

Examining Choice Overload across Single-list and Multi-list User Interfaces

Alain D. Starke^{1,2,*}, Justyna Sedkowska³, Mihir Chouhan³ and Bruce Ferwerda³

¹Marketing and Consumer Behaviour Group, Wageningen University & Research, Hollandseweg 1, 6706KN, Wageningen, Netherlands

²MediaFutures, University of Bergen, Lars Hilles gate 30, 5008, Bergen, Norway

³Department of Computer Science and Informatics, Jönköping University, Jönköping, Sweden

Abstract

Recommender systems are prone to triggering choice overload among users due to the typically large set sizes. Various applications have been developed that aim to overcome this through interface design, notably by so-called multi-list recommender systems. However, to what extent such user interface design actually reduces choice overload compared to single-list interfaces has yet to be examined. In a user study ($N = 150$), we compared three common user interfaces (UIs) in the context of recipe recommendation: a single-list UI, a grid UI and a multi-list UI. Whereas earlier studies found differences in choice difficulty and choice satisfaction across grid-based and multi-list recommender interfaces, we observed no such differences, as the explanations were possibly not sufficiently helpful. Instead, we found that grid-based UIs and multi-list UIs had a higher perceived ease of use than a single-list UI, which in turn reduced choice difficulty. The benefits of such interfaces, thus, may lie in the organization of the UI, at least in the recipe domain.

Keywords

Choice Overload, User Interface, User Experience, Recommender Systems, Food

1. Introduction


It is often assumed that larger choice sets are desirable. However, humans have a limited cognitive capacity, which can lead to difficulties in the decision-making process. As a result, larger choice sets can lead to dissatisfaction, regret or even choice deferral among decision-makers [1, 2]. This phenomenon is more commonly known as choice overload [3]. It occurs when the number of options within a choice set exceeds the amount that one's working memory can cope with. Limiting the number of options to reduce choice overload is not always a feasible nor a desirable solution. For example, it can create the impression that a platform has little to offer while, on the contrary, there is a continuous increase of content available that recommender systems need to deal with [4, 5]. The common rationale is that through the use of personalization (i.e., items that are relevant to individual users), choice overload can be

InTRS'22: Joint Workshop on Interfaces and Human Decision Making for Recommender Systems, September 22, 2022, Seattle, US (hybrid event).

*Corresponding author.

✉ alain.starke@wur.nl (A. D. Starke); seju17zd@student.ju.se (J. Sedkowska); chmi21xe@student.ju.se (M. Chouhan); Bruce.Ferwerda@ju.se (B. Ferwerda)

ORCID 0000-0002-9873-8016 (A. D. Starke); 0000-0003-4344-9986 (B. Ferwerda)

 © 2022 Copyright © 2022 for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

 CEUR Workshop Proceedings (CEUR-WS.org)

mitigated even though choice difficulty might still be relatively high [6]. However, more recent work has found that the extent to which users experience choice overload does not only depend on the number of options presented but is influenced by how items are presented in the user interface (UI) [5, 7, 8]. The main aim of this study is to investigate the impact of different UIs on choice overload, in the context of a recipe recommender system. We investigated how UIs can influence a user's evaluation, by differentiating between UIs that organize in lists, grids, or multi lists.

1.1. Problem

Iyengar and Lepper [1] showed that choice overload takes place in common, 'offline' decision-making environments with a large number of items. Many studies have examined choice overload in different context since then, most notably that choice overload occurs in online contexts as well [6, 9]. This is particularly a problem due to the abundance of choices that are available in online environments [9], which typically exceeds that of brick-and-mortar businesses.

Recommender systems aim to aid the decision-making process by mitigating choice overload [10, 11]. A typical way to support decision-making is through personalization by presenting content that is most relevant to users [12]. A common side effect of highly relevant items is that they are often similar and therefore hard to compare. Hence, although personalization is able to mitigate choice overload, it does not necessarily fully mitigate choice difficulty [6]. Although it is possible to diversify the recommended content and, as a result, reduce the experienced choice difficulty [13, 14], recent developments in recommender research have started to explore other methods to mitigate choice difficulties. A recent direction is to adjust the information architecture by re-organizing the content into different UIs [5, 15]. In doing so, the choice architecture of the decision making environment is changed [16], rather than the content.

This study considers three different popular UIs in terms of how items are presented to users: single-list interfaces, grids, and multi-list UIs (see Figure 1). Each interface is defined to have different characteristics. In a single-list UI, items are stacked on top of each other in a single column and can be explored by scrolling vertically. Such a design is commonly used to display search engine results [17]. Research on the single-list UI has shown that users tend to pay more attention to the items that are presented at the top of a list [17, 18, 19, 20]. In contrast, a grid UI consists of multiple rows of items with multiple items in each row [21]. Grids are particularly popular on e-commerce websites, because they are capable of providing a comparative overview of many similar items. This interface has been found to force users' to evaluate items across the different axes, in a more balanced way than one would for a single-list UI [20]. Furthermore, users of grid-like interfaces tend to examine more items than they would in other UIs [19].

The third type of UI examined in this study is the Multi-list UI. It has been adopted unan- imously by video streaming services [4], such as Netflix, Disney+ and HBO Max. Multi-list UIs are typically used in the context of recommender systems [5, 8, 15], stacking multiple lists of personalized algorithms on top of each other. Some studies refer to these 'lists within a multi-list UI' as carousels [22]. In a typical Multi-list UI, each list or carousel is accompanied by an explanation that describes what category all the items in the list below belong to [5, 8], or justify how the content is generated [22]. Additionally, further items can be discovered

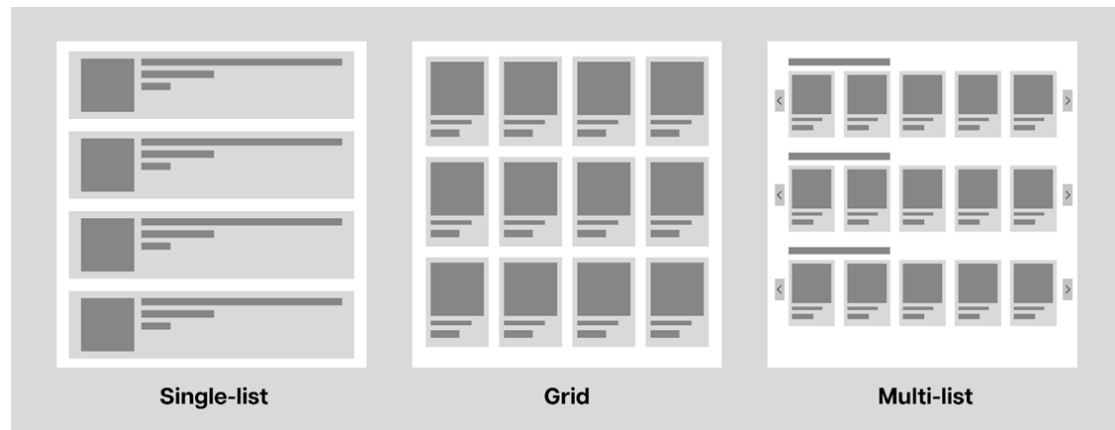


Figure 1: Visual representations of the different UIs in this study: single-list UI, grid UI and multi-list UI.

in each list through horizontal scrolling, which can include dozens of recommended items. Multi-list UIs typically allow the inclusion of items within different categories on one web page, combining constraint-based recommender approaches with content-based recommendation or collaborative filtering [5, 8]. Unlike for single-list and grid UIs, it is not clear yet whether items with a specific position in a list or carousel are more likely to be selected, while it seems likely that users are more likely to select *any* item from a list or carousel that appears higher up within the multi-list UI [4, 8].

Within the context of recommender systems, the definitions used for different types of UIs vary. Based on their own definition, Jannach et al. [5] compare a single-list UI to a multi-list UI, but what they define as a single-list UI is defined as a grid UI by Yener and Dundar [21]; the latter being consistent with the definition in this paper. In Jannach et al. [5], the single difference between the two UIs is that a multi-list UI has a label above each row of items explaining the genre of content for the row below, which is in line with the definitions used in Starke et al. [8]. In contrast, Kammerer and Gerjets [20], as well as Resnick et al. [23] define single-list UIs as a list in the way that it is also defined in the current paper. This differentiation is also consistent with other studies [21, 24].

Possibly also due to the ambiguity regarding definitions for different list UIs, it is currently unclear what kind of effect different UIs have on a user's decision-making and evaluation in the context of choice overload and recommender systems. Previous work has shown that various interfaces evoke different information processing behavior [20], and determine which items are more likely to receive users' attention [24]. That this would also apply to choice overload is likely, due to the preliminary findings in the context of multi-list recommender systems [5, 8]. However, a comparative study is still missing, also with regard to a user's perception and evaluation.

1.2. Research Questions

This study aims to determine if different UIs contribute to the occurrence of choice overload. Users are asked to choose a recipe they like and to evaluate their perception and experience of the system. We posit the following research questions:

[RQ1]: To what extent do users experience choice overload when interacting with a grid-based user interface and a multi-list interface, compared to a single-list UI?

We furthermore explore how the different UIs influence the perceived ease of use of the interfaces as it can further affect the evaluation of the recommender system as a whole [7, 25, 26]. The ease of use is referred to as to whether users can complete tasks quickly with ease and without frustration [27]. It is therefore of interest to examine how choice overload, perceived ease of use and the use of different UIs relate to each other:

[RQ2]: How does a user's perceived ease of use depend on the presented user interface?

2. Related work

2.1. Choice Overload

One of the first studies to examine choice overload was conducted by Iyengar and Lepper [1]. Their work involves three experiments, which show that while many choices may seem desirable, people are actually more likely to refrain from making a choice. In this study, this involved purchasing a product. Additionally, more options are found to result in higher difficulty to choose one option, and if a choice is made, post-purchase regret is more likely.

A meta-analysis on 50 studies on choice overload, conducted by Scheibehenne et al. [3], has established that choice overload is rather context-dependent. For example, people with preferences or expertise for certain products tend to prefer larger choice sets. They further describe that a lack of familiarity or preferences, which entails that people will not choose something they have prior knowledge of, is a precondition of choice overload to occur. Moreover, according to Scheibehenne et al. [3], if the options have "complementary or unique features that are not directly comparable", making a choice becomes more difficult. This might be worsened by the lack of a dominant item or an item that would clearly be preferred. The latter is also observed in the context of recommender systems, where highly attractive, similar options suggested to users tend to increase the choice difficulty [6, 28]. In an offline supermarket context, an increase in the number of products makes it harder to distinguish between items [29].

2.2. User Interfaces

A user interface is defined as the graphical representation of a system [30]. A lot of varieties are on offer to present items. The most basic one is a single-list UI, where items are stacked on top of each other and can be scrolled through vertically (See Figure 1). Such a design can be found on pages displaying search engine results (e.g., [17]), as well as on many e-commerce websites. A second common way to display items is through a grid. Grids usually consist of multiple rows with three to six items in each row (See Figure 1). Grids are especially popular on e-commerce websites, for they allow interface designers to show many items within a limited space, which would be optimized to desktop screens.

Recent research in recommender systems has introduced multi-list UIs as a third popular type (see Figure 1), particularly in movie streaming services [4]. The UI includes multiple algorithms, which typically optimize for different user models and/or constraints [4], and stack them on top of each other. In the context of academic research, multi-list recommender systems have also been used to promote healthier eating in recipe recommender systems [8], as well as to support movie decision-making [5, 22].

Among the most notable related work, Jannach et al. [5] have explored the decision-making behavior of users for a grid-like UI (defined by them as a single-list UI) and a multi-list movie recommender. They find that users tend to spend longer when interacting with the multi-list UI, resulting in a higher level of effort, while not affecting choice satisfaction. Starke et al. [8] follow a similar research design in the context of recipe recommendation, but also differentiate between smaller (5 items) and larger (25 items) set size, presented in either a grid or a multi-list UI. The multi-list recommender leads to a higher level of choice satisfaction, arguably due to the larger number of options to choose from, along with a higher diversity due to the use of multiple algorithms, compared to a grid). However, Starke et al. [8] also observe higher levels of choice difficulty when using a multi-list UI, which would be more in line with the ‘classical’ choice overload studies: people tend to prefer larger sets, but this comes at the cost of a higher level of choice difficulty [3].

2.3. UIs and User-centric evaluation of Recommender Systems

Over the past two decades, researchers have started to advocate for a more user-centric approach to the design of recommender system. Algorithmic accuracy is deemed to not be enough to optimize for [31, 32]. Instead, recommendation lists should consist of diverse options so that users may discover unique items. Diversity among items increased may increase perceived attractiveness [13, 32], while presenting many items that are too similar may trigger choice overload [6, 28].

Previous recommender studies have examined how effortful an interface is to use (cf. [33]). In this study, we consider perceived ease of use, which examines a user’s ability to complete tasks without feeling frustrated [27]. A recommender’s UI design may affect the effort that a user perceives when using the system (cf. [33, 34]). Several studies have proposed design guidelines for UI to create more effective recommender systems [26, 35], advocating that poorly designed UIs may discourage users from making purchases in e-commerce, akin to choice deferral in choice overload contexts [36]. Additionally, studies have also shown that a system’s UI can be an influential element in terms of how a user evaluates the recommender system’s quality [7, 25, 26]. Nonetheless, research on how recommendations are compiled into UI lists and how this would affect a user’s perception has received little attention in recommender research, even though multiple studies have argued that a good combination of UIs and algorithms can improve the quality of a recommender system [7, 26, 35, 37].

3. Method

3.1. Dataset

To examine our research questions, we designed a user interface that presented dinner recipes to users. This domain was selected due to the popularity and abundance of cooking recipes online [38], which may lead to choice overload. Since familiarity may mitigate the experienced choice overload [1, 3], we decided to only include vegetarian and vegan recipes, as these are consumed less frequently [39]. The domain selection was thus aimed to trigger choice overload among users. Forty recipes were selected for our evaluation, which were either vegetarian or vegan. Recipes were sampled from the popular Swedish recipe websites ICA.se and undertian.com, and fell into four different categories: ten pasta recipes, ten vegan recipes, ten stews, and ten salads.

3.2. Participants

A total of 150 participants were sampled from a convenience sample. They were recruited from various sources, as we applied a snowball sampling strategy. The majority of participants were recruited from surveypw.io, a tit-for-tat survey exchange platform. Participants recruited on that platform were compensated with points that allowed them to recruit participants for their own studies. All participants were at least 18 years old and were fluent in English. Unfortunately, no details were obtained on the gender and nationality of the participants. Note that the obtained data and analysis scripts are available in our repository: <https://osf.io/26u9g/>.

3.3. Research Design and Procedure

Participants were randomly presented one out of three user interfaces, either a single-list UI (n=53), a grid UI (n=55), or a multi-list UI (n=44). Each participants were asked to agree with the informed consent, after which they were presented the following scenario:

You and your three friends have planned to have dinner together this weekend and you have been chosen to decide what you will all eat. It is your responsibility to find a recipe that you believe will be well received by all of your friends. Since two of your friends are vegetarian you will have to take that into consideration. For inspiration, you go online to a recipe website in order to find the most suitable vegetarian recipe.

Afterwards, participants were presented a user interface with 40 recipes, which were not personalized to the user. Each participant was asked to inspect the presented recipes and to chose one recipe they liked the most. Afterwards, participants were asked to evaluate their choice and the system, inquiring on choice difficulty, choice satisfaction, and ease of use.

3.4. Interface

The recipe website comprised the 40 recipes in a single recommendation interface. The three different UIs are depicted in Figure 2. To avoid serial position effect the placements of the recipes were randomized for each interface [18, 40, 41, 42]. For the multi-list UI, which consisted of multiple rows of recipes, each row belonged to one of the four specific categories: pasta, salad, stew, and vegan. While the recipes in each row were randomized in that UI, the vertical order

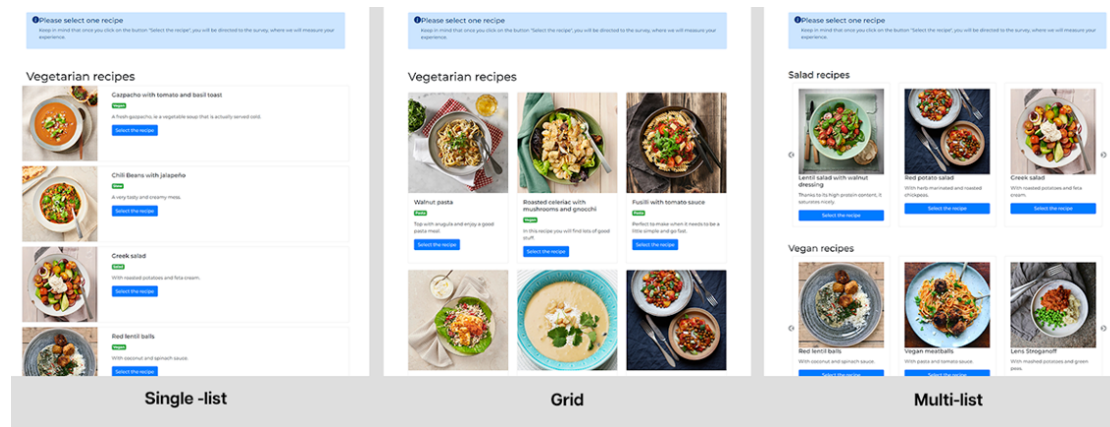


Figure 2: The interface in which recipes were presented to users. Depicted are the different UIs: single-list, grid and multi-list, as defined in the current study.

of the rows was always similar. In this case, all recipes in the salad category were displayed at the top, followed by the vegan, pasta and stew recipes – in this order.

Each recipe displayed an image of the dish, its name, a short description, and what category it belonged to. In order to minimize the effect an image may have had on the participants' decision, images were selected to be taken from similar 'helicopter view' angles, perpendicular to the dish. In addition, the selected images only depicted the dish itself, with the exception of possible utensils.

3.5. Evaluation Measures

To address [RQ1] and [RQ2], user perception and experience aspects were adapted from earlier work on UI design and recommender systems. The approach of measuring such aspects is in line with the recommender system evaluation framework of Knijnenburg et al. [10]. The propositions used in the study were adapted from earlier studies: for choice difficulty [6, 14, 28], choice satisfaction [6, 8], and ease of use [24, 43].

A principal component factor analysis was performed on the user responses, which were measured on 7-point Likert scales. The results are outlined in Table 1, which showed that we indeed could reliably infer three different user evaluation aspects: choice difficulty, choice satisfaction, and ease of use. In doing so, we applied promax rotation to allow for correlation between the different user aspects. One item was omitted due to low factor loadings (< 0.4), while the internal consistency of all aspects was found to be at least good (Cronbach's Alpha > 0.7).

Table 1

Results of the principal component factor analysis, which were inferred using promax rotation. Questionnaire items of the evaluation aspects were measured on 7-point Likert Scales. Items in gray were omitted from analysis due to low factor loadings.

Aspect	Item	Loading
Choice difficulty $\alpha = 0.81$	It was easy to choose a recipe.	-0.699
	The choice task was overwhelming.	0.787
	I found it difficult to choose a recipe from this list.	0.880
	I changed my mind several times before making a decision.	0.825
Choice satisfaction $\alpha = 0.75$	I am not satisfied with my chosen recipe.	0.842
	I like the recipe I've chosen.	-0.753
	I think I chose the best recipe among the available options.	0.887
Ease of use $\alpha = 0.81$	The layout of the website made it hard to consider all the recipes.	0.883
	It was easy to use the website.	0.890
	I found it easy to use the layout to search for recipes.	-0.685
	The website is user friendly.	0.665

4. Results

4.1. Choice Difficulty and Choice Satisfaction (RQ1)

We examined to what extent users experienced choice overload when interacting with our different recipe recommendation UIs. We used a two-way ANOVA to examine whether both a grid UI and a multi-list UI were evaluated more positively than a single-list UI. For choice difficulty, we did not observe any differences across the different conditions. Although choice difficulty was highest for the single-list UI ($M = 0.082$, $SD = 1.07$), it was not significantly higher than the difficulty experienced when using the grid UI ($M = -0.034$, $SD = 0.96$): $F(1, 147) = 0.35$, $p = 0.55$, nor significantly higher than the choice difficulty experienced when engaging with the multi-list UI ($M = -0.056$, $SD = 0.98$): $F(1, 147) = 0.44$, $p = 0.51$. This indicated that both the grid and the multi-list UIs did not significantly reduce choice difficulty when interacting with a recipe recommendation interface, which is also depicted in Figure 3.

For choice satisfaction, we neither observed any differences across conditions. Users were not more satisfied when choosing from a grid UI ($M = -0.036$, $SD = 0.99$) than when picking a recipe from a single-list UI ($M = 0.014$, $SD = 0.98$): $F(1, 147) = 0.07$, $p = 0.80$. In a similar vein, multi-list UIs ($M = 0.030$, $SD = 1.07$) neither led to a higher level of choice satisfaction, compared to single-list UIs: $F(1, 147) = 0.01$, $p = 0.94$. This indicated that the decision-making process was not evaluated more positively in grid-based or multi-list UIs, compared to a traditional single-list UI. These results are also depicted in Figure 4.

4.2. Perceived Ease of Use (RQ2)

We further examined differences in perceived ease of use, and whether this related to choice difficulty and choice satisfaction. A two-way ANOVA revealed that both the grid-based ($M =$

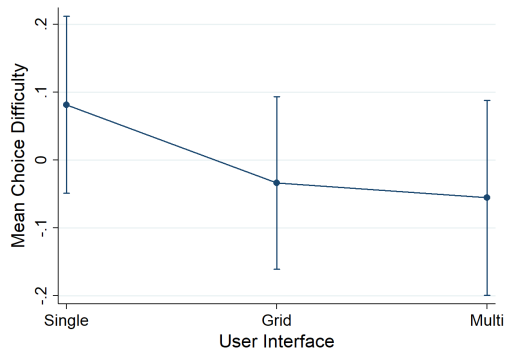


Figure 3: Standardized scores for the choice difficulty experience aspect across conditions. Errors bars represent 1 S.E.

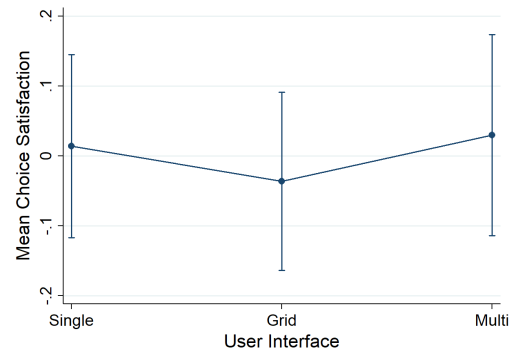


Figure 4: Standardized scores for the choice satisfaction experience aspect across conditions. Errors bars represent 1 S.E.

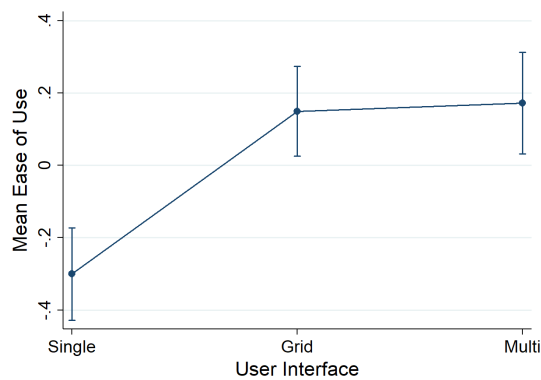


Figure 5: Standardized scores for the ease of use perception aspect across conditions. Error bars represent 1 S.E.

0.15, $SD = 0.92$) and a multi-list UIs ($M = 0.17$, $SD = 1.02$) were perceived as easier to use than the single-list UI ($M = -0.30$, $SD = 1.02$); for the grid UI: $F(1, 147) = 5.61$, $p = 0.019$; for the multi-list UI: $F(1, 147) = 5.45$, $p = 0.021$. This indicated that the a grid-oriented interface design, regardless of whether it involved explanations or categorization, led to a higher perceived of use. This result is also depicted in Figure 5.

We further examined whether ease of use was related to the choice difficulty and choice satisfaction experience aspects. To do so, we ran three different multiple linear regression models. Model 1 predicted choice difficulty using perceived ease of use ($F(1, 148) = 19.48$, $p < 0.001$), which indicated that ease of use was significantly and negatively related to choice difficulty: $\beta = -0.34$, $p < 0.001$. This indicated that users who perceived an UI as easy to use also experienced lower choice difficulty. This is also described in Table 2.

Model 2 predicted choice satisfaction using perceived ease of use. We found that it positively predicted choice satisfaction: $\beta = 0.20$, $p = 0.013$, which indicated that users who found an UI easy to use also tended to be more satisfied with their chosen recipe. Model 3 examined to

Table 2

Results of three different Multiple Linear Regression analyses. Model 1 predicted Choice Difficulty, while Models 2 and 3 predicted Choice Satisfaction. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

	Choice Difficulty	Choice Satisfaction	
	Model 1 β (S.E.)	Model 2 β (S.E.)	Model 3 β (S.E.)
Choice Difficulty			-0.14 (0.085)
Ease of Use	-0.34 (0.077)***	0.20 (0.080)*	0.16 (0.085)
R^2	0.116***	0.0410*	0.0572*

what extent the results from Model 2 would be consistent if choice difficulty was also added as predictor. Although a significant model was inferred ($F(2, 147) = 4.46$, $p < 0.01$), it did not reveal any significant relation between choice satisfaction and any of the two predictors: not for choice difficulty ($\beta = -0.14$, $p = 0.11$), nor for ease of use ($\beta = 0.16$, $p = 0.07$). Although another analysis indicated that choice difficulty and choice satisfaction were significantly related, it seemed that including both choice difficulty and ease of use into Model 3 led to neither predictor being significantly related to choice satisfaction. Taken together, these findings did suggest that the benefits of perceived ease of use seemed to translate to the two choice overload experience aspects, but that a clear path could not be established towards choice satisfaction when all three aspects are involved.

5. Discussion

We examined the role of different UIs on choice overload in the context of a recommender system scenario in the food recipe domain. In doing so, we have focused on the role of improving the UI rather than the presented content, while presenting content that is not necessarily personalized, following the approach of Starke et al. [8]. We have found that users of the single-list UI report significantly lower levels of perceived ease of use, than users of grid-based and multi-list UIs. In contrast, we have not observed any direct differences for choice satisfaction and choice difficulty as a result of our UI conditions. Although a small effect may have been present, we have not been able to observe this with the current sample size ($N = 150$), which would allow for medium effect sizes given the current research design.

Our findings for choice satisfaction and choice difficulty are consistent with Jannach et al. [5], who also report no differences across grid-based and multi-list UI conditions. However, it is at odds with Starke et al. [8] in that respect, because they report higher levels of choice difficulty for the multi-list conditions, compared to a grid-like condition¹.

The main difference in impact between the single-list UI and the two grid-like UIs (grid and multi-list) is the increase in the perceived ease of use. This concept has mainly been used in studies on UI design (e.g., [24]), and somewhat more uncommonly in recommender system studies (e.g., [31, 44]). It has been argued that grid-like interfaces allow for easier comparison

¹This is defined as a single-list condition by Starke et al. [8]

between items, such as in an e-commerce context [20]. The main weak point of single-list UIs becomes arguably apparent in contexts where the presented content is not necessarily tailored to the user, or in cases where there are not a few prominent items that stand out from the other options [6]; which arguably also applied to the current study. Although the direct experiential benefits of a multi-list recommender interface are limited in the current study, an increase in ease of use has also led to a reduction in choice difficulty and an increase in choice satisfaction; two indirect effects. Hence, it seems that mostly users who find an UI to be easy to use also make effective decisions, also regardless of whether this is because of the specific UI design. Such a possible path from objective system aspects through perception aspects to experience aspects (cf. [10]) is consistent with Starke et al. [8], who also report on a recommender system study that presents recipes that are not necessarily personalized. The main difference is that their perception aspect is diversity, which concerns the presented items rather than the UI, which is increased by a multi-list UI and, in turn, leads to a decrease in choice difficulty.

To come back to our research questions, we have only found indirect evidence that choice overload can be reduced through a multi-list interface. While choice difficulty and satisfaction have not directly varied across UIs, perceived ease of use does increase for a multi-list UI, which in turn has reduced choice difficulty and has also seemed to increase choice satisfaction. This study shows that different UIs can impact users' evaluation and possibly the content they interact with. It must also be noted that higher levels of choice satisfaction may be related to actual changes in behavior [33], but this is beyond the current study's design.

Our findings suggest that the benefits of a multi-list recommender interface may not be as profound as its widespread use suggests. Based on the evaluation aspects examined in this paper, its main benefits stem from an increase in ease of use, which may also improve a user's experience with a system. The main UI aspect in this case is the organization of the recommended items, rather than the presented explanations, for we have not observed any differences between grid-based and multi-list UIs. Thus, in the food domain, in an UI with a strong visual focus, multi-list UIs may be easier to use than an UI that organizes its item in a vertical way, but a recipe website may also present its content in a grid.

5.1. Limitations

Although much of the related literature and this study has touched upon the recommender system domain, this study has not investigated the aspect of algorithmic accuracy or quality of recommendations. Since the primary focus has been on the UI, we have included a recommender system scenario, albeit with no real personalization involved. In a follow-up study, this comparison between different UIs should also be performed in the context of personalized recommendations. Hence, previous studies have pointed out how varying inter-item and user-item similarity may affect the extent to which choice difficulty is induced [6].

Some behavioral and perception aspects have not been measured that could have further explained our findings. Among others, we have not allowed users to refrain from choosing a recipe (i.e., choice deferral). Moreover, a lack of familiarity with the options is a significant factor that moderate the extent to which choice overload is experienced, which should also be considered in a follow-up study. Instead, the items selected for this study, i.e., vegetarian recipes, have been selected as the overall familiarity with vegetarian recipes is likely to be low.

Nonetheless, we like to emphasize that our approach and method design overlaps with other studies investigating choice overload in an online context, with regard to the inquired aspects [5, 6, 8, 20, 24].

5.2. Future Work

Future research should also consider qualitative research methods to further examine the merits of a multi-list recommender interface. Such an approach could help to better understand users' views on the topic and to determine why they have preferences for certain UIs. In this sense, we are considering to conduct a case study on an existing website that involves personalized content. This would likely increase the realism of the task at hand. Moreover, as mentioned earlier, a future study should also examine whether user experience aspects are related to choice behavior. On top of that, other measures such as time spent on each UI, familiarity with the presented items, and allowing for the possibility of non-choices (i.e., choice deferral [36]), will also be included.

Acknowledgments

This work was supported by funding from the Wageningen University Digital Twin Program. In addition, it was supported by industry partners and the Research Council of Norway with funding to MediaFutures: Research Centre for Responsible Media Technology and Innovation, through the centers for Research-based Innovation scheme, project number 309339.

References

- [1] S. S. Iyengar, M. R. Lepper, When choice is demotivating: Can one desire too much of a good thing?, *Journal of personality and social psychology* 79 (2000) 995.
- [2] J.-Y. Park, S. S. Jang, Confused by too many choices? choice overload in tourism, *Tourism Management* 35 (2013) 1–12.
- [3] B. Scheibehenne, R. Greifeneder, P. M. Todd, Can there ever be too many options? a meta-analytic review of choice overload, *Journal of consumer research* 37 (2010) 409–425.
- [4] C. A. Gomez-Urbe, N. Hunt, The netflix recommender system: Algorithms, business value, and innovation, *ACM Transactions on Management Information Systems (TMIS)* 6 (2015) 1–19.
- [5] D. Jannach, M. Jesse, M. Jugovac, C. Trattner, Exploring multi-list user interfaces for similar-item recommendations, in: *Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization*, 2021, pp. 224–228.
- [6] D. Bollen, B. P. Knijnenburg, M. C. Willemsen, M. Graus, Understanding choice overload in recommender systems, in: *Proceedings of the fourth ACM conference on Recommender systems*, 2010, pp. 63–70.
- [7] M. Ge, C. Delgado-Battenfeld, D. Jannach, User-perceived recommendation quality-factoring in the user interface, *ACM RecSys* (2010) 22–25.

- [8] A. Starke, E. Asotic, C. Trattner, “serving each user”: Supporting different eating goals through a multi-list recommender interface, in: *Fifteenth ACM Conference on Recommender Systems*, ACM, 2021, pp. 124–132.
- [9] F. Wang, M. Wang, Y. Zheng, J. Jin, Y. Pan, Consumer vigilance and choice overload in online shopping, *International Journal of Electronic Commerce* 25 (2021) 364–390.
- [10] B. P. Knijnenburg, M. C. Willemsen, Z. Gantner, H. Soncu, C. Newell, Explaining the user experience of recommender systems, *User modeling and user-adapted interaction* 22 (2012) 441–504.
- [11] T. N. T. Tran, A. Felfernig, C. Trattner, A. Holzinger, Recommender systems in the healthcare domain: state-of-the-art and research issues, *Journal of Intelligent Information Systems* 57 (2021) 171–201.
- [12] F. Ricci, L. Rokach, B. Shapira, Recommender systems: introduction and challenges, in: *Recommender systems handbook*, Springer, 2015, pp. 1–34.
- [13] B. Ferwerda, M. P. Graus, A. Vall, M. Tkalcic, M. Schedl, How item discovery enabled by diversity leads to increased recommendation list attractiveness, in: *Proceedings of the Symposium on Applied Computing*, 2017, pp. 1693–1696.
- [14] M. C. Willemsen, M. P. Graus, B. P. Knijnenburg, Understanding the role of latent feature diversification on choice difficulty and satisfaction, *User Modeling and User-Adapted Interaction* 26 (2016) 347–389.
- [15] A. D. Starke, C. Trattner, Promoting healthy food choices online: a case for multi-list recommender systems, in: *Proceedings of the ACM IUI 2021 Workshops*, 2021.
- [16] E. J. Johnson, S. B. Shu, B. G. Dellaert, C. Fox, D. G. Goldstein, G. Häubl, R. P. Larrick, J. W. Payne, E. Peters, D. Schkade, et al., Beyond nudges: Tools of a choice architecture, *Marketing letters* 23 (2012) 487–504.
- [17] A. D. Starke, M. C. Willemsen, C. Trattner, Nudging healthy choices in food search through visual attractiveness, *Frontiers in Artificial Intelligence* 4 (2021) 621743.
- [18] M. Bar-Hillel, Position effects in choice from simultaneous displays: A conundrum solved, *Perspectives on Psychological Science* 10 (2015) 419–433.
- [19] L. Chen, P. Pu, Eye-tracking study of user behavior in recommender interfaces, in: *International conference on user modeling, adaptation, and personalization*, Springer, 2010, pp. 375–380.
- [20] Y. Kammerer, P. Gerjets, The role of search result position and source trustworthiness in the selection of web search results when using a list or a grid interface, *International Journal of Human-Computer Interaction* 30 (2014) 177–191.
- [21] M. Yener, O. Dundar, *Expert Android Studio*, John Wiley & Sons, 2016.
- [22] B. Rahdari, B. Kveton, P. Brusilovsky, The magic of carousels: Single vs. multi-list recommender systems, in: *Proceedings of the 33rd ACM Conference on Hypertext and Social Media*, 2022, pp. 166–174.
- [23] M. L. Resnick, C. Maldonado, J. M. Santos, R. Lergier, Modeling on-line search behavior using alternative output structures, in: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 45, SAGE Publications Sage CA: Los Angeles, CA, 2001, pp. 1166–1170.
- [24] L. Chen, H. K. Tsoi, Users’ decision behavior in recommender interfaces: Impact of layout design, in: *RecSys’ 11 Workshop on Human Decision Making in Recommender Systems*,

2011.

- [25] J. Beel, H. Dixon, The ‘unreasonable’ effectiveness of graphical user interfaces for recommender systems, in: *Adjunct Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization*, 2021, pp. 22–28.
- [26] A. A. Ozok, Q. Fan, A. F. Norcio, Design guidelines for effective recommender system interfaces based on a usability criteria conceptual model: results from a college student population, *Behaviour & Information Technology* 29 (2010) 57–83.
- [27] P. Pu, L. Chen, R. Hu, A user-centric evaluation framework for recommender systems, in: *Proceedings of the fifth ACM conference on Recommender systems*, 2011, pp. 157–164.
- [28] M. C. Willemsen, B. P. Knijnenburg, M. P. Graus, L. Velter-Bremmers, K. Fu, Using latent features diversification to reduce choice difficulty in recommendation lists, *RecSys* 11 (2011) 14–20.
- [29] B. Fasolo, R. Hertwig, M. Huber, M. Ludwig, Size, entropy, and density: What is the difference that makes the difference between small and large real-world assortments?, *Psychology & Marketing* 26 (2009) 254–279.
- [30] R. Mátrai, Z. T. Kosztyán, Navigation strategies in case of different kind of user interfaces, in: *User Interfaces*, IntechOpen, 2010.
- [31] R. Hu, P. Pu, Enhancing recommendation diversity with organization interfaces, in: *Proceedings of the 16th international conference on Intelligent user interfaces*, 2011, pp. 347–350.
- [32] S. M. McNee, J. Riedl, J. A. Konstan, Being accurate is not enough: how accuracy metrics have hurt recommender systems, in: *CHI’06 extended abstracts on Human factors in computing systems*, 2006, pp. 1097–1101.
- [33] A. Starke, M. Willemsen, C. Snijders, Effective user interface designs to increase energy-efficient behavior in a rasch-based energy recommender system, in: *Proceedings of the eleventh ACM conference on recommender systems*, 2017, pp. 65–73.
- [34] P. Pu, L. Chen, Trust-inspiring explanation interfaces for recommender systems, *Knowledge-Based Systems* 20 (2007) 542–556.
- [35] E. Murphy-Hill, G. C. Murphy, Recommendation delivery, in: *Recommendation systems in software engineering*, Springer, 2014, pp. 223–242.
- [36] R. Dhar, Context and task effects on choice deferral, *Marketing letters* 8 (1997) 119–130.
- [37] B. P. Knijnenburg, N. J. Reijmer, M. C. Willemsen, Each to his own: how different users call for different interaction methods in recommender systems, in: *Proceedings of the fifth ACM conference on Recommender systems*, 2011, pp. 141–148.
- [38] C. Trattner, D. Elswiler, Investigating the healthiness of internet-sourced recipes: implications for meal planning and recommender systems, in: *Proceedings of the 26th international conference on world wide web*, 2017, pp. 489–498.
- [39] S. M. Hargreaves, A. Raposo, A. Saraiva, R. P. Zandonadi, Vegetarian diet: an overview through the perspective of quality of life domains, *International journal of environmental research and public health* 18 (2021) 4067.
- [40] M. Mandl, A. Felfernig, E. Teppan, M. Schubert, Consumer decision making in knowledge-based recommendation, *Journal of Intelligent Information Systems* 37 (2011) 1–22.
- [41] A. Mantonakis, P. Rodero, I. Lesschaeve, R. Hastie, Order in choice: Effects of serial position on preferences, *Psychological Science* 20 (2009) 1309–1312.

- [42] J. Murphy, C. Hofacker, R. Mizerski, Primacy and recency effects on clicking behavior, *Journal of computer-mediated communication* 11 (2006) 522–535.
- [43] A. M. Lund, Measuring usability with the use questionnaire12, *Usability interface* 8 (2001) 3–6.
- [44] R. Hu, P. Pu, A study on user perception of personality-based recommender systems, in: *International conference on user modeling, adaptation, and personalization*, Springer, 2010, pp. 291–302.