence systems development 13. To what extent does the technology platform affect the development cost? (e.g., IOS vs. Android; if both IOS and Android are offered if both mobile app and web app are offered) 14. What other technological aspects would affect the development cost

In the interviews, experts were asked to rate expertise, effort, and cost required to develop functional, non-functional, and system requirements in the HealthHabor app. For each requirement, the experts rated low, low-medium, medium, medium-high, and high (cost, expertise, and effort), which were then represented in increasing order from 1 to 5 (1=low, 2=low-medium, 3=medium, 4=medium-high, and 5=high). The ratings were then described using descriptive statistics. Specifically, the minimum, maximum, and median values. The sum ratings for the evaluators were also calculated and the median for each category was calculated using this information. The next step was to compare the expertise, effort, and cost ratings of each feature. The Weighted Sum Model (WSM) was applied to evaluate the overall estimates for expertise, effort, and cost. Using the direct rating method, equal importance was assigned to all the factors and allocated an equal weight (w) of 0.33, considering that skilled labor and effort are important factors that can translate into financial implications. The weighted score for the persuasive elements was determined using the following formula:

$$Weight\ Score = weffort \times Effort + wexpertise \times Expertise + wcost \times Cost$$

Dollar signs were used to describe the qualitative values based on the weighted score of each persuasive feature.

Requirements	Category	System Features	Level of Expertise					Overall Effort					Overall Cost				
			Low	Low-medium	Medium	Medium-High	High	Low	Low-medium	Medium	Medium-High	High	Low	Low-medium	Medium	Medium-High	High
Functional	PRIM	Reduction															
		Tunneling															
		Tailoring															
		Personalization															
		Self-Monitoring															
		Simulation															
		Rehearsal															
	DIAL	Praise															
		Reward															
		Reminders															
		Suggestions															
		Similarities															
		Liking															
		Social role															
	CRED	Trustworthiness															
		Expertise															
		Surface credibility															
		Real-world feel															
		Authority															
		3rd party endorsements															
		Verifiability															
	SOCI	Social learning															
		Social comparison															
		Normative influence															
		Social facilitation															
		Cooperation															
		Competition															
		Recognition															
Non-Functional	UI/UX Design (Postulates)																
System	Backend																
	Testing																
	Maintenance																

Figure 1: Questions 1-3 rating sheet

4. Findings

Overall, the grouping of experts into system designers, software programmers, and managers showed variations in their estimates. System designers displayed higher estimates of the required expertise, effort, and cost in all categories. Typically, the forecasts provided by programmers were lower than those of system designers, while estimates from experts in managerial roles generally fell between these two.

4.1. Expertise required to develop persuasive elements

The level of expertise was denoted as 1= beginner, 2 = advanced beginner, 3 = competent, 4= proficient, or 5 = expert. Under the PRIM category in Table 2, Personalization requires the highest level of proficiency to develop because of its complexity. On the contrary, Self-monitoring is perceived as the least complex feature, hence the low estimates. Depending on the context, Reduction, Tailoring, and Simulation estimates ranged from advanced beginner to expert skill level, while Rehearsal estimates varied from beginner to expert level.

In the DIAL category, significant importance is placed on Suggestion, Liking, and Social role, indicating a need for special competence. Developing these features requires consideration of both content and context complexity. Features such as Praise, Rewards, Reminders, and Similarity are perceived to require less skills than the abovementioned. In the CRED category, making a system believable requires proficient skills for effective development. The participants perceived expertise varied from beginner to expert for Real-world feel, Expertise, and Authority. Real-world feel requires the least software skills to develop in this category. The SOCI category requires a high-level competence to develop all features except Recognition, which can be accomplished with advanced beginner skill levels. Effective development of these features requires careful planning and design to achieve the desired impact.

4.2. Effort required to develop persuasive elements

Traditionally, software development efforts have been measured in individual working hours. Effort estimates for developing persuasive features differ based on complexity, as shown in Table 2. Primary task support features demand significant effort, followed by social and credibility support features. Personalization and Tailoring (for PRIM) had the highest total effort estimates, requiring extensive data collection and analysis for a customized experience. Meanwhile, Praise, Reward, Reminders, Similarity (DIAL), Real-world feel and Authority (CRED), and Recognition (SOCI) require the slightest effort to develop. These features are often relatively easy to develop and may be achieved with standard ready-made components. The remaining features across categories require a medium effort to develop them effectively. The perceived effort necessary for Expertise, Surface credibility, Authority and Third-party endorsements, Normative influence, Social facilitation, and Cooperation varies greatly as it ranges from a minimum of 1 to a maximum of 5.

4.3. Direct cost to develop persuasive features

Developing PRIM features such as Personalization and Tailoring is regarded as the most expensive among experts, as highlighted in Table 2. However, the cost can vary depending on the level of Personalization or Tailoring [19]. Providing more advanced Personalization would necessitate a lot of data collection and analysis. The cost would also differ based on whether it is done on an individual or group level; the need for individual customization will often increase costs compared to group-based tailoring. In contrast, perceived low-cost features include Praise, Reward, Reminders, Suggestion, Similarity (DIAL), Verifiability, Third-party endorsements, Authority, Real-world feel (CRED), and Normative influence and Recognition (SOCI). These features are considered simpler to develop. Although most of the features in the PRIM category were rated as having medium to high development costs, Self-monitoring required the least developmental cost within this category. The cost of development varied from 1 to 5 for Reduction and Self-monitoring (PRIM), Praise (DIAL), Third-party endorsements, Expertise and Authority (CRED), Social learning, Social comparison, Normative influence, and Cooperation (SOCI).

Table 2
Expert ratings for expertise, effort, and cost

		Expertise				Effort				Cost			
Category	System Features	Minimum	Maximum	Median	Sum Rating	Minimum	Maximum	Median	Sum Rating	Minimum	Maximum	Median	Sum Rating
	Reduction	2	5	5	28	1	3	3	19	1	5	3	19
	Tunneling	3	4	3	24	3	5	3	25	1	4	3	21
	Tailoring	2	5	4	26	2	5	4	29	2	5	4	28
PRIM	Personalization	3	5	5	31	2	5	5	30	3	5	5	32
	Self-Monitoring	1	4	2	15	2	5	3	25	1	5	2	20
	Simulation	2	5	3	25	2	5	3	25	2	5	3	22
	Rehearsal	1	5	3	21	1	4	3	20	1	3	3	16
	Praise	1	5	2	15	1	3	1	11	1	5	1	14
	Reward	1	3	2	15	1	3	2	14	1	3	2	14
	Reminders	1	5	2	19	1	4	2	15	1	4	1	11
DIAL	Suggestion	2	5	3	23	1	4	3	16	1	3	1	13
	Similarity	1	4	2	16	1	4	2	15	1	4	2	14
	Liking	1	5	3	19	1	3	3	19	1	3	3	18
	Social role	2	4	3	22	2	5	3	22	2	5	3	22
	Trustworthiness	1	5	4	26	1	4	3	20	1	4	3	18
	Expertise	1	5	3	23	1	5	4	23	1	5	3	20
	Surface Credibility	2	4	3	21	1	5	3	20	2	5	3	23
CRED	Real-world feel	1	5	1	15	1	3	1	13	1	3	1	13
	Authority	1	5	3	18	1	5	1	15	1	5	2	16
	3rd party endorsements	1	3	2	14	1	5	3	17	1	5	2	16
	Verifiability	1	3	3	17	1	4	3	19	1	4	2	15
	Social learning	2	5	3	25	2	5	3	25	1	5	3	22
	Social comparison	2	5	3	23	2	5	3	24	1	5	3	23
	Normative influence	1	5	3	19	1	5	3	20	1	5	2	16
SOCI	Social facilitation	2	5	3	21	1	5	3	22	1	4	3	19
	Cooperation	2	5	3	23	1	5	3	21	1	5	3	21
	Competition	1	5	3	22	2	5	3	22	2	5	3	22
	Recognition	1	3	2	14	1	4	2	15	1	4	2	15

4.4. Overall cost assessment (expertise, effort, and cost)

The Weighted Sum Model (WSM) was used to compare the cost of persuasive features across categories. We found that the PRIM category consistently remains at the top, followed by SOCI, CRED, and DIAL categories according to the weighted scores presented in Table 3. The development cost of primary task support features is substantially higher than that of dialogue support features. In the PRIM category, Personalization holds the highest total weighted score (30.7), followed by Tailoring (27.4). Social learning (23.8) and Social comparison (23.1) acquired the highest weighted scores in the SOCI category, and Expertise (21.8), Trustworthiness (21.1), and Surface credibility (21.1) weighed the most in the CRED category. In the DIAL category, the Social role (21.8) weighed the most. Rehearsal, Praise, Real-world feel, and Recognition have the lowest weighted total scores in their respective categories.

Table 3

Persuasive features weighted scores

Weights		**0.33**	**0.33**	**0.33**		**Weighted Score (WS)**				
Category	**System features**	**Exp**	**Eff**	**Cos**		**Exp**	**Eff**	**Cos**	**Total WS**	**$ Value**
PRIM	Personalization	31	30	32		10.2	9.9	10.6	30.7	$$$$$
	Tailoring	26	29	28		8.6	9.6	9.2	27.4	$$$$
	Simulation	25	25	22		8.3	8.3	7.3	23.8	$$$
	Tunneling	24	25	21		7.9	8.3	6.9	23.1	$$$
	Reduction	28	19	19		9.2	6.3	6.3	21.8	$$$
	Self-monitoring	15	25	20		5.0	8.3	6.6	19.8	$$
	Rehearsal	21	20	16		6.9	6.6	5.3	18.8	$$
DIAL	Social role	22	22	22		7.3	7.3	7.3	21.8	$$$
	Liking	19	19	18		6.3	6.3	5.9	18.5	$$
	Suggestion	23	16	13		7.6	5.3	4.3	17.2	$$
	Reminders	19	15	11		6.3	5.0	3.6	14.9	$
	Similarity	16	15	14		5.3	5.0	4.6	14.9	$
	Reward	15	14	14		5.0	4.6	4.6	14.2	$
	Praise	15	11	14		5.0	3.6	4.6	13.2	$
CRED	Expertise	23	23	20		7.6	7.6	6.6	21.8	$$$
	Trustworthiness	26	20	18		8.6	6.6	5.9	21.1	$$$
	Surface credibility	21	20	23		6.9	6.6	7.6	21.1	$$$
	Verifiability	17	19	15		5.6	6.3	5.0	16.8	$
	Authority	18	15	16		5.9	5.0	5.3	16.2	$
	3rd party endorsement	14	17	16		4.6	5.6	5.3	15.5	$
	Real-world feel	15	13	13		5.0	4.3	4.3	13.5	$
SOCI	Social learning	25	25	22		8.3	8.3	7.3	23.8	$$$
	Social comparison	23	24	23		7.6	7.9	7.6	23.1	$$$
	Competition	22	22	22		7.3	7.3	7.3	21.8	$$$
	Cooperation	23	21	21		7.6	6.9	6.9	21.5	$$$
	Social facilitation	21	22	19		6.9	7.3	6.3	20.5	$$
	Normative influence	19	20	16		6.3	6.6	5.3	18.2	$$
	Recognition	14	15	15		4.6	5.0	5.0	14.5	$

Using the total weighted scores, features were grouped with a range from below 17 to above 29. Features with a score between 21 and 24.9 include Reduction, Tunneling, Simulation, Social role, Trustworthiness, Expertise, Surface credibility, Social learning, Social comparison, Cooperation, and Competition. These features are perceived to require a moderate financial commitment compared to the features above. Features with a total weighted score between 17 and 20.9, such as Self-monitoring, Rehearsal, Suggestion, Liking, Normative influence, and Social facilitation, are relatively more affordable. Lastly, Reminders, Similarity, Rewards, Praise, Verifiability, Authority, Third-party endorsements, Real-world feel, and Recognition have a score below 17. Development of these features may be achieved at a lower cost than the others.

For illustration purposes, these groupings were later described using dollar signs, as seen in Figure 2. These represent the perceived monetary value necessary to develop each feature successfully, which thus also reflects the perceived economic impact of each feature in ultimately guiding the decision-making on resource allocation. The financial commitment required for developing features in the lower category is considerably less than that required for Personalization or Tailoring.

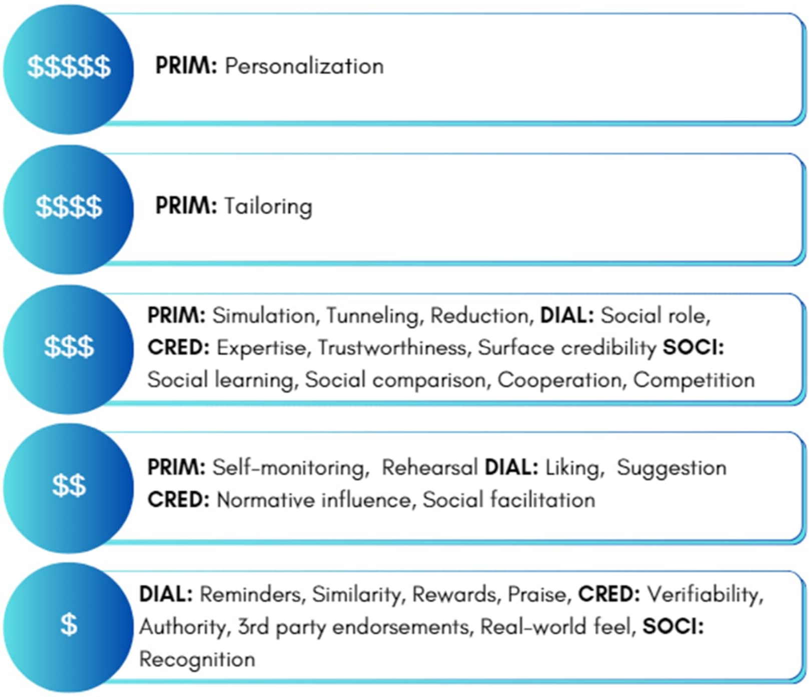

Figure 2: Feature grouping by weighted score

4.5. Non-functional and system requirements

When developing HBCSS, a strong understanding of user interface and user experience (UI/UX) design is essential as it directly affects general user appeal and persuasion indirectly. While proficient developers are needed for backend development and testing, UI/UX design is perceived as challenging and requires expert-level skills. Backend development can be time-consuming, but the effort required depends on the system's complexity. Testing and maintenance are perceived to require less effort, which, however, is crucial for software quality. Back-end development and UI/UX design are generally perceived as costly features. However, perceptions of the level of expertise required for varied widely, with some experts rating it as low as one and as high as five.

5. Discussion

This study conducted expert interviews to explore the costs of creating persuasive elements in health behavior change support systems. The results suggest that further research is needed to fully understand the costs of developing these features. The study also provides novel insights into prioritizing components across different categories. Results highlight that dialogue and credibility support features may require fewer financial resources than social and primary support features. Primary support features were found to be the most costly, indicating their significant development costs.

PRIM's Personalization and Tailoring features were viewed as the most expensive due to the complex nature of their development. Personalization involves providing individualized content to a user, while Tailoring requires the system to provide information targeted to specific user group needs [20]. However, the findings highlight that developing these features is context-dependent and depends on the level of Personalization or Tailoring required. According to Oinas-Kukkonen, the intensity of personalization development in systems varies between low-level (weak) and high (strong) personalization. Sometimes, systems might be perceived as personalized when they are not [19]. Conversely, strong personalization demands extensive data collection and analysis, in addition to the use of advanced technologies. High personalization in HBCSS helps to realize sustainable behavior, requiring more complex algorithms to understand users' needs and preferences[18]. This would require high costs compared with low personalization. Additionally, personalizing software functionality increases the difficulty level more than personalizing content through Praise or Simulation [14], ultimately increasing development costs. Overall, the scope, level of automation, and personal versus general and contextual support are important in estimating the development cost.

Low-cost dialogue, credibility, and social support components are perceived as reasonably easy to develop. In addition, some features, such as virtual Rewards and Social comparison [37], are prevalent in persuasive systems, lowering the cost owing to the reusability of components. The application of reusable components not only reduces overall costs but can also improve efficiency. Features under credibility support are perceived as straightforward, suggesting that these persuasive elements can be efficiently developed within the HBCSS without such a significant financial burden. It is important to note that while developing credibility support features may be less costly than primary or social support features, enhancing trust might still require considerable effort over a long period and increase costs.

As each feature has been rated for the perceived expertise, effort, and cost required to develop it, coupling certain features may reduce the overall development cost. Research by [26] showed that various PSD elements operate synergistically rather than in isolation. For instance, Personalization might be developed together with other features such as Suggestion and Rewards, Self-monitoring and Reminders [26]. Pairing resource-intensive features with low-cost features or combining features in primary task support, such as Simulation and Rehearsal, may reduce the overall development costs. In addition, Self-monitoring may act as a foundation feature by providing essential data that serves as the basis for developing other features, such as Personalization, Tailoring, Social comparison, and in more generally, social support. Utilizing the data acquired through Self-monitoring to customize content may reduce initial costs and enhance a system's effectiveness. However, it is important to remember that the number of persuasive features doesn't determine their effectiveness [17]. Adding too many features can, in fact, make the system less convincing [25]. Instead, designers should focus on the most important features and evaluate them against the resources (expertise, effort, and cost) needed to develop.

The findings also provide some insights into the high costs related to UI/UX design as well as backend development. Given that more resources are needed to maintain the system's usability, security, and appeal, the high costs are justified. Additionally, the cost of developing HBCSS by regulatory measures such as the General Data Protection Regulation (GDPR) [3] or Medical Device Regulation (MDR) [2] applies to digital health technologies. Although these non-functional requirements do not directly contribute to a system's core functionality, they are critical in making a successful system [21], which highlights their impact on the effectiveness of HBCSS.

This study provides insight for decision-makers, system designers, and software developers who apply persuasive systems design. Understanding the resource dynamics of each persuasive software feature can help healthcare providers, policymakers, and other stakeholders optimize resources to create effective HBCSS. While this study used weight management as a scenario, the results can help develop other HBCSS closely related to weight, like metabolic syndrome [24], cardiovascular disease, and type 2 diabetes [7, 35] or even beyond the health application domain.

The small sample size is a limitation of this study. Additionally, the subjective nature of experts' evaluations may affect the results' generalizability. Furthermore, the interviewees were provided with the application scenario at the start of the interview, which may have potentially influenced the interview scores. Notwithstanding these limitations, we believe that the method of evaluating the cost is valuable and can be replicated to determine the cost of developing persuasive features and inform decision-making.

This study examined the cost of developing persuasive features in a weight management application scenario. Interview ratings of expertise, effort, and cost were utilized. A more in-depth analysis of qualitative responses would be useful for eliciting more nuanced insights and emergent themes, thus enhancing our understanding of the complex nature of costing persuasive features. Future research could investigate the PSD model in its entirety by including persuasion context and postulates. In addition, it could explore the feasibility of reusing persuasive components in diverse contexts.

6. Conclusion

This study utilized semi-structured interviews to gain valuable insights into estimating the expertise, effort, and cost of developing persuasive elements in Health Behavior Change Support Systems. Based on expert interviews, we investigated the perceived cost implications of persuasive elements to identify the cost impact of different PSD elements and provide insights for developing HBCSS and cost-saving design strategies. These findings invite us to examine the costs of developing persuasive elements in more detail. They provide insights into how different components should be prioritized across categories. The findings also reveal that dialogue and credibility support features may usually require fewer resources than social and primary support features. Our approach to evaluating the cost can be adopted by also others to determine the cost of developing persuasive features and inform decision-making.

Acknowledgments

We would like to acknowledge the interviewees for participating in this study and providing us with their highly valuable insight.

References

[1] E. Agyei, J. Miettunen, H. Oinas-Kukkonen, Effective interventions and features for coronary heart disease: a meta-analysis. Behaviour and Information Technology. (2023). https://doi.org/10.1080/0144929X.2023.2213342.

[2] E.E.Y.F. Agyei, S. Pohjolainen, H. Oinas-Kukkonen, Impact of Medical Device Regulation on Developing Health Behavior Change Support Systems. In: Lecture Notes in Computer

Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). pp. 1–15 Springer Science and Business Media Deutschland GmbH (2022). https://doi.org/10.1007/978-3-030-98438-0_1.

[3] E.E.Y.F. Agyei, H. Oinas-Kukkonen, GDPR and Systems for Health Behavior Change: A Systematic Review. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). pp. 234–246 Springer (2020). https://doi.org/10.1007/978-3-030-45712-9_18.

[4] R.D. Banker, H. Chang, C.F. Kemerer, Evidence on economies of scale in software development. Inf Softw Technol. 36, 5, 275–282 (1994). https://doi.org/https://doi.org/10.1016/0950-5849(94)90083-3.

[5] B.W. Boehm, Software Engineering Economics. IEEE Transactions on Software Engineering. SE-10, 1, 4–21 (1984). https://doi.org/10.1109/TSE.1984.5010193.

[6] S. Hollinghurst, T.J Peters, S. Kaur, N, Wiles, G. Lewis, D. Kessler, Cost-effectiveness of therapist-delivered online cognitive-behavioural therapy for depression: Randomised controlled trial. British Journal of Psychiatry. 197, 4, 297–304 (2010). https://doi.org/10.1192/bjp.bp.109.073080.

[7] C. Koliaki, S. Liatis, A. Kokkinos, Obesity and cardiovascular disease: revisiting an old relationship, (2019). https://doi.org/10.1016/j.metabol.2018.10.011.

[8] E.G. Lattie, E.C. Adkins, N. Winquist, C. Stiles-Shields, Q.E. Wafford, A.K. Graham, Digital mental health interventions for depression, anxiety and enhancement of psychological well-being among college students: Systematic review. J Med Internet Res. 21, 7, (2019). https://doi.org/10.2196/12869.

[9] T. Lehto, H. Oinas-Kukkonen, Examining the Persuasive Potential of Web-based Health Behavior Change Support Systems. AIS Transactions on Human-Computer Interaction. 7, 3, 126–140 (2015). https://doi.org/10.17705/1thci.00069.

[10] T. Lehto, H. Oinas-Kukkonen, Persuasive features in web-based alcohol and smoking interventions: A systematic review of the literature, (2011). https://doi.org/10.2196/jmir.1559.

[11] G.C. Low, D.R. Jeffery, Function points in the estimation and evaluation of the software process. IEEE Transactions on Software Engineering. 16, 1, 64–71 (1990). https://doi.org/10.1109/32.44364.

[12] M.C. Harrell, M.A. Bradley, Data Collection Methods. Semi-Structured Interviews and Focus Groups. (2009).

[13] J.O. Markkanen, N. Oikarinen, M.J. Savolainen, V. Nyman, V. Salminen et al., Mobile health behaviour change support system as independent treatment tool for obesity: a randomized controlled trial. Int J Obes. (2023). https://doi.org/10.1038/s41366-023-01426-x.

[14] J. Matthews, K.T. Win, H. Oinas-Kukkonen, M. Freeman, Persuasive Technology in Mobile Applications Promoting Physical Activity: a Systematic Review. J Med Syst. 40, 3, 1–13 (2016). https://doi.org/10.1007/s10916-015-0425-x.

[15] T. Menzies, Z. Chen, J. Hihn, K. Lum, Selecting Best Practices for Effort Estimation (2006). https://doi.org/10.1109/TSE.2006.114

[16] S. Michie, L. Yardley, R. West, K. Patrick, F. Greaves, Developing and evaluating digital interventions to promote behavior change in health and health care: Recommendations resulting from an international workshop, (2017). https://doi.org/10.2196/jmir.7126.

[17] H. Oinas-Kukkonen, A foundation for the study of behavior change support systems. Pers Ubiquitous Comput. 17, 6, 1223–1235 (2013). https://doi.org/10.1007/s00779-012-0591-5.

[18] H. Oinas-Kukkonen, S. Pohjolainen, E. Agyei, Mitigating Issues With/of/for True Personalization. Front Artif Intell. 5, (2022). https://doi.org/10.3389/frai.2022.844817.

[19] H. Oinas-Kukkonen, Personalization myopia: A viewpoint to true personalization of information systems. In: ACM International Conference Proceeding Series. pp. 88–91 Association for Computing Machinery (2018). https://doi.org/10.1145/3275116.3275121.

[20] H. Oinas-Kukkonen, M. Harjumaa, Persuasive systems design: Key issues, process model, and system features. Communications of the Association for Information Systems. 24, 1, 485–500 (2009). https://doi.org/10.17705/1cais.02428.

[21] H. Oinas-Kukkonen, M. Harjumaa, A Systematic Framework for Designing and Evaluating Persuasive Systems. In: Oinas-Kukkonen Harri and Hasle, P. and H.M. and S.K. and Ø.P. (ed.) Persuasive Technology. pp. 164–176 Springer Berlin Heidelberg, Berlin, Heidelberg (2008).

[22] O. Oyebode, C. Ndulue, M. Alhasani, R. Orji, Persuasive Mobile Apps for Health and Wellness: A Comparative Systematic Review. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). pp. 163–181 Springer (2020). https://doi.org/10.1007/978-3-030-45712-9_13.

[23] P. Pandey, Analysis of the techniques for software cost estimation. In: International Conference on Advanced Computing and Communication Technologies, ACCT. pp. 16–19 (2013). https://doi.org/10.1109/ACCT.2013.13.

[24] H. Park, S. Jun, H.A. Lee, H.S. Kim, Y.S. Hong, H. Park, The Effect of Childhood Obesity or Sarcopenic Obesity on Metabolic Syndrome Risk in Adolescence: The Ewha Birth and Growth Study. Metabolites. 13, 1, (2023). https://doi.org/10.3390/metabo13010133.

[25] R.E. Petty, J.T Cacioppo, Communication and persuasion: Central and peripheral routes to attitude change. (2012).

[26] T. Ploug, P. Hasle, Practical Findings from Applying the PSD Model for Evaluating Software Design Specifications. (2010).

[27] K.M. Rabarison, C.L. Bish, M.S. Massoudi, W.H. Giles, Economic evaluation enhances public health decision making. Front Public Health. 3, JUN, (2015). https://doi.org/10.3389/fpubh.2015.00164.

[28] G. Rinaldi, A. Hijazi, H. Haghparast-Bidgoli, Cost and cost-effectiveness of mHealth interventions for the prevention and control of type 2 diabetes mellitus: A systematic review, (2020). https://doi.org/10.1016/j.diabres.2020.108084.

[29] Y.G. Seo, T. Salonurmi, T. Jokelainen, P. Karppinen, A.M. Teeriniemi, J. Han, Lifestyle counselling by persuasive information and communications technology reduces prevalence of metabolic syndrome in a dose–response manner: a randomized clinical trial (PrevMetSyn). Ann Med. 52, 6, 321–330 (2020). https://doi.org/10.1080/07853890.2020.1783455.

[30] X. Shao, H. Oinas-Kukkonen, Thinking about persuasive technology from the strategic business perspective: A call for research on cost-based competitive advantage. (2018).

[31] E.G. Spanakis, S. Santana, M. Tsiknakis, K. Marias, V. Sakkalis, A. Teixeira, Technology-based innovations to foster personalized healthy lifestyles and well-being:a targeted review. J Med Internet Res. 18, 6, (2016). https://doi.org/10.2196/jmir.4863.

[32] A.M, Teeriniemi, T. Salonurmi, T. Jokelainen, H. Vähänikkilä, T. Alahäivälä, P. Karppinen et al., A randomized clinical trial of the effectiveness of a Web-based health behaviour change support system and group lifestyle counselling on body weight loss in overweight and obese subjects: 2-year outcomes. J Intern Med. 284, 5, 534–545 (2018). https://doi.org/10.1111/joim.12802.

[33] K. Torning, H. Oinas-Kukkonen, Persuasive system design : State of the art and future directions. ACM International Conference Proceeding Series. 350, (2009). https://doi.org/10.1145/1541948.1541989.

[34] E. Triantaphyllou, Multi-Criteria Decision Making Methods. In: Multi-criteria Decision Making Methods: A Comparative Study. pp. 5–21 Springer US, Boston, MA (2000). https://doi.org/10.1007/978-1-4757-3157-6_2.

[35] J. Upadhyay, O. Farr, N. Perakakis, W. Ghaly, C. Mantzoros, Obesity as a Disease, (2018). https://doi.org/10.1016/j.mcna.2017.08.004.

[36] K.T. Win, M.R.H. Roberts, H. Oinas-Kukkonen, Persuasive system features in computer-mediated lifestyle modification interventions for physical activity, (2019). https://doi.org/10.1080/17538157.2018.1511565.

[37] O. Zuckerman, A. Gal-Oz, Deconstructing gamification: evaluating the effectiveness of continuous measurement, virtual rewards, and social comparison for promoting physical activity. Pers Ubiquitous Comput. 18, 7, 1705–1719 (2014). https://doi.org/10.1007/s00779-014-0783-2.