GRIFFITH UNIVERSITY

A tale of two perspectives: A conceptual framework of user expectations and experiences of instructional fitness apps

Author

Aladwan, A, Kelly, RM, Baker, S, Velloso, E

Published

2019

Conference Title

CHI '19: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems

Version

Accepted Manuscript (AM)

DOI

10.1145/3290605.3300624

Rights statement

© ACM, 2019. This is the author's version of the work. It is posted here by permission of ACM for your personal use. Not for redistribution. The definitive version was published in CHI '19: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, ISBN: 978-1-4503-5970-2, https://doi.org/10.1145/3290605.3300624

Downloaded from

http://hdl.handle.net/10072/402252

Griffith Research Online

https://research-repository.griffith.edu.au

A Tale of Two Perspectives: A Conceptual Framework of User Expectations and Experiences of Instructional Fitness Apps

Ahed Aladwan
Interaction Design Lab, School of Computing and Information Systems, The University of Melbourne
Melbourne, Australia
aaladwan@student.unimelb.edu.au

Ryan M. Kelly
Interaction Design Lab, School of Computing and Information Systems, The University of Melbourne
Melbourne, Australia
ryan.kelly@unimelb.edu.au

Steven Baker
Interaction Design Lab, School of Computing and Information Systems, The University of Melbourne
Melbourne, Australia
steven.baker@unimelb.edu.au

Eduardo Velloso
Interaction Design Lab, School of Computing and Information Systems, The University of Melbourne
Melbourne, Australia
eduardo.velloso@unimelb.edu.au

ABSTRACT

We present a conceptual framework grounded in both users' reviews and HCI theories, residing between practices and theories as a form of intermediate-level knowledge in interaction design. Previous research has examined different forms of intermediary knowledge such as conceptual structures, strong concepts, and bridging concepts. Within HCI, these forms are generic and rise either from theories or particular instances. In this work, we created and evaluated a conceptual framework for a specific domain (instructional fitness apps). We first extracted the particular instances using users' online reviews and conceptualised them as an expectations and experiences framework. Second, within the framework, we evaluated the artefact related constructs using Norman's design principles. Third, we evaluated beyond the artefact related constructs using distributed cognition theory. We present an analysis of such intermediate-level knowledge with the aim of informing future designs.

CCS CONCEPTS

• **Human-centered computing** → **User models**; *HCI theory, concepts and models.*

CHI 2019, May 4–9, 2019, Glasgow, Scotland Uk

© 2019 Copyright held by the owner/author(s). Publication rights licensed to ACM.

This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in *CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019), May 4–9, 2019, Glasgow, Scotland Uk*, https://doi.org/10.1145/3290605.3300624.

KEYWORDS

Fitness; smartphone; online reviews; expectations; experience; conceptual framework; intermediate-level knowledge

ACM Reference Format:
Ahed Aladwan, Ryan M. Kelly, Steven Baker, and Eduardo Velloso. 2019. A Tale of Two Perspectives: A Conceptual Framework of User Expectations and Experiences of Instructional Fitness Apps. In *CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019), May 4–9, 2019, Glasgow, Scotland Uk*. ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3290605.3300624

1 INTRODUCTION

Regular physical activity is a known factor that prevents or treats a range of physical [10, 37, 54] and mental [41, 62, 67] health conditions. Thus, physical activity promotes better health which is a crucial resource that allows everyday living, leading to a better quality of life [63, p.12].

Currently, more than one-quarter of adults globally (1.4 billion) are physically inactive [24], which puts them at risks of developing the aforementioned conditions. The direct health care cost of physical inactivity mounts to $54 billion globally [63, p.16]. The World Health Organisation embraced a global action plan to achieve a 15% relative reduction in physical inactivity by 2030 [63, p.21]. One important pillar of this action plan is to create active people via involving stakeholders such as research and development institutes. One research theme of such involvement is to investigate and expand the potential of digital technologies and innovations to promote physical activity [63, p.85].

Health-related apps on mobile devices are a well-known health intervention platform that targets many health conditions [36]. These apps are more accessible in comparison with traditional face-to-face interventions [56] and have many

technological features that facilitate and enhance intervention. More and more people are using smartphone devices, which positions them as a practical platform with extended reach for a broader public health impact. Smartphone apps that promote physical activity and well-being are widespread. To illustrate, there are over 100,000 health and fitness apps in the iTunes store in 2018 [55].

Studies that analyse the content and features of physical activity and fitness apps in the major app stores have found that they lack physical activity guidelines [9]. In addition, there is a need to assess their content and design quality [8], they are not based on users' involvement [9], and they lack a foundation in research theories, especially behavioural change theory [16, 18, 20, 47, 66]. Specifically, instructional fitness apps - a specific type of physical activity apps that teaches the users a specific exercise - have received little attention and often excluded from studies of activity tracking [33]. This is problematic because apps that instruct weight training exercises, for example, may lead to injury [29, 34]. This makes the investigation of their informational, instructional, and interactional aspects very crucial to safe and effective performance of weight training exercises.

The above issues indicate that designers of these apps are not taking advantage of research knowledge related to people, behaviour, and technology, leading to suboptimal designs [1, 5, 12, 22, 23, 52]. This phenomenon is a manifestation of what is known as the *research-practice* gap that has been widely acknowledged within HCI [12, 14, 23, 27, 52, 59]. Practitioners, in turn, report that academic research outcomes are not communicated in a way that is relevant to their practice and ready to be used [15, 52].

One way to bridge the gap is to describe how empirical findings affect designs or systems conceptualised as implications [61]. Another way is to ground design instances in well-established theories which might lead to the creation of intermediate-level knowledge such as strong concepts [30] and bridging concepts [19]. In the words of Hollan et al. [28, p.180] *"we believe that what was lacking was a method that could identify the critical features of the interactions between pilots and the old instrument and a theoretical language in which these features could be expressed in a sufficiently abstract form"*. This indicates the need to create intermediate-level knowledge as *"translational resources"* [15] that facilitate the design of physical activity and instructional fitness apps.

In this paper we analysed users' online reviews of the top 10 instructional fitness apps in the iOS app store. We holistically framed users' expectations and experiences in a conceptual framework (a proposed form of intermediate-level knowledge) and evaluated the framework using Norman's design principles [53] and Hutchins' distributed cognition theory [31]. They both are built on an understanding of the emotional, psychological, and cognitive aspects of human computer interaction.

This work is novel because the creation and evaluation of the conceptual framework employ different forms of user-centred design knowledge within HCI [61]: "fieldwork-design knowledge" based on analysing users' online reviews that is evaluated using "practice-informed design knowledge" and "human-science-informed design principles". Our contributions are: 1) an intermediate-level knowledge framework built from users' reviews that facilitate the design of instructional fitness apps, and 2) users' views and voices, presented within the framework categories, that can be used for design advocacy—practitioners often use resources to present and justify their design choices including data that humanise these choices based on users' problems and needs [15].

2 RELATED WORK

Our work builds on and contributes to the agenda of *fitness apps* design and use, the use of *online reviews* to understand people interaction with technology, and the creation and application of different forms of *intermediate-level knowledge*.

Smartphone Fitness Apps

There have been several research apps that investigated physical activity and general fitness training. Buttussi et al. [13] developed a system that supervises exercising to investigate context-aware and user-adaptive techniques as input for outdoor fitness activity programming and customisation. Consolvo et al. [17] studied on-body activity sensing and inference effect on encouragement and emphasised the importance of having an interactive application that allows manipulating inferred data. Ahtinen et al. [2] highlighted the role of social sharing and playfulness as motivational factors. These investigations aimed to understand specific aspects of the interaction for general physical activity. Moreover, the design of *instructional* fitness apps is often excluded from studies of fitness and activity tracking technologies [e.g. 33].

Investigation of commercial apps includes the work of Mollee et al. [49], who developed a techniques and functionalities framework for physical activity smartphone features. These features were previously reported in the literature [11, 47, 57] and were scored based on a set of physical activity apps features analysis. The framework includes categories such as measuring and monitoring, information and analysis, support and feedback, adaptation and social, usability, and other. Unlike our work,their framework followed a top-down approach and was not grounded in users' opinions of design instances or HCI theories.

Finally, Tholander and Nylander [64] pointed out the need to inform sports interactive technologies research via understanding people's first-hand experience of sports practices and technology use. Kjeldskov and Skov [35] also found that

the community extensively used usability evaluation, which led to better interfaces for people to use "a mobile device in a mobile context". The challenge today, according to them, is to design devices, services, and interactions that "fit well into people's complex lives, for work and leisure" when they use a multitude of technologies.

Users' Online Reviews

Previous research identified that it is important to qualitatively account for the users' different opinions at the appropriate level of detail. Liang et al. [39] demonstrated that app-users' textual reviews had a significant impact on app sales ratings. They observed that to get a valuable analysis, there is a need to get to the details of different consumer opinions at a finer granularity of the analysis —this can provide new insights and improve theoretical exploration. In comparison to techniques involving text mining of app reviews [3], the constant comparative inductive approach equips the researcher with sensitivity to the data that allows for multiple levels of abstraction; something that might be hard to achieve using text mining techniques.

Users' opinions and online review authenticity are important components for a proper data analysis. Li et al. [38] evaluated the impact of crowd-sourced generation of apps fake reviews. In their sample of targeted apps (114 apps with 50,461 reviews), they reported that such manipulation mainly targets apps in the game category (70%) and only 1.9% of the health and fitness apps category. Such manipulations come in bursts and use standard repeated text. Apps targeted by the manipulation were rated significantly higher on average.

Intermediate-level Knowledge in HCI

The creation and communication of knowledge is a key activity in the Interaction Design community. One approach is through designing new artefacts and evaluating them via user studies. Höök and Löwgren [30] pointed out that this is the predominate approach in HCI —using empirical research for knowledge production at the theoretical and instances level. While it is useful to create such forms of knowledge, it may be hard for designers to build upon these outcome as they are particular instances and not built around reuse across other cases [6]. However, in design-oriented research there are other forms of knowledge that are neither generalisable like theories nor particular like instances. Examples of these forms of knowledge include guidelines [40], annotated portfolios [21], design heuristics [51], strong concepts [30], and bridging concepts [19].

Höök and Löwgren define strong concepts as *"design elements abstracted beyond particular instances which have the potential to be appropriated by designers and researchers to extend their repertoires and enable new particulars instantiations"* [30, p.23:5]. They isolate elements of a particular concrete artefact and abstract them for reuse across applications, different situations, or whole design genres. Therefore, such strong concepts provide a richer understanding of different interaction design phenomena via commonalities examination. Strong concepts have the following characteristics: they have an interactive behaviour; they are situated at the interface between people and technology; they encapsulate a core design idea; and they are abstract.

Dalsgaard and Dindler states that bridging concepts are *"distinguished by their ability to facilitate exchange both ways between overarching theory and practice, rather than by being developed from theory or practice or with the specific aim of informing either theory or practice."* [19, p. 1637]. They are similar to strong concepts in that they also can be derived from particular design instances, but they differ since they can be developed from theory too.

In our view, while these two recent approaches differ, they are both focused on knowledge reuse. Their fundamental contribution to the field is to allow for abstract concepts to be utilised across applications and use situations. In our work, we have no preconception of what form of intermediate-level knowledge has better advantages compared to others; we think this can only be judged by those who evaluate and utilise them. However, what we *can* contribute here is a suggestion of how to generate this form of knowledge for our specific case study inspired by the above two forms of intermediate-level knowledge and very recent developments [6].

3 METHOD

Our approach is to qualitatively analyse users' online reviews of instructional fitness apps. We address the question: what are the reviewers' expectations and experiences when they interact with these apps? We construct our conceptual framework based on an analysis of these expectations and experiences.

Data Collection

The focus of our data collection was on apps that prescribe (text, voice, images, video) body-weight or weightlifting exercises as they teach a motor-skill through instructional content.

We began by identifying the top-10 instructional fitness apps in the iOS Top App Charts in the category health and fitness. These were: *Full Fitness Exercise Workout Trainer, 7 Minute Workout Challenge, Fitness Buddyplus Gym Workout Exercise and Home Trainer, Livefit with Jamie Eason, Modern Physique with Steve Cook, Jefit Pro Workout - Fitness & Exercise Tracking System, 12 Minute Athlete HIIT Workouts, Gymstreak Pro - Bodybuilding Tracker, Fitness and Bodybuilding by Vgfit, Bodyweight Training You Are Your Own Gym.*

We first obtained apps ranking using *AppAnnie* [1]. The advantage of this utility is that it keeps historical records of

[1] App Annie, App Store Stats, https://www.appannie.com

the ranking. The date of the sampling was August, 1st 2017. Apple does not reveal the exact details of the iOS Top App Charts ranking methodology. However, it is generally accepted within the App Store Optimisation (ASO) community that these factors include: the average app store rating; rating/review volume; download and install counts; uninstalls (retention rate); app usage statistics; and weighted recent growth trends [65]. We believe this approach is superior to focusing on only one attribute—such as downloads or rating. It has the secondary advantage of minimising the effect of fake reviews on ranking. The list of chosen apps does not include any fitness app for other sports like running, swimming, or cycling. We also discarded apps that require additional accessories or wearables because we wanted to investigate the smartphone as a device for exercise instruction and prescription. We only considered paid apps because previous research showed that the length of the review is correlated with the price of the app [44, 45] and a longer review increases our chances of finding meaningful content. An additional advantage is that paid apps usually contain more features compared to free versions.

We chose *Appbot* [2] as a platform for collecting reviews. The search keywords in *Appbot* were: instruction, video, picture, photo, image, and text. These were chosen as they described content related to exercises and thus they supported our goal of understanding the instructional aspects of the apps. We collected 450 reviews and we scanned them to determine the range of keywords that describe exercises. Based on this, we expanded the keywords collection to 37 keywords with a focus on instruction attributes (hear, listen, audio, sound, voice, tone, pitch, see, video, movie, film, motion, animation, avatar, figure, image, picture, photo, illustration, graph, read, text, word, sentence, paragraph, learn, understand, remember, follow, repeat, replay, play, recognise, recognise, instruction, form, information). Based on the revised keywords we compiled 2536 reviews written between September 1st 2014 and September 1st 2017. We discarded any repeated review from our dataset and minimised the fake reviews burst effect [38] as we sampled reviews over a 3 year period. The reviews information included app id, app name, review id, country, ratings, date, and review text. This became the dataset for our analysis.

Data Analysis

We employed a constant comparative method for our analysis [50]. Because we have a large dataset, we decided to systematically sample the data and then iteratively analyse it until no new insights were revealed. We started with an initial sample that consisted of every 10th review totalling 230 reviews. In stage two, we sampled every 5th review, an additional 214 reviews, which resulted in 444 reviews in total. We reached data saturation after this stage indicating no further need for sampling. The first author engaged in line by line coding of the reviews. After the first iteration of open coding, further iterations included axial coding to consolidate codes under coherent groups. Then, through constant comparisons, these codes formed categories and core concepts that resemble a conceptual framework. For coding we used QDA Miner as a qualitative analysis tool [3].

4 FINDINGS

We consolidated the different categories of users' expectations and experiences into a conceptual framework. We categorised these expectations and experiences into aspects related to the content of the apps (contents), the functional utilities of the apps (utilities), and the psychological features of the interaction (character). Figure 1 shows the conceptual framework. When it comes to information diversity across the different framework categories, we found that content codes represented 47.2% of the total codes, utilities 39.7%, and character 13.2%, respectively. Moreover, the majority of the reviews covered the top 4 apps. Below we describe the framework in greater detail and highlight its underlying properties and attributes, using verbatim quotes from users' reviews.

Content

When users choose an instructional fitness app, they expect specific content (fitness workouts, exercises and their instructional format) that helps them achieve a specific outcome —improved fitness. Within the content construct, we identified distinctive content types, namely, instructional exercises and workouts. Fitness exercises are physical activities that are prescribed using different instructional forms such as text, static images, dynamic visualisations/animations, audio instructions, and videos. They are grouped by a specific purpose or goal in a workout program. We categorised the content attributes as informational, organisational, and instructional.

Informational Aspects. This group refers to the instructional content as information used for learning a new motor skill, which we classify under the headings: clarity, granularity, control, and correctness. To illustrate, **clarity** refers to the ability to decode instructions and understand them as a basic user need—without clarity, it would be impossible to understand or successfully complete the exercise. For example, text must be intelligible, and videos should not have distortions. *" ... Loads of exercises and workouts with clear pictures." Review #1410. " ... when I go to the explanations how to do something, sometimes the words overlap and you have to either guess what it's saying or just focus on what colour to read." Review #40.*

[2] Appbot, App reviews and ratings analysis for mobile teams, https://appbot.co)

[3] The QDA Miner project file is included in the supplementary material.

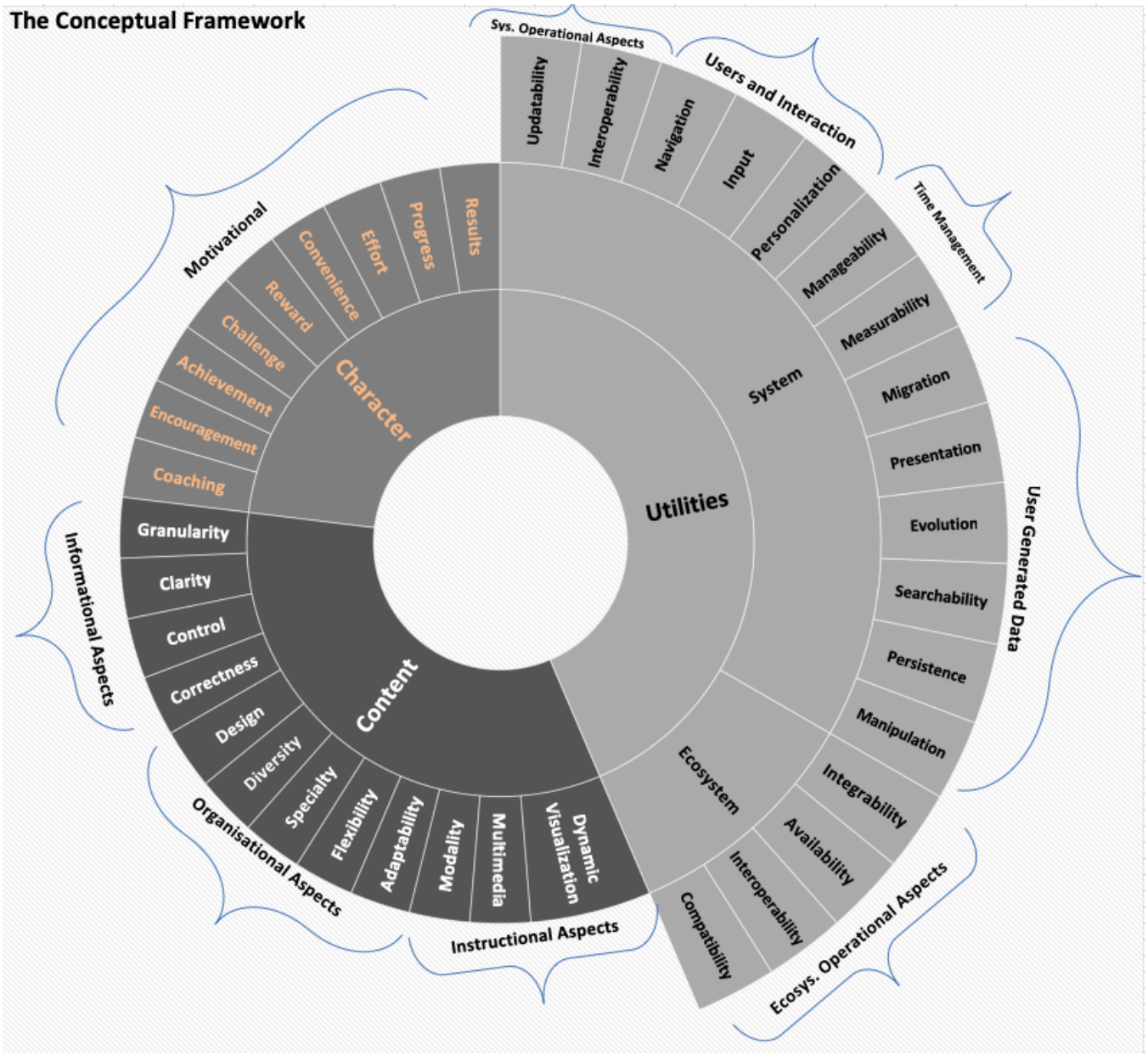

Figure 1: The conceptual framework main categories of content, utilities, and character are at the centre. As we move outward, each one branches out to include their attributes. The external labels/annotations group the outermost attributes based on function, purpose, or role.

Many users pointed out their preference to **control** the pace of information flow which allows them to understand exercise instructions better. In some cases, users suggested having the ability to control and rotate the animated avatar/model showing the exercise. *" ... can't move the 3d model to see the right position" Review #1965. " ... or stop motions of the exercise so you can see how it's properly done" Review #900.*

More reviewers indicated the importance of the exercise instructions' **correctness**. Having correct instructions on how to perform an exercise is very important and can prevent injury. *" ... I found that the internal and external rotator exercise pictures should be reversed. The internal is showing external pictures and vice-versa. I hope no one is trusting this app ... clearly it is not accurate." Review #240.*

Other reviewers referred to the **granularity** of exercise instructions, i.e. the level of detail. For example, some users requested the ability to see a demonstration of the exercise from different angles or to have step-by-step instructional photos. *" ... Lacks multiple angles for each shot" Review #1515. " ... and step by step guidelines for all the workouts!" Review #1805.*

Organisational Aspects. The organisational aspects answer user questions such as: why this content?, what purpose does the content serve?, and how can I organise content to achieve my goals? For example, workouts are a collection of exercises grouped to deliver a specific training program. Users are interested in attributes related to workouts organisation such as design considerations, diversity, speciality, and flexibility.

To illustrate, a workout **design** philosophy refers to the training programming approach to build an exercise program. Examples of these are hypertrophy, strength, or endurance development. *" ... I share Mark's philosophy of building lean muscle to boost metabolism rather than just increasing cardio workouts." Review #2220.* Users also indicated the need for scientific evidence to support training programming and design choice. *" ... That being said, the research behind this study isn't reliable. There's absolutely no data to support the use of this app's workout routine." Review #1015.* One user mentioned the need for a mechanism to create random workouts based on time, equipment, and goal. *"Random exercise generator. Either in the main body part section ... , or in the specific body part section that will allow you to get a random exercise for whatever part u want." Review #1570.*

Another important quality consideration is the **diversity** of workouts and exercises. This allows users to choose and perform designed and tested training programs. *"An extensive array of workouts for beginners to gym rats." Review #560. ... It has a ton of workouts ..." Review #110.* For exercises, diversity refers to the number of different exercises for the same muscle group. *"There are also variations for each exercise so you can make the workout easier or harder depending on your current ability." Review #2320. "There are a ton of exercises for each muscle ... figure out what's best for you and get a routine" Review #855.*

Another highlighted feature is the **speciality** of workout programs to cover special purposes such as a specific sports training, training around an injury, or time/space limitations. *" ... This app is so simple and contains workouts specific to certain parts of the body, and specific sports ... " Review #290. " ... only thing that I would recommend ... is to have more training "styles" and to go more into detail with those styles. Say, Mountain Climbing ..." Review #120.*

Furthermore, reviewers demanded **flexibility**, meaning the possibility of carrying out the workouts in different environments using different or no equipment at all. *" ... As a father of young kids, having an effective, motivating strength building program I can do at home is great ..." Review #2220. " ... no need to invent workouts, no need for gym or fancy workout items - just door, chair and a floor :)" Review #2130.* Lastly, reviewers repeatedly referred to **adaptability**, where exercises should adapt to their age, skill, and fitness level in addition to showing regression and progression variations. *" ... as you can modify the exercises to match your fitness level across all types of movements ..." Review #2220.*

Instructional Aspects. This category covers aspects related to instruction design such as comprehension and legibility. For example, **multimedia** refers to the use of different instruction formats, like text, picture, and video to help users learn an exercise. Users referred positively to different combinations of media, notably videos and step-by-step pictures. *"It's still a good app, and one that I could certainly recommend, especially for the nice touch of the instructions, photos and video clips" Review #1150.* Users referred to different instruction **modalities** indicating their adequacy and appropriateness. Singling out a specific modality such as audio or visual indicates that users need and value instructions via different modalities. For example, users referred to audio/voice instructions as clear or motivational. *" ... The audio cues during the workout are perfect!" Review #950." ... voice commands are motivational" Review #1090.*

Moreover, visual instructions are the number one modality that users prefer because it allows them to understand and learn a fitness exercise thoroughly. *" ... really like the videos to explain the exercises if you have a question about the correct form" Review #415.* Using **dynamic visualisation** and animations improve learning over textual or static pictures. Dynamic visualisation shows a human or an avatar performing a physical activity continuously. *"... and the animated pictures showing how to do a move is great" Review #1525." ... demonstrations are using static photo, don't waste your money on this app" Review #285.*

Utilities

Utilities is the collection of features, services, components, and their attributes that facilitate delivering the app's core purpose. The utilities of an instructional fitness app provide features that allow it to function within the context of training delivery as an everyday activity. The **system** dimension within utilities examines different attributes—and their properties—at the device app, software, services, and hardware level. Here we consider attributes of local nature, meaning, not part of the broader environment beyond the device. This comprises all the different components that make a system (in this case a smartphone OS and hardware). The **ecosystem**, by contrast, refers to the collection of devices, applications, and services that are beyond the smartphone's

system (app/software/hardware) level. We grouped different aspects related to utilities under topics such as user-generated data, time management, users and interaction related aspects, system capabilities, and ecosystem data and devices.

User Generated Data. The interaction between the user and the device may generate data. The data attribute describes how the system deals with the user data at both the system and ecosystem and what kind of features/properties it should exhibit to fulfil their needs. For example, data **manipulation** refers to tasks, features and options within an app that allows the user to manipulate data. It includes activities such as managing workout collections, creating templates, and modifying workouts. *" ... you can even create your own personal workout ..." Review #1.* Moreover, users prefer to have specific functions that help them manage and organise their exercises. Examples of these include adding exercises to their programs during workouts, adding notes related to their performance, modifying exercises settings, and editing logged data. *"It would be nice to be able to add specific notes to each exercise in a workout." Review #1660. "Settings for exercises are too basic; no super sets or tri sets, not able to set different repetitions for different sets." Review #1535.*

Many users pointed out their need for data **persistence**, meaning that the user should be able to have their data backed up or restored whenever they deem necessary. This is an important property within an exercising context because the overall progression of the exercise regime directly influences performance. Users need to have access to their data when offline too. *" ... I have a TON of progress that I kept in there, only to be disappointed that I can't access or see it anymore ..." Review #200. " ... internet required to watch 10 second videos??? Why not embed them into the app so they can be watched anywhere ..." Review #195.* Another aspects is the ability to **search** workouts, exercises, and logged data accurately and using multiple filter criteria that produce consistent results. *"Provide a search option in the "Exercises" section to find a specific exercise rather than having to scroll to find it." Review #2230.*

One of the most important reported aspects is data **evolution**. Evolution of data in the context of exercising describes how data change over time reflects the evolution of performance, physical attributes, and exercise programming. For users, this is an essential utility that allows them to quantify regression or progression. *"When I work out, I like to see what I was able to perform last week in regards to weight and reps ... Keeping track of what I was able to do was the whole point of this app, and now that's gone." Review #550.* The following review speaks loudly to the significance of data evolution: *"Where the heck is my workout history? Before I could look at my custom workout and see my progression over weeks. Now it's gone. Why would you take something like that away? Isn't the whole idea to TRACK OUR PROGRESSION." Review #1620.*

Moreover, reviewers repeatedly highlighted data **presentation** issues. For example, showing graphs of exercise attributes and performance evolution, showing the remaining exercises during a workout, or showing the next exercise to be performed. *" ... It is more difficult to see what exercises I have left to do ..." Review #1600. "It's easy to know what's coming up next and check your form." Review #1145.* Additionally, exercise analytics is an integral part of performing exercises as it provides the user with interesting insights. *"Would love some form of graphing of activities to watch progress, or something to that affect." Review #2395.* Finally, users referred to data **migration** where they prefer to export raw data to a different system for better graphics or more analysis. *"I use this app specifically for daily weight tracking. The one thing that I wish I could do is export the 6 months of data I've collected, so I can see that data on a bigger scale." Review #45.*

Time Management. Within an exercising context, time is a valuable resource. Not having time is one of the most common reasons users give for not exercising; therefore users appreciate features that allow them to manage their time. Utilities include **measuring** time by using timers; features such as the ability to adjust timers; and to have them run along with other utilities, such as music, without interruption. Counting is another feature that is important and relevant within an exercising context too. *"The timers are automated and integrated into the workout so you don't have to keep track of the intervals." Review #1945. "Counting reps when doing what Lauren calls Ladders without the app can be a royal pain the app helps solve the issue." Review #2510.* Scheduling exercise and workout sessions allow users to plan and **manage** their time effectively. *" ... Plus the scheduling tool and free cloud backups- you guys are great. Thanks so much!!!" Review #740. "Nice addition with the calendar." Review #2210.*

Users and Interaction Related Aspects. A unique user usually uses an exercise app. Utilities that allow **personalisation** based on users' identification, preferences, and needs are essential. *"When you click on "calorie counter", then "user profile", it does not display the name of the current user, whereas the application claims to be able to manage several user accounts." Review #920. "Despite the change of language, all texts remain in English." Review #920. "... It would be good if it were to take the region from the system settings and display the appropriate date format, e.g. DD/MM for Europe." Review #1105. "I really wish this app would let me play my own music list ..." Review #1210.*

Some users highlighted general **interaction** aspects of the system such as input, navigation and user interface. To illustrate, reviewers want to be able to customise data **input**. *"Even adding a feature that lets you just name your own exercise and put it in your workout would be sufficient" Review #1970.* They also prefer a practical, predictive and convenient data input

method. *"After each set I have to re-type in the weight and reps even if they are the same as the previous set, which a great deal of the time is the case, which adds wasted time to my workouts" Review #545. "It no longer takes advantage of predictive text in the notes ..." Review #770.* Some users demanded convenient **navigation** that allows them to navigate the app easily especially during exercise, because exercising can hinder proper human-device interaction. *"Too much going on. Too much to sift thru when trying to workout" Review #1570. "Navigating the workouts that I created myself (and therefore am intimately familiar with) is now impossibly confusing" Review #1625.* Some users suggested a friendly navigation method via voice command. *"As soon as this app gets voice input I'll rate it with five stars. This would be useful in order to use the app handsfree :)" Review #2315.*

Finally, a few reviewers commented on **user interface** issues using terms such as 'sloppy', 'clunky', 'worthless', 'simple', 'causes eye strain', 'has a nice sharp look', 'nicely designed', or 'has a steep learning curve'. *"This app is not worth its steep learning curve. It takes ages to create custom workouts. It is not easy to see the notes you create during a workout either, rendering the app almost useless. I want to see more info than the weights and reps from my last workouts." Review #1405.*

System Operational Aspects. System operational aspects refer to different components, functionalities and attributes at the system level and how they operate as a whole unit. For example, reviewers noted issues related to how different functionalities, apps, and services **interoperate**. For example, app status persistence when switching to another app and back (iOS multitasking). *" ... and most annoyingly, switching to another app and back again often causes the app to come out of your currently selected exercise." Review #255.* Another example is hearing the exercise app timer while playing one's music. *"You can play your own music and that won't affect the timer" Review #1950.* Or, that the sleep mode will not affect app status. *"voiceover that works with the display in sleep mode" Review #1060.*

Finally, the app should stay **up-to-date** to keep evolving with the system software/hardware upgrades and development. Reviewers complained of being unable to keep the app up-to-date with new devices and operating system introductions. In some cases, they were not happy with an update and wished for the possibility of rolling back. *"This app hasn't been updated in over a year, the ui hasn't been updated for the latest iOS version or the iPhone 6/6+" Review #150. " ... how do I go back to the old version?" Review #630.*

Ecosystem Operational Aspects. Data within an ecosystem context refers to its availability on different mediums and platforms and the integrability with other services and devices. Reviewers highlighted the need for data to be accessible and **available** on different mediums such as websites, desktops, tablets, and wearables. *"But the results tracker, that tracks my weight, was blank on the iPhone, though the data was still there on the iPad mini" Review #1150. "but would just like to ask why you don't see your custom workouts on the Apple Watch?" Review #2065.*

Many reviewers pointed out the importance of **integrating** the app data with health services, particularly Apple health kit, which is usually used by other apps and devices as a central repository for health data exchange. *"However if we get a fitness app that's compatible with the apple watch cant we also make it compatible with the health app etc. It's such a shame to do a great workout but it doesnt get counted along with my cyclometer or activity data." Review #1130. " ... however these days customers expect to see a health app integrate with other apps, health devices like the hambone, and especially the apple health app." Review #150.* On the other hand, other devices might exist within an ecosystem which necessitates compatibility.

Here **compatibility** refers to the app being operational and with the same look and feel on different platforms within the ecosystem. *"It was built for iPhone 4's and doesn't fit well on the iPhone 6 or 6s" Review #400. "Apple Watch does not work (hangs on an empty 'Next' screen)" Review #1120. " ... and does not display on a Braille reader." Review #1310.*

Previously we described **interoperability** at the system level, in an ecosystem context it refers to the app's ability to synchronise status and operate on different devices simultaneously. It implements features such as iOS handshaking. *"Would be so much better if you could just look at your watch to see your next exercise instead of getting your phone out every time." Review #2065.*

Character

The character construct of our framework refers to psychological and experiential aspects of users' interaction with their instructional fitness app. Users highlighted a few motivational aspects of the experience in their reviews such as motivation by results, progress, challenges and achievements. Here we considered all the different reasons why users kept using the app; or in their words, what 'kept them going'.

Motivational. The motivational aspects of the experience have been highlighted in users' reviews, including observations related to motivation by results, progress, challenges and achievements. Some users pointed out that seeing **results** is important for them and one of the main reasons to review the app positively. *"I started using this app when I was a chubby (fat) 14 year old. After using this app for almost 2 years I can safely say that this app is the best workout app available. Not only does this workout trim fat but also tones the body. This workout app has toned and tightened my abdominals and pecs."*

Review #1170. "And when I see my arms and legs getting more defined, it keeps me coming back for more." Review #1110.

Other users were motivated by **progress** which relates to a situation when the app has the feature of measuring and showing progress. *"It's motivational and so nice to see where you are compared to where you've been" Review #845. "it really keeps me motivated to work out because every time I log a set I see myself pushing toward a goal instead of just randomly doing things." Review #1495.* Some users pointed out that they enjoyed putting an **effort** and to 'break a sweat'. *"I feel like I've done a lot in the few minutes!!!!" Review #1045. "Intense workout. Have had it for 3 days now. By the 4th exercise, I was absolutely pooped! It worked every muscle in my body." Review #1165. "This app is amazing. It really gets u tired and heart pumping" Review #1245.*

Moreover, many users reported the importance of exercising at **convenient** times and convenient places. *"Great app so easy to follow and can use anytime anyplace - working well so far! Should be easy to keep up the motivation for the above reasons!" Review #1190. "For those with "not enough time" there is no excuse it's just 7min." Review #940. "For me this app really takes the stress out of exercising for me." Review #1140.*

Other reviewers highlighted the motivation by **reward** theme. *" ... and the little rewards you earn are very motivational! Worth every penny and more!!!" Review #965. "it keeps track of which days you use it and rewards you with updates if you work out for long enough in a row." Review #950.* Users pointed out another motivational factor around framing the exercising routine as a **challenging** endeavour. *"I use it like a video game to beat my previous months progress" Review #300. "The workouts challenge you to see how many sets you can do without breaking." Review #1245.* Reviewers made reference to the **achievement** motivational aspect —unlocking additional features, workouts, or levels. *"However, I notice that even though when I complete a workout it says that I have hit certain achievements (e.g. "3 Days In A Row", "1 Month Non-Stop", "Never Skipped A Beat"), these achievements do not get updated in my list of achievements." Review #1115.*

Some apps provided **coaches** that guide users throughout the exercise, be it a virtual or human coach. Reviewers reported that they found it very motivating to be guided by a coach during exercises. *"Guided personal programs with your own virtual trainer" Review #2465. "I liked having Mark guide the workout because it made it easier for me to do it" Review #2400. "this version has fewer fun options like different personal trainer voices." Review #1335.* Some reviewers appreciated encouragement strategies such as encouraging audio cues. *"Voice queues keep you going." Review #995. "Thought there would be music or a person helping to coach/urge you along ..." Review #1345.*

5 DISCUSSION

Our conceptual framework is a form of intermediate knowledge describing mental states, expectations, and experiences of instructional fitness apps reviewers. The framework gives designers insight into users' priorities, voices, and opinions conceptualised in a holistic manner, rather than as individual instances or isolated insights. Analysing the first-person reactions and reviews of application users can support the creation of designs that are reflective of real-world experience while being easy to use and understand. This speaks to Norman's design principles [53] and Hutchins' distributed cognition theory [28, 31], which we will use to ground our conceptual framework in theoretical thinking related to (and beyond) the artefact.

Norman's design principles [53] are built on an understanding of psychological, emotional and cognitive aspects of technology use, providing a suitable guideline to evaluate the immediate user experience at the artefact level, as demonstrated by [43]. Distributed cognition theory [28, 31] evaluates beyond the artefact aspects concerning information flow within the whole ecosystem between different apps, services, and stakeholders. This perspective is crucial because it allows understanding and framing users experiences and expectations when they transcend the immediate user experience of the artefact and reflect on other apps, services, and actors within the distributed app ecosystem. The work of [32, 46, 60] are examples of its use to evaluate cognitive systems, while [26, 58] suggested to use it to evaluate HCI aspects in a healthcare context.

In the discussion below, we will first introduce a theoretical or design principle, then we will discuss them in view of relevant users reviews from our data sample, and finally, we will explain how our conceptual framework bridges the gap between both.

Rethinking the Goals

At the centre of Norman's gulfs of execution and evaluation and the seven stages of action is a user goal [53]. Many of the reviewers highlighted situations where they did not know how to plan, specify, perform, perceive, interpret, or compare an action. Capturing pitfalls during these stages allows designers to enhance the design of their apps incrementally. For example, a few users reported confusion when trying to enter their exercise data. Their goal was clear for them, and they expected the system to show how to do it, but to no avail. Others reported confusion related to evaluating their actions, but as Norman has pointed out, they blame themselves instead.

Radical ideas require rethinking the user's goal [53]. For example, one reviewer wanted to stop the motion of the video so they could understand the proper technique. Another reviewer pointed out that the exercise instructional video shows

one view, usually from the front. Here, there is a difference in goals between the designer and the user; the designer wanted to show the exercise execution, the user wanted to learn as they do in their natural setting by observing someone executing the exercise from multiple viewpoints. There is a difference between showing and teaching a motor skill. A radical enhancement is to rethink what the user's real goal is—why some are willing to step out of their comfort zone and start exercising. To elaborate, a user wants the exercise app to track their fitness progress and use that information to program and customise their future workout. That is a different goal compared to a design that focuses on a utility that shows workout history. Another user demands integration with other apps within the ecosystem. From the design perspective, this could be a 'nice to have' feature, but for the user it is why they chose the app in the first place—their goal is to lose weight, and to integrate their exercise app with their nutrition app is crucial.

The question that should concern the designer is not what kind of exercises should this app have, but rather, what is the fitness goal of the user and how do we accommodate different goals for different users—not only at the app level but within the ecosystem too. Our conceptual framework responds to many of these design goals by including both instructional aspects and informational ones so as to maximise learning. It also highlights design thinking at two different levels (the system and the ecosystem) which helps the designer to conceptualise broader design issues and challenges.

Designing for all Levels of Processing

Norman defines three different levels of processing —visceral, behavioural, and reflective [53]. In a physical exercise setting, especially at the gym, people experience stress, anxiety and anticipation. This necessitates an understanding of how the design of the app relates to visceral responses. In our analysis, users referred to the stress of the exercising environment, like high heart rate, sweating, and being out of breath. Yet it remains a design challenge to understand visceral response within this context and how to design for "attraction or repulsion" [53].

At the behavioural level the subconscious mind is in control, and this is a typical situation within a sports context according to Norman [53, see p. 51]. We are unaware of our actions as our responses are too quick for our conscious mind to process. This could relate to the learned skills aspect, meaning that a user develops specific rituals and automatic habits as part of their exercising routine and their instructional fitness app becomes part of it. This could explain why many users' reviews communicated anger and frustration after an app update that affected the app behaviour; thus, their automatic habits. One could imagine a user having a physically stressful exercise session and trying to understand how to deal with their app after an update; a clear example of an expectation violation in a stressful situation.

Finally, an instructional fitness app helps users to learn a new motor skill and to reflect on their progress and plans. These are clear markers of the reflective processing level; which, according to Norman, is perhaps the most important processing level to the designer. We share this view and believe that instructional fitness apps operate mostly at the reflective level. We argue that learning and evaluation requires understanding and analysis, which is mostly *"cognitive, deep, and slow"* [53]. The conceptual framework includes content aspects that maximise learning and utilities that allow data analysis and progress tracking.

Designing for Flow

Norman [53] refers to Csikszentmihalyi's concept of *flow* as an emotional state in which the person becomes one with the task they are performing and feels fully immersed in the activity without any sense of time or the outside environment [48]. In a physical exercise setting, users did describe such a state, and they referred to the importance of having a technology that supports and sustains that: *"The app now inexplicably jumps into some screwy landscape mode during a workout and won't return to normal making it impossible to log exercises, and suddenly you find the* ***flow*** *of your workout ruined while you fuss with your phone." Review #1670.*

It is clear here that any issue at any level can affect the flow state, be it at the usability or user experience level. Designing for flow in our opinion includes all the above issues, from understanding the user's goals, accounting for their expectations, and the design for all processing levels. Our framework allows designers to think about these issues holistically, thus incorporating many aspects that ensure a smooth and robust experience. Recent work by Havlucu et al. contributes to this by highlighting how rituals, distraction, and interruption might affect flow experience [25]. They indicated the importance of detecting the flow state and provide unambiguous feedback to sustain it. We agree that these are legitimate design challenges and future research topics.

Designing for Distributed Cognition

Hollan et al. [28, p. 175] state that distributed cognition theory *"provides an effective theoretical foundation for understanding human-computer interaction and a fertile framework for designing and evaluating digital artefacts"*. The theory identifies cognitive agents, processes, events within an environment and beyond the *"skin or skull of an individual"* [28]. For example, in a case study of an aircraft cockpit [32], what concerns a passenger is not the set of skills the pilot has, but rather how the whole cockpit environment that involves people, technology and the interaction among them function. In an

analogous way, but for a different unit of analysis, our evaluation includes the user; the app within the boundaries of the smartphone; and the app beyond the boundaries of the smartphone—the whole ecosystem. The question for us is not to capture how well the app is performing, but rather, how the user view the interaction between them, the app (system), and the whole ecosystem of other apps, services, and technologies. A more critical question to answer is whether the interaction between all these different elements serves the user's goal. In other words, by capturing the user's view within such boundaries, we hope to formulate a new understanding of the whole interaction among users, system, and ecosystem of instructional fitness apps. We hope such an understanding can inform future discussion and decisions related to the evolution of instructional fitness apps.

Back to the Future. Actions, events, and processes within our unit of analysis have a temporal aspect. While all of these can be treated as isolated incidents, a holistic account to understand how earlier incidents affect later ones becomes necessary. Hutchins and Klausen [32] refer to this as *"automated gathering of activity histories"*. In the cockpit example, the Captain used information such as current and cleared altitudes and all information available about current and previous performance to construct a plan of action on how to get to the cleared altitude. The amount of distributed information that describes the current system state is the tip of a previous and current information iceberg. In a very recent work, Rogerson et al. [60] refer to a similar phenomenon, but in a different context—how knowledge of a boardgame's state is distributed amongst and moves between players and material elements.

Once we reflect on this, in a setting where the fitness goal is to lose weight or improve physical strength, one cannot underestimate the importance of a related benchmark and the continuous performance measurement against it. It is clear that such a setting is not limited to a single attribute, such as exercise frequency for example, but a multitude of variables distributed within a cognitive, physical, and social context. Information in these contexts has different sources, trajectories and representation as it propagates through the environment and it is crucial to process and frame this information in a way that is meaningful and useful for the user. In other words, collecting information about all the aspects of an instructional fitness app use will inform future use occurrences. As we have outlined in our findings, for some users, that was the sole purpose of the app and that is represented in the user generated data section of the conceptual framework.

Show me the Data. The concept of *"zoomable multiscale interfaces"* where *"zooming and panning"* [7] are important interactional aspects in a system of distributed cognition is very relevant, in our evaluation, to finding and 'seeing' data. As we have discussed before, data is tied to actions, events and process and is provided within a temporal/historical context. However, searching, filtering, and viewing data within a distributed information environment of virtual *"multiscale spaces"* [7] is challenging. This speaks to the concept of moving closer to data and the magnifying optics metaphor [28] due to the distributed nature of data and its volume. In the context of instructional fitness apps there are multiple agents such as the user, the personal trainer, the nutritionist, and the physiotherapist who might all at one time generate and consume data. How to find, filter, and graphically represent relevant data to a specific use case is a challenge. This explains why data has the most diverse attributes in our conceptual framework.

Intersubjectivity Matters. Designing for intersubjectivity plays a significant role in instructional fitness apps because different actors are involved in the process of generating information and hence affect its trajectory. A system of distributed cognition requires a shared understanding of how the system operates—such understanding is based on the different actors' expectations, experiences, or methods (in the case of non-human actors). The very nature of such a system requires coordinated action grounded in shared knowledge to make the overall goal attainable. Intersubjectivity builds on and makes use of information distribution and representation discussed previously. Albinsson et al. [4] refer to some intersubjectivity concepts in the relationship between a client and a trainer. Concepts such as a trainers' philosophy, vision, and what perfection means are negotiated in terms of the client goal, age, and physical abilities to reach a shared understanding. While it is crucial for such a negotiation to happen, we recognise the challenge of mediating relationships via technology. For example, in the cockpit case study, the "silent look" metaphor explains a very sophisticated form of intersubjectivity that is very hard for a non-human actor to use for interaction. While that is true, from Norman's analysis of the autopilot example [Norman, 1990], we realise the importance of intersubjectivity to include non-human actors. Meaning, as we embed more intelligence into instructional fitness apps, it becomes necessary for non-human actors to explain their actions as an input to the overall shared understanding of how the system operates.

God in the Machine. "While no member of the crew is taxed by these circumstances, the system as a whole may still be doing more cognitive work than could be done by any individual alone" [32, p. 7]. While the interaction between an individual and an app is not as complex and diverse as an aeroplane cockpit, the whole health ecosystem is still a complex one. To reflect on contemporary and future instructional fitness apps, we envisage many use scenarios. For example, currently, the app provides interaction between a coach and the user; the user and their fitness tribe; other health apps; other health

unrelated utilities such as the calendar; the shopping app; and many different sorts of wearables. For the future, that would involve a more complex setup that includes the gym equipment, the gym classes schedule, IoT devices such as the fridge, different health practitioners, and many more. In other words, as the whole ecosystem is becoming more complex and diverse, it becomes very crucial to design for integration and interoperability; two essential aspects in our conceptual framework.

6 LIMITATIONS

One limitation of analysing online user reviews is interpretation. The researcher has no means of clarifying ambiguous statements or digging deeper to have a more comprehensive understanding. Still, the fact that our sample is large can smooth out errors and capitalise on the wisdom of the crowd. Another limitation is the criteria of choosing the apps because the top 10 apps in the iOS Top App Charts did not cover the whole spectrum of instructional fitness apps. Additionally, the majority of reviews covered the top 4 apps. However, the sheer size of the sample and the fact that we focus on apps that prescribe exercises, workouts and simple and complex motor skills ensures that the framework is specific and not general. Another limitation is the reviewers' interpretation of features. Researching the apps' features and interviewing the developers might clarify such gaps and provide more insights. Still, the fact that the same features have different reviews helps to identify and eliminate such misinterpretations. Additionally, the first author has previous experience working as a smartphone apps developer and a personal trainer, which helped to understand the reported apps' features.

7 CONCLUSION

The designer or engineer understand the function of an airspeed indicator, but it is the user (pilot) account of how they used the device and how they expected it to function in their context is of true significance. This study contributes a framework that describes user expectations and experiences using instructional fitness apps. Such a framework can be utilised to guide future design approaches and methodologies. Capturing user needs is a significant step towards designing a successful intervention strategy using smartphone apps. Our conceptual framework links vertically downwards to users experiences and opinions of design instances and upward to the distributed cognition theory as a theory that views interaction with interactive digital artefacts as a flow of information in different directions and trajectories. Horizontally, it connects with design guidelines. Löwgren considers design guidelines as intermediate level knowledge [42] embedded in Norman's design principles; which in their own right are integrated vertically downward to many design instances and upward to theories from a wide variety of disciplines. Finally, our framework can serve as an example to not only guide the design of instructional fitness apps, but to also structure similar frameworks created for other kinds of apps, e.g. educational, entertainment, and productivity.

ACKNOWLEDGMENTS

Ahed Aladwan's work is supported by the Australian Government Research and Training Program scholarship. Eduardo Velloso is the recipient of an Australian Research Council Discovery Early Career Award (Project Number: DE180100315) funded by the Australian Government. We would also like to thank Jenny Waycott and Deepti Aggarwal for their feedback on early drafts of the paper.

REFERENCES

[1] Charles Abraham and Susan Michie. 2008. A Taxonomy of Behavior Change Techniques Used in Interventions. *Health Psychology* 27, 3 (2008), 379–387. https://doi.org/10.1037/0278-6133.27.3.379 arXiv:NIHMS150003

[2] Aino Ahtinen, Pertti Huuskonen, and Jonna Häkkilä. 2010. Let's All Get Up and Walk to the North Pole: Design and Evaluation of a Mobile Wellness Application. In *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries (NordiCHI '10)*. ACM, New York, NY, USA, 3–12. https://doi.org/10.1145/1868914.1868920

[3] Mohammad A. Al-Ramahi, Jun Liu, and Omar F. El-Gayar. 2017. Discovering Design Principles for Health Behavioral Change Support Systems: A Text Mining Approach. *ACM Transactions on Management Information Systems* 8, 2-3, Article 5 (June 2017), 24 pages. https://doi.org/10.1145/3055534

[4] Pia A. Albinsson, B. Yasanthi Perera, and G. David Shows. 2017. Pursuing fitness: how dialectic goal striving and intersubjectivity influence consumer outcomes. *Consumption Markets & Culture* 20, 1 (2017), 35–58. https://doi.org/10.1080/10253866.2016.1172213

[5] Neil Anderson, Peter Herriot, and Gerard P. Hodgkinson. 2001. The practitioner-researcher divide in industrial, work and organizational (IWO) psychology: Where are we now, and where do we go from here? *Journal of Occupational and Organizational Psychology* 74, 4 (nov 2001), 391–411. https://doi.org/10.1348/096317901167451

[6] Wolmet Barendregt, Tilde Bekker, Peter Börjesson, Eva Eriksson, Asimina Vasalou, and Olof Torgersson. 2018. Intermediate-level Knowledge in Child-computer Interaction. In *Proceedings of the 17th ACM Conference on Interaction Design and Children (IDC '18)*. ACM, New York, NY, USA, 699–704. https://doi.org/10.1145/3202185.3205865

[7] Benjamin B. Bederson, James D. Hollan, Ken Perlin, Jonathan Meyer, David Bacon, and George Furnas. 1996. Pad++: A Zoomable Graphical Sketchpad For Exploring Alternate Interface Physics. *Journal of Visual Languages & Computing* 7, 1 (1996), 3 – 32. https://doi.org/10.1006/jvlc.1996.0002

[8] Nasser F. BinDhim, Alexandra Hawkey, and Lyndal Trevena. 2015. A Systematic Review of Quality Assessment Methods for Smartphone Health Apps. *Telemedicine and e-Health* 21, 2 (2015), 97–104. https://doi.org/10.1089/tmj.2014.0088 PMID: 25469795.

[9] Paulina Bondaronek, Ghadah Alkhaldi, April Slee, Fiona L Hamilton, and Elizabeth Murray. 2018. Quality of Publicly Available Physical Activity Apps: Review and Content Analysis. *JMIR Mhealth Uhealth* 6, 3 (21 Mar 2018), e53. https://doi.org/10.2196/mhealth.9069

[10] Frank W. Booth, Christian K. Roberts, and Matthew J. Laye. 2012. *Lack of Exercise Is a Major Cause of Chronic Diseases.* American Cancer Society, 1143–1211. https://doi.org/10.1002/cphy.c110025

[11] Judit Bort-Roig, Nicholas D. Gilson, Anna Puig-Ribera, Ruth S. Contreras, and Stewart G. Trost. 2014. Measuring and Influencing Physical Activity with Smartphone Technology: A Systematic Review. *Sports Medicine* 44, 5 (01 May 2014), 671–686. https://doi.org/10.1007/s40279-014-0142-5

[12] Elizabeth Buie, Clare J. Hooper, and Aaron Houssian. 2013. Research-practice Interaction: Building Bridges, Closing the Gap. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems (CHI EA '13)*. ACM, New York, NY, USA, 2493–2496. https://doi.org/10.1145/2468356.2468813

[13] Fabio Buttussi and Luca Chittaro. 2008. MOPET: A context-aware and user-adaptive wearable system for fitness training. *Artificial Intelligence in Medicine* 42, 2 (2008), 153 – 163. https://doi.org/10.1016/j.artmed.2007.11.004

[14] John M. Carroll and Mary Beth Rosson. 1992. Getting Around the Task-artifact Cycle: How to Make Claims and Design by Scenario. *ACM Transactions on Information Systems* 10, 2 (April 1992), 181–212. https://doi.org/10.1145/146802.146834

[15] Lucas Colusso, Cynthia L. Bennett, Gary Hsieh, and Sean A. Munson. 2017. Translational Resources: Reducing the Gap Between Academic Research and HCI Practice. In *Proceedings of the 2017 Conference on Designing Interactive Systems (DIS '17)*. ACM, New York, NY, USA, 957–968. https://doi.org/10.1145/3064663.3064667

[16] David E. Conroy, Chih-Hsiang Yang, and Jaclyn P. Maher. 2014. Behavior Change Techniques in Top-Ranked Mobile Apps for Physical Activity. *American Journal of Preventive Medicine* 46, 6 (2014), 649 – 652. https://doi.org/10.1016/j.amepre.2014.01.010

[17] Sunny Consolvo, David W. McDonald, Tammy Toscos, Mike Y. Chen, Jon Froehlich, Beverly Harrison, Predrag Klasnja, Anthony LaMarca, Louis LeGrand, Ryan Libby, Ian Smith, and James A. Landay. 2008. Activity Sensing in the Wild: A Field Trial of Ubifit Garden. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '08)*. ACM, New York, NY, USA, 1797–1806. https://doi.org/10.1145/1357054.1357335

[18] Logan T. Cowan, Sarah A. Van Wagenen, Brittany A. Brown, Riley J. Hedin, Yukiko Seino-Stephan, P. Cougar Hall, and Joshua H. West. 2013. Apps of Steel: Are Exercise Apps Providing Consumers With Realistic Expectations?: A Content Analysis of Exercise Apps for Presence of Behavior Change Theory. *Health Education & Behavior* 40, 2 (2013), 133–139. https://doi.org/10.1177/1090198112452126 PMID: 22991048.

[19] Peter Dalsgaard and Christian Dindler. 2014. Between Theory and Practice: Bridging Concepts in HCI Research. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 1635–1644. https://doi.org/10.1145/2556288.2557342

[20] Artur Direito, Leila Pfaeffli Dale, Emma Shields, Rosie Dobson, Robyn Whittaker, and Ralph Maddison. 2014. Do physical activity and dietary smartphone applications incorporate evidence-based behaviour change techniques? *BMC Public Health* 14, 1 (dec 2014), 646. https://doi.org/10.1186/1471-2458-14-646

[21] William Gaver. 2012. What Should We Expect from Research Through Design?. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. ACM, New York, NY, USA, 937–946. https://doi.org/10.1145/2207676.2208538

[22] Sabine Geldof and Joannes Vandermeulen. 2007. a Practitioner's View of Human-Computer Interaction Research and Practice. *Artifact: Journal of Visual Design* 1, 3 (dec 2007), 134–141. https://doi.org/10.1080/17493460701800181

[23] Colin M. Gray, Erik Stolterman, and Martin A. Siegel. 2014. Reprioritizing the Relationship Between HCI Research and Practice: Bubble-up and Trickle-down Effects. In *Proceedings of the 2014 Conference on Designing Interactive Systems (DIS '14)*. ACM, New York, NY, USA, 725–734. https://doi.org/10.1145/2598510.2598595

[24] Regina Guthold, Gretchen A Stevens, Leanne M Riley, and Fiona C Bull. 2018. Worldwide trends in insufficient physical activity from 2001 to 2016: a pooled analysis of 358 population-based surveys with 1.9 million participants. *The Lancet Global Health* 6, 10 (2018), e1077 – e1086. https://doi.org/10.1016/S2214-109X(18)30357-7

[25] Hayati Havlucu, Terry Eskenazi, Bariş Akgün, Mehmet Cengiz Onbaşlı, Aykut Coşkun, and Oğuzhan Özcan. 2018. Flow State Feedback Through Sports Wearables: A Case Study on Tennis. In *Proceedings of the 2018 Designing Interactive Systems Conference (DIS '18)*. ACM, New York, NY, USA, 1025–1039. https://doi.org/10.1145/3196709.3196807

[26] Brian Hazlehurst, Paul N. Gorman, and Carmit K. McMullen. 2008. Distributed cognition: An alternative model of cognition for medical informatics. *International Journal of Medical Informatics* 77, 4 (2008), 226 – 234. https://doi.org/10.1016/j.ijmedinf.2007.04.008

[27] Eric B. Hekler, Predrag Klasnja, Jon E. Froehlich, and Matthew P. Buman. 2013. Mind the Theoretical Gap: Interpreting, Using, and Developing Behavioral Theory in HCI Research. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 3307–3316. https://doi.org/10.1145/2470654.2466452

[28] James Hollan, Edwin Hutchins, and David Kirsh. 2000. Distributed Cognition: Toward a New Foundation for Human-computer Interaction Research. *ACM Transactions on Computer-Human Interaction* 7, 2 (June 2000), 174–196. https://doi.org/10.1145/353485.353487

[29] Wilhelmina E Hoogendoorn, Mireille NM van Poppel, Paulien M Bongers, Bart W Koes, and Lex M Bouter. 1999. Physical load during work and leisure time as risk factors for back pain. *Scandinavian Journal of Work, Environment & Health* 25, 5 (1999), 387–403. https://doi.org/10.5271/sjweh.451

[30] Kristina Höök and Jonas Löwgren. 2012. Strong Concepts: Intermediate-level Knowledge in Interaction Design Research. *ACM Transactions on Computer-Human Interaction* 19, 3, Article 23 (Oct. 2012), 18 pages. https://doi.org/10.1145/2362364.2362371

[31] Edwin Hutchins. 1995. *Cognition in the Wild*. MIT Press. https://books.google.com.au/books?id=CGIaNc3F1MgC

[32] Edwin Hutchins and Tove Klausen. 1996. *Distributed cognition in an airline cockpit*. Cambridge University Press, 15–34. https://doi.org/10.1017/CBO9781139174077.002

[33] Luke Hutton, Blaine A Price, Ryan Kelly, Ciaran McCormick, Arosha K Bandara, Tally Hatzakis, Maureen Meadows, and Bashar Nuseibeh. 2018. Assessing the Privacy of mHealth Apps for Self-Tracking: Heuristic Evaluation Approach. *JMIR Mhealth Uhealth* 6, 10 (22 Oct 2018), e185. https://doi.org/10.2196/mhealth.9217

[34] Zachary Y. Kerr, Christy L. Collins, and R. Dawn Comstock. 2010. Epidemiology of Weight Training-Related Injuries Presenting to United States Emergency Departments, 1990 to 2007. *The American Journal of Sports Medicine* 38, 4 (2010), 765–771. https://doi.org/10.1177/0363546509351560 PMID: 20139328.

[35] Jesper Kjeldskov and Mikael B. Skov. 2014. Was It Worth the Hassle?: Ten Years of Mobile HCI Research Discussions on Lab and Field Evaluations. In *Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices & Services (MobileHCI '14)*. ACM, New York, NY, USA, 43–52. https://doi.org/10.1145/2628363.2628398

[36] Predrag Klasnja and Wanda Pratt. 2012. Healthcare in the pocket: Mapping the space of mobile-phone health interventions. *Journal of Biomedical Informatics* 45, 1 (2012), 184 – 198. https://doi.org/10.1016/j.jbi.2011.08.017

[37] Ruud Knols, Neil K. Aaronson, Daniel Uebelhart, Jaap Fransen, and Geert Aufdemkampe. 2005. Physical Exercise in Cancer Patients During and After Medical Treatment: A Systematic Review of Randomized and Controlled Clinical Trials. *Journal of Clinical Oncology* 23, 16 (2005), 3830–3842. https://doi.org/10.1200/JCO.2005.02.148 PMID: 15923576.

[38] Shanshan Li, James Caverlee, Wei Niu, and Parisa Kaghazgaran. 2017. Crowdsourced App Review Manipulation. In *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '17)*. ACM, New York, NY, USA, 1137–1140. https://doi.org/10.1145/3077136.3080741

[39] Ting-Peng Liang, Xin Li, Chin-Tsung Yang, and Mengyue Wang. 2015. What in Consumer Reviews Affects the Sales of Mobile Apps: A Multifacet Sentiment Analysis Approach. *International Journal of Electronic Commerce* 20, 2 (2015), 236–260. https://doi.org/10.1080/10864415.2016.1087823

[40] W. Lidwell, K. Holden, J. Butler, and K. Elam. 2010. *Universal Principles of Design, Revised and Updated: 125 Ways to Enhance Usability, Influence Perception, Increase Appeal, Make Better Design Decisions, and Teach Through Design*. Rockport Publishers. https://books.google.com.au/books?id=3RFyaF7jCZsC

[41] Gill Livingston, Andrew Sommerlad, Vasiliki Orgeta, Sergi G Costafreda, Jonathan Huntley, David Ames, Clive Ballard, Sube Banerjee, Alistair Burns, Jiska Cohen-Mansfield, Claudia Cooper, Nick Fox, Laura N Gitlin, Robert Howard, Helen C Kales, Eric B Larson, Karen Ritchie, Kenneth Rockwood, Elizabeth L Sampson, Quincy Samus, Lon S Schneider, Geir SelbÃęk, Linda Teri, and Naaheed Mukadam. 2017. Dementia prevention, intervention, and care. *The Lancet* 390, 10113 (2017), 2673 – 2734. https://doi.org/10.1016/S0140-6736(17)31363-6

[42] Jonas Löwgren. 2013. Annotated Portfolios and Other Forms of Intermediate-level Knowledge. *Interactions* 20, 1 (Jan. 2013), 30–34. https://doi.org/10.1145/2405716.2405725

[43] Ewa Luger and Abigail Sellen. 2016. "Like Having a Really Bad PA": The Gulf Between User Expectation and Experience of Conversational Agents. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 5286–5297. https://doi.org/10.1145/2858036.2858288

[44] Walid Maalej, Zijad Kurtanović, Hadeer Nabil, and Christoph Stanik. 2016. On the automatic classification of app reviews. *Requirements Engineering* 21, 3 (01 Sep 2016), 311–331. https://doi.org/10.1007/s00766-016-0251-9

[45] D. Martens and T. Johann. 2017. On the Emotion of Users in App Reviews. In *Proceedings of the 2017 IEEE/ACM 2nd International Workshop on Emotion Awareness in Software Engineering (SEmotion)*. 8–14. https://doi.org/10.1109/SEmotion.2017.6

[46] Joseph McKnight and Gavin Doherty. 2008. Distributed Cognition and Mobile Healthcare Work. In *Proceedings of the 22Nd British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction - Volume 2 (BCS-HCI '08)*. BCS Learning & Development Ltd., Swindon, UK, 35–38. http://dl.acm.org/citation.cfm?id=1531826.1531836

[47] Anouk Middelweerd, Danielle M. van der Laan, Maartje M. van Stralen, Julia S. Mollee, Mirjam Stuij, Saskia J. te Velde, and Johannes Brug. 2015. What features do Dutch university students prefer in a smartphone application for promotion of physical activity? A qualitative approach. *International Journal of Behavioral Nutrition and Physical Activity* 12, 1 (01 Mar 2015), 31. https://doi.org/10.1186/s12966-015-0189-1

[48] Mihaly Csikszentmihalyi. 1997. *Finding Flow: The Psychology of Engagement With Everyday Life*. Basic Books, New York, NY, USA.

[49] J.S. Mollee, A. Middelweerd, R.L. Kurvers, and M.C.A. Klein. 2017. What technological features are used in smartphone apps that promote physical activity? A review and content analysis. *Personal and Ubiquitous Computing* 21, 4 (01 Aug 2017), 633–643. https://doi.org/10.1007/s00779-017-1023-3

[50] Michael Muller. 2014. Curiosity, creativity, and surprise as analytic tools: Grounded theory method. In *Ways of Knowing in HCI*. 24–48. https://doi.org/10.1007/978-1-4939-0378-8_2

[51] Jakob Nielsen and Rolf Molich. 1990. Heuristic Evaluation of User Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '90)*. ACM, New York, NY, USA, 249–256. https://doi.org/10.1145/97243.97281

[52] Donald A. Norman. 2010. The Research-practice Gap: The Need for Translational Developers. *Interactions* 17, 4 (jul 2010), 9–12. https://doi.org/10.1145/1806491.1806494

[53] Donald A. Norman. 2013. *The design of everyday things: Revised and expanded edition*. Basic Books.

[54] B. K. Pedersen and B. Saltin. 2015. Exercise as medicine - evidence for prescribing exercise as therapy in 26 different chronic diseases. *Scandinavian Journal of Medicine & Science in Sports* 25, S3 (2015), 1–72. https://doi.org/10.1111/sms.12581

[55] PocketGamer. 2018. App Store Metrics | Pocket Gamer.biz | PGbiz. http://www.pocketgamer.biz/metrics/app-store/categories/

[56] Matthew Price, Erica K. Yuen, Elizabeth M. Goetter, James D. Herbert, Evan M. Forman, Ron Acierno, and Kenneth J. Ruggiero. 2014. mHealth: A Mechanism to Deliver More Accessible, More Effective Mental Health Care. *Clinical Psychology & Psychotherapy* 21, 5 (2014), 427–436. https://doi.org/10.1002/cpp.1855

[57] Carolyn Rabin and Beth Bock. 2011. Desired Features of Smartphone Applications Promoting Physical Activity. *Telemedicine and e-Health* 17, 10 (2011), 801–803. https://doi.org/10.1089/tmj.2011.0055 PMID: 22010977.

[58] Atish Rajkomar and Ann Blandford. 2011. Distributed Cognition for Evaluating Healthcare Technology. In *Proceedings of the 25th BCS Conference on Human-Computer Interaction (BCS-HCI '11)*. British Computer Society, Swinton, UK, UK, 341–350. http://dl.acm.org/citation.cfm?id=2305316.2305375

[59] Christian Remy, Silke Gegenbauer, and Elaine M. Huang. 2015. Bridging the Theory-Practice Gap: Lessons and Challenges of Applying the Attachment Framework for Sustainable HCI Design. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 1305–1314. https://doi.org/10.1145/2702123.2702567

[60] Melissa J. Rogerson, Martin R. Gibbs, and Wally Smith. 2018. Cooperating to Compete: The Mutuality of Cooperation and Competition in Boardgame Play. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 193, 13 pages. https://doi.org/10.1145/3173574.3173767

[61] Corina Sas, Steve Whittaker, Steven Dow, Jodi Forlizzi, and John Zimmerman. 2014. Generating Implications for Design Through Design Research. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 1971–1980. https://doi.org/10.1145/2556288.2557357

[62] Felipe B. Schuch, Davy Vancampfort, Justin Richards, Simon Rosenbaum, Philip B. Ward, and Brendon Stubbs. 2016. Exercise as a treatment for depression: A meta-analysis adjusting for publication bias. *Journal of Psychiatric Research* 77 (2016), 42 – 51. https://doi.org/10.1016/j.jpsychires.2016.02.023

[63] Seventy-first World Health Assembly. 2018. WHO global action plan on physical activity 2018-2030. http://apps.who.int/iris/bitstream/handle/10665/272722/9789241514187-eng.pdf

[64] Jakob Tholander and Stina Nylander. 2015. Snot, Sweat, Pain, Mud, and Snow: Performance and Experience in the Use of Sports Watches. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 2913–2922. https://doi.org/10.1145/2702123.2702482

[65] Alex Walz. 2015. Deconstructing the app store rankings formula with a little mad science. Retrieved Nov 2, 2017 from https://moz.com/blog/app-store-rankings-formula-deconstructed-in-5-mad-science-experiments

[66] Joshua H West, P. Cougar Hall, Carl L Hanson, Michael D Barnes, Christophe Giraud-Carrier, and James Barrett. 2012. There's an App for That: Content Analysis of Paid Health and Fitness Apps. *Journal of Medical Internet Research* 14, 3 (14 May 2012), e72. https://doi.org/10.2196/jmir.1977

[67] Elisabeth Zschucke, Katharina Gaudlitz, and Andreas Ströhle. 2013. Exercise and Physical Activity in Mental Disorders: Clinical and Experimental Evidence. *Journal of Preventive Medicine & Public Health* 46, Suppl 1 (2013), S12–S21. https://doi.org/10.3961/jpmph.2013.46.S.S12

University Library

A gateway to Melbourne's research publications

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Aladwan, A; Kelly, R; Baker, S; Velloso, E

Title:
A Tale of Two Perspectives: A Conceptual Framework of User Expectations and Experiences of Instructional Fitness Apps

Date:
2019

Citation:
Aladwan, A., Kelly, R., Baker, S. & Velloso, E. (2019). A Tale of Two Perspectives: A Conceptual Framework of User Expectations and Experiences of Instructional Fitness Apps. CHI 2019: PROCEEDINGS OF THE 2019 CHI CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS, Association for Computing Machinery. https://doi.org/10.1145/3290605.3300624.

Persistent Link:
http://hdl.handle.net/11343/227786

File Description:
Accepted version
License:
Unknown