

FinMARS: A Mobile App Rating Scale for Finance Apps

Johannes Huebner¹, Carlo Schmid², Mehdi Bouguerra¹, Alexander Ilic¹

¹ETH Zurich, Zurich, Switzerland

²University of St. Gallen, St. Gallen, Switzerland

jhuebner,bmehdi,ailic@ethz.ch, carlo.schmid@student.unisg.ch

ABSTRACT

Mobile apps increasingly replace face-to-face interactions between financial service providers and their customers. Therefore, it is critical for developers of finance apps to understand users' perception thereof, and to be able to assess the quality of their own app and their competitors' apps. Star ratings as provided by mobile app stores suffer from multiple shortcomings and are not detailed enough to fulfil this purpose. In this work, we thus developed a reliable, objective, multidimensional measure of the quality of mobile finance apps, which includes both generic and domain-specific aspects. We used an iterative approach and expanded on related work in the Health domain, and validated the scale empirically. The resulting app rating scale for finance apps is a reliable, objective measure of app quality, comprised of six subscales and a total of 34 items. It exhibits excellent internal consistency ($\alpha=.93$) and very good interrater reliability ($ICC=.74$).

CCS Concepts

• **Software and its engineering**→ **Software creation and management** • **Information systems**→ **Information systems applications** • **Human-centered computing**→ **Human computer interaction (HCI)**

Keywords

Mobile app quality; app rating scale development; finance apps; FinTech

1. INTRODUCTION

Smartphones and mobile apps have changed many aspects our lives dramatically and will likely continue to do so in the near future. Estimates for 2019 place the number of smartphone users at 2.5 billion worldwide [1]. The banking and finance industry is one example of how much consumer-facing processes have already changed in the last decades, and of how they continue to evolve. This manifests not only in the annual 4.1 billion downloads for finance apps on the Apple app store alone (2018 estimate [1]), but also becomes evident in what has been dubbed the “FinTech revolution”, i.e. the emergence of a wave of new market entrants seeking to challenge every part of the financial value chain. Even the underlying infrastructure might be shaken

up in the years to come, a development particularly driven by crypto-currencies and distributed ledger technologies [2]. Perhaps most strikingly, consumers' interactions with financial services providers increasingly happen through digital-only channels and especially mobile apps, instead of face-to-face in a bank branch. Because of this, it is crucial for financial service providers to understand what users are looking for, and to be able to assess the quality of their own app and their competitors' apps.

Yet, no standardized, objective and efficient way exists to assess the quality of finance apps. While the overall star ratings of mobile apps fulfils an important signaling function for new users, they suffer from several shortcomings, as pointed out by previous research, such as [3]–[6]:

- They are biased towards extreme ratings
- The textual content of a review and the star rating sometimes mismatch, and the five-star scale is subject to individual interpretation
- Reviews of previous app versions might not be valid for the present version anymore
- Individual bugs or issues may be addressed in many reviews, thus further biasing the overall rating
- The quantity and sentiment of reviews may be heavily influenced by how users are prompted to leave reviews when using the app
- Developers may even simply buy beneficial reviews
- Competitors may systematically try to leave negative reviews for an app

In addition, it is questionable whether following the feedback of the most vocal reviewers would invariably lead to an improved app quality (and other metrics like revenues). It is further debatable whether such reviews are informative enough to help developers infer in which areas their app might need improving, or why exactly their competitors receive better or worse reviews. Nonetheless, star ratings – in spite of their crudeness – provide at least some indication of the quality of an app, and are arguably the best widely available measure for gauging users' opinion of any given app, which is why this study also incorporated them in a validation step.

We thus systematically developed and empirically validated a domain-specific mobile app rating scale for finance apps geared towards researchers and app development practitioners (FinMARS). While developing this scale, the intrinsic nature of the finance industry was considered: strict privacy requirements, highly sensitive information, heavy regulations, and a large variety of personal circumstances that affect how individuals manage their finances – all of the above criteria need to be considered when building a high-quality mobile finance app, and they might be less crucial for apps in other categories, say a music streaming service or a productivity app. The resulting scale is designed to help researchers, market analysts, and (prospective) app developers objectively assess the quality of a set of finance app(s) of interest.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org

ICICM 2019, August 23–26, 2019, Prague, Czech Republic

© 2019 Association for Computing Machinery.

ACM ISBN 978-1-4503-7188-9/19/08...\$15.00

<https://doi.org/10.1145/3357419.3357428>

2. Related Work

2.1 User-Generated Reviews of Mobile Apps

User reviews can be a valuable data source for developers looking to improve their existing mobile apps. Prior work has proposed multiple approaches to help developers make sense of their own app's reviews in an automated way, e.g. by classifying reviews into bug reports or feature requests [7], by identifying the most "informative" reviews [8], or even identifying code passages responsible for a particular issue [9]. Sentiment analysis has been implemented to enrich such approaches [3]. Moreover, reviews not only serve as a feedback mechanism to app developers, but they also fulfill a signaling function for prospective, new users, which can affect app sales [10]. Of course, the most salient signal for app quality is the average star rating displayed right next to an app's name in all major app stores.

However, as outlined above, star ratings are not free of biases, and many apps receive only few reviews. Prior research has further found that some aspects like an app's user interface (i) are mentioned more frequently in reviews than other aspects (e.g. privacy), and (ii) they also have a larger impact on star ratings [3]. Therefore, average star ratings should be taken with a grain of salt, and are not necessarily a perfect measure of app quality.

2.2 Mobile App Rating Scales

For comparing apps in a more detailed, objective way, a quality assessment tool is thus needed. Stoyanov et. al created a Mobile App Rating Scale (MARS) for mobile health apps with the goal to develop a reliable multidimensional scale capturing all major aspects that determine the quality of apps in this category [11]. As they outlined in their paper, prior work often focused on technical features of applications and did not include the quality of those features. Other attempts to create evaluation frameworks for apps were found to be either too generic, or complex. The resulting scale has been proven to be a reliable tool in assessing health apps' quality, and it has been used by numerous researchers since its conception. Stoyanov et. al also developed a simplified version geared towards end users (uMARS, the user version of MARS), which uses simpler language in the item and response descriptions [12]. While the app rating scales presented by Stoyanov et. al have been proven helpful in rating health apps, it is not directly applicable to other app categories, since some of the MARS criteria apply to a much lesser extent to non-health apps. Additionally, other aspects might be more crucial for finance apps instead, such as the trustworthiness of the app's developer, privacy, or the possibility for users to conduct certain use cases on different platforms (e.g. for retirement planning, a desktop format might be better-suited than a mobile app).

Similar to health apps, though, finance apps can be categorized in many subcategories, which poses a challenge in developing an overarching scale applicable to all subtypes. While in some cases it might be beneficial to focus on an even narrower subcategory and include even more fine-granular evaluation criteria, we follow the proven approach of Stoyanov et. al and focus on criteria that apply to all finance apps [11], [12].

2.3 Mobile Finance Apps

The use of IT in the finance industry, as well as research on this topic, is hardly a recent phenomenon. Core banking systems, credit cards, interbank payment systems, self-service systems like ATMs as well as online banking are all examples of how financial service providers have heavily relied on the use of IT [13]. Yet, the term "FinTech" (short for Financial Technology), has only

recently emerged as part of the so-called "FinTech revolution", which has not simply yielded yet another customer channel (mobile apps, see [14] for a summary of prior research), but also changed the financial industry more profoundly. For example, new business models have emerged (e.g. crowd-funding, peer-to-peer lending), machine learning techniques are used more intensely (e.g. for robo-advisors), and perhaps most importantly, the customer interface of financial service providers is transforming. Instead of face-to-face interactions in physical bank branches, customers increasingly revert to mobile apps. While similar effects of the digital transformation can certainly be observed in other industries, this change is particularly dramatic for financial service providers, whose main differentiators from competitors often are their superior customer service and their ability to provide trustworthy financial advice when needed. Maintaining these differentiating qualities with digital-only channels is a great challenge that requires a very careful design of the corresponding mobile finance apps, which further substantiates the need for an objective quality assessment scale. While such scales exist for health apps [11], [12], to the best of our knowledge, no comparable measures are available for finance apps.

3. Methodology

3.1 Scale Development

This work followed the proven scale development approach proposed by DeVellis [15]. After determining what should be measured, the next steps were to create an item pool, determine the format of measurement, and a structured review and validation of the items (e.g. expert review, testing on real-world observations). This work drew on the MARS and uMARS for health apps for an initial set of items, from which suitable criteria were adopted (subscales engagement, functionality, and aesthetics), while health-specific criteria that could not be meaningfully applied to finance apps were dismissed. Then, the scale was applied to a random set of 10 finance apps (see next section for details), and adjusted. At the same time, additional finance-specific criteria were developed based on empirical observations. This process was repeated until consent between the authors was reached, i.e. no more items or measurements were added, modified or removed in an iteration. In the following two iterations, we both induced subscales with category-specific quality criteria from ten real-world samples per iteration, and also deduced subscales from the technology adoption model (TAM, [16]) and its successor, the unified theory of acceptance and use of technology (UTAUT, [17]). In total, this led to the addition of three new subscales: (i) an app's *trust signaling*, which indicates the quality of signaling trust towards potential users; (ii) *app value*, which evaluates the quality and value of the actual financial services provided, and (iii) *financial behavior*, which assesses the app's ability to help users change their financial behavior positively. Each of the new subscales consists of six to ten items (see Table 1 below), and each item is measured on a 5-point Likert scale. The average of the items within each subscale determine a subscale score, e.g. the mean score of items 1 through 5 constitute the *engagement* score. The mean of the six subscale scores then in turn constitute the overall FinMARS score of an app.

Table 1. Interrater reliability and internal consistency of the FinMARS subscales and items, corrected item-total correlations and descriptives based on independent ratings of 50 mobile finance apps

#	Subscale / item	Corrected item-total correlation	Mean	SD
Engagement, alpha=.78, ICC=0.67 (95% CI 0.58-0.75)				
1	Entertainment	.64	2.98	1.13
2	Interest	.74	3.21	0.69
3	Customization	.49	3.40	0.97
4	Interactivity	.73	3.17	0.96
5	Target group	.60	4.11	0.81
Functionality, alpha=.62, ICC=0.66 (95% CI 0.59-0.72)				
6	Performance	.70	3.77	1.13
7	Ease of use	.32	3.21	1.02
8	Navigation	.31	3.47	0.88
9	Gestural design	.74	3.94	0.79
Aesthetics, alpha=.92, ICC=0.73 (95% CI 0.63-0.81)				
10	Layout	.75	3.70	0.95
11	Graphics	.76	3.72	1.16
12	Visual appeal	.78	3.70	1.18
Trust Signaling, alpha=.82, ICC=0.55 (95% CI 0.46-0.63)				
13	Accuracy of app description	.54	3.38	0.95
14	Brand awareness	.35	1.51	1.23
15	Continuous update frequency	.66	2.68	1.34
16	Customer service attraction	.62	2.57	1.02
17	Introduction / tutorial	.62	2.43	1.33
18	Perceived security signal	.67	3.49	0.93
19	Device / data permission trust signaling	.59	3.34	0.87
20	App on-boarding / first-time usage	.52	3.28	0.95
21	Privacy signaling	.65	2.96	1.02
22	Monetization transparency	.42	2.70	0.95
App Value, alpha=.66, ICC=0.50 (95% CI 0.35-0.62)				
23	Value proposition	.72	3.04	1.08
24	Cross-platform compatibility	.54	3.02	1.05
25	Notifications	.49	3.04	1.04
26	Financial overview	.82	3.30	1.10
27	Integration of other services	.27	3.15	1.01
28	Export functions	.30	2.44	1.50
Financial Behavior, alpha=.68, ICC=0.62 (95% CI 0.51-0.70)				
29	Financial literacy requirement	.04	3.34	1.40
30	Usage effort	.41	2.91	1.32
31	Financial feedback	.82	3.26	1.18
32	Active learning inputs	.72	2.30	1.38
33	Financial awareness	.53	2.96	1.14
34	Perceived impact on financial behavior	.50	2.62	0.99

3.2 Selection of Mobile Finance Apps and Acquisition of User-Generated Review Data

Data on a large set of real-world mobile finance apps was acquired through the mobile app analysis company 42matters in August 2018 using the following inclusion criteria: The app must be (i) available in English, (ii) available in the Swiss iOS app store for free, (iii) in the category “Finance”, (iv) it must have at least 50 ratings received, and (v) it must be usable without an existing banking relationship (e.g. bank account). The fourth criterion was included due to the necessity for robust average user ratings in our empirical validation. Since no structured information on the fifth criterion was available, we first made a selection based on criteria (i) through (iv), leading to 1’985 apps. Then, all apps were manually screened to double-check the accuracy of the first four criteria, and for the final criterion. Apps with descriptions in other languages than English, and statements such as “must be customer” or “enrolled in our banking service” were removed. This led to a dataset of 523 apps eligible for our study. From this set of apps, we draw random samples for each of the iterations during scale development, and a sample of 50 apps for a validation step (see next section).

For the empirical validation step, 42matters provided us with detailed review data (review text, date, star rating) for the selected apps. This review data had been crawled between Dec 2016 and Oct 2018.

3.3 Empirical Validation using App Store Star Ratings

To validate the scale, we first asked two expert raters of the research team with extensive experience in the financial services industry and with finance apps to independently rate the same 50 randomly selected finance apps, which is in line with the proven approach of [11]. In doing so, they were instructed to first use each app for 10 minutes and then assess its quality using the FinMARS. This process was conducted between October and November 2018 and yielded very good interrater reliability between the two raters (two-way mixed ICC=.74^{***}, 95% CI .67-.80), indicating that the FinMARS as developed in this work is a reliable measure of finance apps’ quality.

For the empirical validation, we compared FinMARS scores with mean star ratings of the same apps using correlation analyses. However, five of the 50 selected apps had a suspiciously high share of five-star reviews ($\geq 75\%$) even though they did not seem outstandingly high-quality or useful, and required a paid subscription to use them. This unusual distribution indicated that the app developers may have bought positive reviews, as also enumerated in the shortcomings of star ratings in the introduction of this paper. Seven other apps had not been updated in over a year, and predominantly had reviews related to previous app versions that were not necessarily accurate for the current app store version anymore. These 12 apps with biased app store reviews were thus excluded from the correlation analysis, leading to a set of 38 apps.

4. Results and Discussion

The resulting FinMARS (see Table 1) provides a reliable, multi-dimensional scale to objectively assess finance apps’ quality, comprised of 34 questions in the following subscales: (i) engagement, (ii) functionality, (iii) aesthetics, (iv) trust signaling, (v) app value, and (vi) financial behavior. The first three subscales focus on general criteria for mobile apps, while the remaining, newly-developed subscales are specifically geared towards mobile

finance apps. While additional subscales are conceivable, e.g. to capture intricacies of a particular subtype of finance apps, we deliberately focused on the most crucial ones that apply to all finance apps and help developers ensure a high level of quality in a pragmatic fashion. The scale exhibited very good interrater reliability (ICC=.74^{***}) and internal consistency (Cronbach alpha=.93, N=50 apps). The subscales also exhibited high internal consistencies and interrater reliabilities (see Table 1).

A correlation analysis between 38 eligible apps’ FinMARS scores (see previous chapter) and their mean app store ratings revealed moderate positive correlations (Pearson’s $r=.49^{**}$, $p=.0019$). Correlation tests using rank correlation coefficients (Spearman’s Rho and Kendall’s Tau) yielded similar results. Discrepancies between FinMARS scores and app store ratings tended to be highest for apps that were frequently updated by the developers, meaning some of the older app store reviews contributing to the mean rating might not have been 100% accurate for the latest version anymore, which was rated using FinMARS. We thus also conducted correlation analyses using only app store reviews from the last few months (i.e. only last 3, 6, 9, 12 or 24 months), which partially led to stronger correlations (e.g. $r=.55^{**}$ for a cut-off at 9 months). Ideally, the decision about the maximum age of reviews taken into account should be made on a per-app basis such that for apps with mostly minor, cosmetic updates (even though these might happen frequently) reviews are considered valid for longer than for those apps that are changed more fundamentally.

As discussed in the introduction of this paper, even though they are subject to biases, app store ratings on an aggregated basis are still a good high-level indicator of general app quality and user satisfaction for most apps. This is why we conducted a correlation analysis as a plausibility check, expecting that the correlation between FinMARS scores and app store ratings should be positive, but not necessarily extremely strong. A perfect positive correlation would indicate that FinMARS also perfectly reproduces the biases of star ratings, which would be detrimental to the objectivity of the scale. Still, a moderately positive correlation indicates alignment with the general sentiment of users’ ratings, which is certainly desirable.

5. Conclusion and outlook

This work systematically developed and empirically validated a detailed yet simple, objective, and robust scale to assess the quality of mobile finance apps. In line with prior work (e.g. [11]), the resulting scale has been applied to 50 apps by two independent raters, yielding very high interrater reliability and internal consistency. The FinMARS scores additionally exhibited a strong positive correlation with mean app store ratings. FinMARS thus is a robust, objective measure of app quality that captures the users’ perception of a given finance app to a great degree, while providing substantially more details.

However, this work is not free of limitations. While FinMARS has been specifically adopted to meet category-specific aspects of finance apps, there still exists a variety of apps within this category (e.g. budget planning or trading apps). Future work should thus capture the intricacies of those subcategories better, possibly by assigning non-equal weights to the six subscales. In this work, we addressed this issue by focusing on aspects that are applicable to a wide range of finance apps. In addition, it may be a worthwhile endeavor to create a second version of FinMARS geared more towards end users that require less expertise, similar to what [12] have done. Another idea to make FinMARS scores more accessible to end users could be to develop a simple,

perhaps visual abstraction such as a traffic light indicating the FinMARS scores or subscale scores. However, we argue that the most immediate beneficiaries of such app rating scales are the professionals involved in mobile app development, which is the target group this work focuses on.

6. ACKNOWLEDGMENTS

We thank our project partner, 42matters AG, for the unique opportunity to conduct research with a large volume of app review data, and we gratefully acknowledge the grant from the Swiss CTI (26989.1 PFES-ES)

7. REFERENCES

- [1]. Statista, "Most popular Apple App Store categories in September 2018," 2018.
- [2]. J. McWaters, G. Bruno, R. Galaski, and S. Chatterjee, "The future of financial infrastructure - An ambitious look at how blockchain can reshape financial services," 2016.
- [3]. J. Huebner, R. M. Frey, C. Ammendola, E. Fleisch, and A. Ilic, "What People Like in Mobile Finance Apps: An Analysis of User Reviews," in *International Conference on Mobile and Ubiquitous Multimedia*, 2018.
- [4]. H. Khalid, E. Shihab, M. Nagappan, and A. E. Hassan, "What do mobile app users complain about?," *IEEE Softw.*, vol. 32, no. 3, 2015.
- [5]. I. M. Ruiz, M. Nagappan, B. Adams, T. Berger, S. Dienst, and A. Hassan, "An examination of the current rating system used in mobile app stores," *IEEE Softw.*, 2017.
- [6]. M. Kuehnhausen and V. S. Frost, "Trusting smartphone apps? To install or not to install, that is the question," in *2013 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support*, 2013.
- [7]. W. Maalej, Z. Kurtanović, H. Nabil, and C. Stanik, "On the automatic classification of app reviews," *Requir. Eng.*, vol. 21, no. 3, 2016.
- [8]. N. Chen, J. Lin, S. C. H. Hoi, X. Xiao, and B. Zhang, "AR-miner: mining informative reviews for developers from mobile app marketplace," in *Int'l Conf. on Software Engineering*, 2014.
- [9]. A. Ciurumelea, A. Schaufelbühl, S. Panichella, and H. C. Gall, "Analyzing reviews and code of mobile apps for better release planning," in *International Conference on Software Analysis, Evolution and Reengineering*, 2017.
- [10]. T.-P. Liang, X. Li, C.-T. Yang, and M. Wang, "What in consumer reviews affects the sales of mobile apps: A multifacet sentiment analysis approach," *Int. J. Electron. Commer.*, vol. 20, no. 2, 2015.
- [11]. S. R. Stoyanov, L. Hides, D. J. Kavanagh, O. Zelenko, D. Tjondronegoro, and M. Mani, "Mobile app rating scale: a new tool for assessing the quality of health mobile apps," *JMIR mHealth uHealth*, vol. 3, no. 1, 2015.
- [12]. S. R. Stoyanov, L. Hides, D. J. Kavanagh, and H. Wilson, "Development and Validation of the User Version of the Mobile Application Rating Scale (uMARS)," *JMIR mHealth uHealth*, vol. 4, no. 2, Jun. 2016.
- [13]. D. W. Arner, J. Barberis, and R. P. Buckley, "The Evolution of Fintech: A New Post-Crisis Paradigm?," pp. 1–45, 2015.

- [14]. A. A. Shaikh and H. Karjaluo, "Mobile banking adoption: A literature review," *Telemat. informatics*, vol. 32, no. 1, pp. 129–142, 2015.
- [15]. R. F. DeVellis, *Scale Development*, 4th ed. SAGE Publications, 2017.
- [16]. F. D. Davis, "Perceived Usefulness, Perceived Ease Of Use, And User Acceptance of Information Technology," *MIS Q.*, vol. 13, no. 3, pp. 319–340, 1989.
- [17]. V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis, "User acceptance of information technology: Toward a unified view," *MIS Q.*, vol. 27, no. 3, 2003.

8. Appendix I – FinMARS Items

8.1 Instructions

Raters should: (i) Turn off VPN/Ad-Blockers to ensure same rating standard, (ii) Use the app and trial it thoroughly for at least 10 minutes, (iii) Determine how easy it is to use, how well it functions and does it do what it purports to do, and (iv) Review app settings, developer information, external links, security features, dedicated website for mobile app etc.

When filling out the questionnaire, circle the number that most accurately represents the quality of the app you are rating. All items are rated on a 5-point scale from "1. Inadequate" to "5. Excellent". Select N/A if the app component is irrelevant. In the actual questionnaire, each point on the scale is annotated with item-specific examples.

8.2 Rating Scale

Section A, Engagement – *fun, interesting, customizable, interactive, has prompts (e.g. sends alerts, messages, reminders, feedback, enables sharing)*

1. Entertainment: Is the app fun/entertaining to use? Does it use any strategies to increase engagement through entertainment?
2. Interest: Is the app interesting to use? Does it use any strategies to increase engagement by presenting its content in an interesting way?
3. Customization: Does it provide/retain all necessary settings/preferences for app features (e.g. sound, content and notifications)?
4. Interactivity: Does it allow user input, provide feedback, contain prompts (reminders, sharing options, notifications, etc.)?
5. Target group: Is the app content (visuals, language, design) appropriate for the target audience? Is finance-related content suitable for potential users?

Section B, Functionality – *app functioning, easy to learn, navigation, flow logic, and gestural design of app*

6. Performance: How accurately/fast do the app features (functions) and components (buttons/menus) work?
7. Ease of use: How easy is it to learn how to use the app; how clear are the menu labels, icons and instructions?
8. Navigation: Does moving between screens make sense; Does app have all necessary links between screens?
9. Gestural design: Are interactions (taps/swipes/pinches/scrolls) intuitive? Are they consistent across all components/screens?

Section C, Aesthetics – *graphic design, overall visual appeal, color scheme, and stylistic consistency*

10. Layout: Is arrangement and size of buttons, icons, menus and content on the screen appropriate?
11. Graphics: How high is the quality/resolution of graphics used for buttons, icons, menus and content?
12. Visual appeal: How good does the app look?

Section D, Trust Signaling – *How is the app perceived before the user decides to commit himself or herself to use the app.*

13. Accuracy of app description (in app store): Does app contain what is described? Is the pricing outlined correctly?
14. Brand awareness: How well is the brand known (in the targeted countries)?
15. Continuous update frequency: How many times was the app updated within the last 12 months?
16. Customer service attraction: How attractive is the customer service of the app? Does it provide many options or is it almost impossible to contact customer service?
17. Tutorial: How much effort did the app developers put into the introduction to facilitate usage of the app?
18. Perceived security signaling: Does the app make an effort to signal security to the user?
19. Device / Data permission trust signaling: How does the app ask for permission / data from the user? Is the app signaling trust by giving reasonable explanation?
20. App on-boarding / First time usage: Does the app establish trust by letting the user preview the app? To which extent has the first-time user to give up personal data during the app on-boarding process compared to other similar apps?
21. Privacy signaling: Is the app trustworthy in regard to privacy?
22. Monetization transparency: How does the app make money? Is pricing for services / subscription visible / available?

Section E, App Value – *How much value is created for the user when using the app. Value in terms of convenience or real value added.*

23. Value proposition: Does the app fulfill it promises in terms of economic value and convenience?
24. Cross-Platform compatibility: Level of cross-platform compatibility?
25. Notifications: How well are the notifications in terms of relevance and quality?
26. Financial overview: Quality of presenting financial numbers in terms of clarity and intuitiveness?
27. Integration of other services: Does the app provide integration of other services (third party-apps, inside app, or sharing functions)? How useful and well implemented are these integrations?
28. Data export: Is there an export function for financial user data? Possible to adjust timeframes, choose columns? (Note: App-websites offer export functions)

Section F, Financial Behavior – *Aspects of the app related to financial behavior: possible barriers or facilitators.*

29. Financial literacy (FL) requirement: Which level of FL is required to understand concepts and services used in app?
30. Usage effort: How much effort is needed from the user to use the app continuously after set-up?
31. Financial feedback: Does the app provide the user with financial feedback over time?
32. Active learning inputs: Does the app try to increase financial literacy / understanding with active learning inputs such as newsletters, articles, blogs, explanations or other forms?
33. Financial awareness: Does the app increase financial awareness (e.g. understanding risks, saving opportunities)?
34. Perceived impact on financial behavior: By using the app will the user change his financial behavior? Is the perceived impact positive or negative? How likely is it that user will use the app in terms of usage factor (High / Low effort)?