Date and Rate

The effects of a 5-star rating system in dating apps on target trustworthiness and intrinsic motivation

Authors
Sofia Wendt & Axel Ydenius

May 2021
Stockholm School of Economics

Abstract

The dating app industry is currently experiencing tremendous growth in the number of users. Simultaneously, ratings are becoming common practice on various platforms to communicate the reliability and trustworthiness of their users. While the effects of ratings have been the subject of countless previous studies, ratings in the dating app context remain neglected. The main purpose of this study is therefore to determine the effect of a 5-star rating system on target trustworthiness and contrast it with the suggested undermining effects on intrinsic motivation as a result of measurement in enjoyable activities. By using two quantitative experimental studies, we show that high ratings indeed increase the perceived trustworthiness of a user. Additionally, we find that the exposure to rating systems affects intrinsic motivation negatively as it makes the dating app activity seem more like work and less fun. However, we find no significant evidence of its effect on enjoyment. The studies also find that the difficulty of choosing among available profiles increases with the presence of ratings for women, but has the inverse effect for men, and that less attractive dating profiles increases trustworthiness.

Acknowledgements

Micael Dahlén

Thank you for always believing in us and our ideas. Thank you for also making them possible when we were in doubt. Thank you for being the "snabbaste professorn i stan".

Tavish Gantz

Thank you for knowing so much about the English language. Invaluable to say the least.

Alexander Mafael

Thank you for your interest and invaluable literature suggestions.

Anonymous Models

Even though you will never read this, thank you for making the simulations possible.

The Authors

Sofia Wendt

50546

Axel Ydenius

50619

Introduction

1.1 Background

Do ratings in dating make it more trustworthy but less fun? The presence of ratings is undeniably increasing in multiple contexts and they are becoming a greater part of everyday life. Today, we rate movies, products, restaurants, hotels, chauffeurs, doctors, and even professors (https://www.ratemyprofessors.com/). It is therefore arguably only a matter of time before ratings become common in the dating app realm as well.

Dating apps have since the launch of Tinder in 2012 revolutionized the dating culture, and the implementation of ratings could be the next step. The apps have become increasingly popular in recent years and the trend does not seem to weaken any time soon. A survey conducted by Stanford University illustrates that in the recent period of 1995-2017 the share of heterosexuals who first meet their partner online has risen from 0 to 39 percent and is now the most common way to initiate a relationship (US) (Rosenfeld et al., 2019). It is also forecasted that close to 655 million people will use dating applications worldwide in 2024, which is an increase of close to 45 percent compared to the number of users in 2020 (Dating Services, 2021). In other words, the dating app industry is experiencing high growth, and while observing the increasing presence of ratings in society, we believe researching their effects in dating to be a 'perfect match'.

1.2 Problem Area

Ways to increase trust online are essential on platforms that require human interaction since their relative anonymity means few ramifications for poor behavior (Drouin et al., 2016). On dating apps, for example, statistics show that 80 percent users lie in some respect (Roumeiu, 2011) and, while most dishonesty is practically harmless, online dating scams involving monetary losses are increasingly rife (Fletcher, 2021; TV4, 2021). Gender-specific trust issues are also a growing concern. 70 percent of female users cite offensive behavior as the most unsatisfactory part of their experience while a further 50 percent ultimately stop using the app altogether (Lopes & Vogel, 2017). Male users on the other hand experience that females to a greater extent lie about or misrepresent their physical appearance (Whitty, 2008), leading to mismatched expectations prior to meeting the date in real life.

The solution is not obvious. Adding ratings as a measurement of another person's on- and offline behavior can provide information about the date's trustworthiness, but any attempt to measure enjoyable activities has also been shown to have an undermining effect on a person's intrinsic motivation (Etkin, 2016; Deci et al., 1973).

1.3 Purpose and Expected Research Contribution

To date, research has been conducted on the effects of reputation systems in hotel bookings (Dahlén & Thorbjornsen, 2021), online marketplaces (Bente et al., 2012; Malinen & Ojala, 2011) and a number of P2P business models (Liu et al., 2016; Wang et al., 2015). However, to our knowledge, no research has been conducted on reputation systems in a dating app context, and certainly not on the effects of trustworthiness. Furthermore, while the positive effects of ratings have been studied to a great extent, the potential negative effects have been much overlooked, one disregarded area being ratings' effect on intrinsic motivation. Previous studies on this subject have researched the undermining effects on intrinsic motivation measuring enjoyable activities people are performing themselves (Etkin, 2016; Deci et al., 1973). We contribute to this literature by investigating the effects of measurement on others in the form of ratings.

While the primary aim of this thesis is to illustrate the effects on trustworthiness and intrinsic motivation, we also contribute to other fields of research. For instance, it has been concluded that ratings have facilitating effects on how consumers choose among available alternatives (Filieri, 2015), and we extend this understanding in our setting of selecting between dating profiles. Cues of others' personalities have been shown to affect the perceived attractiveness of the counterpart. We determine whether ratings facilitate enough conviction in this respect (Fiore et al., 2008). Finally, we investigate if physical attractiveness itself can contribute to higher degrees of trust in dating apps by virtue of the halo effect.

The purpose of our study is to close the current gap in research on ratings and dating and to contrast the previously suggested positive *and* negative effects in this new setting. We seek to illustrate the implications of ratings both for the individual user and for the dating app developers. With two experimental studies, we intend to answer the question *"How does a 5-star rating system affect target trustworthiness, target attractiveness, choice difficulty, and intrinsic motivation in dating apps?"*. We seek to answer the first half of the research

question with an experiment using more and less attractive dating profiles that are highly or poorly rated. The second half will be investigated by simulating a dating app experience where the profiles are manipulated with ratings.

1.4 Delimitation

While the growing industry of dating apps is apparent on a global level, this thesis is limited to a Swedish setting with only Swedish respondents. We also study dating apps on a general level, meaning that we wish for the results to be applicable to most dating app services. However, we still had to manufacture fictitious profiles. We chose an age range of 18 to 50 years, and we used the Tinder design as a template since it is the most used app on the market (AppMagic, 2021).

1.5 Disposition

To fulfill the purpose of this thesis, two separate quantitative studies have been conducted. Study 1 investigates the potential positive outcomes of 5-star rating systems in dating apps and seeks to answer how ratings affect target trustworthiness and attractiveness. Study 2 explores whether ratings contribute to decreased choice difficulty but also to decreased motivation. In both experiments we use a true-to-life Tinder simulation which will be described in depth in the methodology sections. First, the theories underlying each study will be presented. Second, the methodology and the results of the experiments are presented for each study respectively. We then proceed to discuss and contrast the two studies in a discussion that includes practical implications and suggested future research.

Theoretical framework

Trustworthiness

Trust is important in any situation that carries risk, interdependence, or ambiguity (McKnight, 2001). First encounters with strangers entail both ambiguity and uncertainty regarding the other person's intentions and future behavior. This especially applies in a dating app context where only limited information is given before the in-person interaction. Similarly to the reasoning in Liesel (2020), means of reducing uncertainty and ambiguity are therefore especially salient in a dating app setting because individuals benefit from the exposure to more information in the initial stages of their new potential partner prior to the date.

Thus, uncertainty reduction theory (URT) is the starting point and underlying framework assessing trustworthiness of users' dating app profiles in this study. A central assumption of URT is that information seeking is used to reduce the uncertainty when initiating new relationships (Berger & Calabrese, 1975). URT further states that if a person anticipates a future interaction (a date), they will be motivated to gain more information in order to reduce the level of uncertainty associated with that meeting. Additionally, motivation to gain more information increases when the other person behaves in abnormal ways. This is relevant in a dating app context due to the veracity uncertainty issues often present in online dating (Norcie et al., 2013).

It should be mentioned that some critique has been aimed towards URT, for instance by Kramer (1999). His paper is in the context of organizational communication, and the critique is that there are different *low or high motivational* factors for reducing uncertainty. He points out that easily understood situations do not require reduction of uncertainty, but we safely assert that a dating app date cannot be dismissed as an easily understood situation due to the unlimited possible outcomes of different matches.

To conclude, the inherently low tolerance for ambiguity leads people to seek necessary information to render future interactions predictable (Berger & Calabrese, 1975). Thus, the more information is gathered, the less uncertainty is involved and the more the counterpart's trustworthiness can increase.

Our source of additional information to reduce uncertainty in the dating sphere is a 5-star rating system based on the users' personality and behavior on- and offline. With the rise of digital consumption and the peer-to-peer business model, other people's reviews have become increasingly important as a means to reduce levels of uncertainty. Some services, namely Uber and Airbnb, have implemented 5-star rating systems where consumers can review their experience and in turn provide additional information to future customers. Ratings have shown to have a significant impact on people's perception about a product or service. Recent research in a hotel setting conducted by Dahlén and Thorbjornsen (2021) found that people are more likely to trust and choose a hotel that is reviewed with a high rating and a bad review rather than a hotel with a bad rating and a good review. Moreover, ratings can act as an indicator and carry information about the trustworthiness and authenticity of a target, which in turn can help to reduce uncertainty prior to making a choice.

Based on conclusions drawn by aforementioned research, we hypothesize that a 5- star rating system can be implemented to reduce uncertainty in a dating app context. Specifically, a high rating given to a person on a dating profile, in comparison to a low rating, will increase trust towards that person. We therefore hypothesize:

> **H1:** A profile with a higher rating is perceived as more trustworthy than a profile with a lower rating

Attractiveness

Physical attractiveness also carries much information. Men have been shown to be perceived as more attractive when their photos in online dating seem trustworthy and genuine, but not too warm and kind. Women's profiles have on the other hand been rated as more attractive when they looked more feminine and projected low self-centeredness and high self-esteem (Fiore et al., 2008). The same study also concluded that while physical appearance matters, the often disregarded biography option in online dating has the power to affect the overall perceived attractiveness of another's profile. The biography's purpose is to shed a light on the user's good qualities - a concise description of oneself and personality. Adding to this, Paunonen (2006) found perceived attractiveness to be positively affected by a personality trait, namely how honest a person is. One could then argue that the implementation of an additional informational factor, such as ratings that allow users to infer similar traits as a

biography does in describing another person's personality and good qualities, could have an effect on the perceived attractiveness of a dating app profile as a whole.

Other studies have also confirmed that people tend to rely on one or a few positive qualities in a person and transfer those to other unrelated traits, a cognitive bias commonly known as the *halo effect* (Forgas & Laham, 2016). It is therefore probable that the cues of a good rating would imply a desirable personality and thus result in a more attractive profile. Furthermore, Kellerman and Reynolds (1990) found that the predictive power of others' behaviors influences the perceived attractiveness where receipts of positive outcomes increased perceived attractiveness and where negative outcomes had the inverse effect. A rating would then have the function of managing the expectations of the person's future behavior.

Extending this research in our study, we intend to investigate the effect on perceived attractiveness given a high versus a low rating that represents how a person has behaved. We hypothesize that high ratings will increase perceived attractiveness for both more attractive and less attractive profiles.

> **H2:** A profile with a higher rating is perceived as more attractive than a profile with a lower rating

H2 investigates whether the cues of a good personality with the help of ratings influences the perceived attractiveness. Another probable outcome of the halo effect in this setting is instead that the objective physical attractiveness of a picture in a profile influences the trustworthiness of that profile. In short, a more attractive person should be perceived as more trustworthy and a less attractive person as less trustworthy.

> **H3:** A more attractive profile picture is perceived as more trustworthy than a less attractive profile picture

Enjoyment

Deci and Ryan's *Cognitive Evaluation Theory of Motivation* (CET) is designed to explain external factors' facilitating or undermining effects on intrinsic motivation (Deci & Ryan, 1985; Deci & Ryan, 2008). Deci & Ryan describe *intrinsic motivation* as that which is

generated internally and is self-authored, whereas *extrinsic motivation* is founded in external controls such as rewards and feedback. When comparing these two types of motivation, anyone driven by the former typically experiences more interest, excitement, and confidence, which in turn manifests in outcomes such as enhanced enjoyment and well-being (Deci & Ryan, 2008).

Different studies on how intrinsic motivation is affected by different types of rewards have been published ever since the concept was coined (Deci, 1971; Sansone & Harackiewicz, 2000). The rewards studied have been both monetary ones and social approval in the form of verbal feedback. In general, rewards have shown to have negative effects on intrinsic motivation to perform a certain activity (Deci, 1971; Deci et al., 1973; Sansone & Harackiewicz, 2000).

Resting on this theory, recent research conducted by Jordan Etkin (2016) digs into the effects on intrinsic motivation and *personal quantification* (tracking of behavioral output). In several experiments, Etkin demonstrates how measurement (e.g. tracking an enjoyable activity) can boost efficiency while the quantification and measurement is present but also undermines intrinsic motivation by reducing how much people enjoy performing the activity. In our study, we want to contrast these effects. Personal quantification has been shown to make naturally enjoyable activities seem more like work and to reduce enjoyment, which in turn decreases intrinsic motivation for performing that activity (Etkin, 2016). While Etkin tried the effect on activities including coloring and walking, the present study investigates online dating. In other words, our use of 'activity' refers to the dating app swiping session. Further, while the previously mentioned studies have focused on measuring an activity that a person is performing, we instead wish to see if similar results hold true when undertaking the measurement of *another* person. In other words, the measurement that will be studied is a social approval feedback of another individual's behavior and is captured by the quantifiable 5-star rating system.

Building on this research, one of the main focus areas of this study is therefore to examine if similar results hold true in an online dating context. We hypothesise that a 5-star rating system as feedback, similar to Deci et al's (1973) social approval feedback, should decrease the enjoyment of using the dating app and thereby undermine the intrinsic motivation.

H4: Ratings in dating apps reduce the enjoyment of performing the activity

Perceived work

Enjoyable activities are viewed as intrinsically motivated. An activity that feels more like work on the other hand is viewed as being *more* of a controlling situation and thus more extrinsically motivated (Etkin, 2016). To see if the activity is being perceived as more controlled and thus undermining intrinsic motivation, we also hypothesize:

H5: The perceived work of performing the activity is greater with the presence of ratings

Choice difficulty

In contrast to feedback having a negative effect on intrinsic motivation, it has also been shown to have the positive effect of increasing efficiency of performing the activity as long as the feedback is present (Etkin, 2016). In our context, higher efficiency translates to choosing among profile alternatives with less difficulty. Therefore, we believe a 5-star rating system should decrease choice difficulty. Additionally, when studying product ratings, Filieri (2015) concluded that ratings have a facilitating effect on evaluations among available alternatives. It is therefore probable that the presence of ratings (good or bad) will aid in the decision-making process for selecting between profiles. Furthermore, drawing on the assumptions from Frost (2008), that people are experience goods, we expect a corresponding effect of ratings in a dating context. Thus, we hypothesize the following:

H6: Choice difficulty decreases with the presence of ratings in a dating app context

Study 1

This study investigates the first half of our research questions and the effects that a 5-star rating system has on target trustworthiness and perceived attractiveness.

3.1 Methodology

In this section we motivate the choice of method for Study 1 and explain how we conducted the experiment and collected the data.

3.1.1 Study design

In order to test the causal claims between our independent and dependent variables we employed an experimental research design. We argue that a quantitative experimental approach is the most suitable as it allows us to manipulate the independent variables (i.e. rating and attractiveness of profile) in order to compare reactions between groups and further to test the causality in our hypotheses (Söderlund, p.16, 2018). Moreover, as the experiment was intended to test the perception in terms of trustworthiness and attractiveness of a dating app profile depending on the profile's attractiveness and rating score, we replicated a real life Tinder profile design.

The experiment consisted of a fictitious Tinder profile followed by a questionnaire and was distributed online in a survey format. The survey respondents were randomly allocated to one out of four groups where two of the groups had the same rating (either 1 star or 5 stars) and two were of equal attractiveness (less and more attractive).

3.1.2 Preparatory work

Our selection of experiment design entailed the need to choose pictures for the fictitious profiles. As one of two independent variables in this experiment was the attractiveness of the person in the profile, it indicated a need to find people considered *more attractive* and *less attractive*. As we assumed respondents to be attracted to different ages, we decided to include three age ranges (18-29, 30-39, and 40-50) in the survey, hence we needed one profile covering each age range. Also, assuming respondents to be attracted to different genders, we needed both female and male profiles. Taken together, these conditions resulted in 6 different more attractive profiles and 6 different less attractive profiles. We motivate our initial choices

of the more attractive and the less attractive pictures using Gulas & McKeage's (2000) research on the nearly automatic tendency to categorize and evaluate human physical appearance as more attractive or less attractive. With our selection of models completed, we further conducted a pre-study to test the accuracy in attractiveness of the models selected for the profiles. This was done in a short survey where (n=21) respondents were asked to rank each model based on how attractive they thought he/she was on a scale ranging from 1 to 10 (1=less attractive, 10=very attractive). The scores were analysed using one-sample t-tests where the means of each profile pair were compared (i.e. less attractive female 18-29 compared to more attractive female 18-29, less attractive male 18-29 compared to more attractive male 18-29 etc.). The results from the pre-study are presented in **Table 1** and confirmed our assessment of what is considered a more or less attractive model. The profile pictures were retrieved from Unsplash, a database with non-copyrighted photos of anonymous models (https://unsplash.com/).

Table 1

One sample t-test	More attractive profile		Less attractive profile			
	Mean	SD	Mean	SD	Sig	t
Man 18-29	8.81	0.93	2.38	1.66	<0.001	43.48
Man 30-39	8.67	1.35	2.81	1.72	<0.001	29.33
Man 40-50	9.10	1.26	2.52	1.66	<0.001	33.05
Woman 18-29	8.86	1.20	2.76	1.48	<0.001	33.96
Woman 30-39	8.76	1.22	2.81	1.83	<0.001	33.89
Woman 40-50	8.81	1.21	2.14	1.88	<0.001	33.39

1=less attractive, 10=very attractive   n=21

3.1.3 Survey design

The survey consisted of the profile stimuli followed by a questionnaire. First, a brief description of the implication of the rating was introduced, namely, an assessment made by other users of the person's behavior. Followingly, the survey had two questions allocating respondents depending on their individual dating preferences. The reason for this was to mirror reality as closely as possible as these are available functions in most dating apps. The first sorting question was mainly used to allocate respondents to the appropriate profile based on gender choice and therefore asked "What gender would you like to see in the profile?".

The formulation of this question was made with careful consideration to be inclusive of all sexualities and written so that no respondent would get irritated, annoyed or offended. The preferred age range can also be set in most real life dating apps because people are attracted to people of particular ages. Therefore, to ensure both the internal and external validity of our study and make sure that the respondents were exposed to ages appropriate to them, different age ranges could also be chosen by our respondents. The available options to choose from were 18-29, 30-39 and 40-50. The manufacturing of the fictitious profiles was especially time consuming, therefore, it was limited to these three age ranges purely for practical reasons. To conclude, by allowing respondents to choose their gender of interest themselves, the implication of their sexual orientation becomes irrelevant, and by allowing them to pick age of interest they could choose dating objects based on their own preferences. Upon choosing dating preferences, respondents were then randomly allocated to one of the four treatments and subsequently asked questions on perceived trustworthiness and perceived attractiveness of the profile.

When measuring preferences and perceptions (i.e. perceived trustworthiness and attractiveness) we used interval scales (Söderlund, p. 93, 2005). Specifically, we adopted a 10-Point Likert scale. While a scale with even numbers implies that respondents are forced to have an opinion, it directs respondents to not choose the neutral option (Söderlund, p. 93, 2005). To ensure content validity, the overlap between our response options and previous research on the subject (Söderlund, p. 136, 2018), we have chosen variables and response options coherent with the correct definitions collected through qualitative reviews of previous studies and established scales (this applies to both studies).

3.1.3.1 Dependent variables

Target trustworthiness

To measure respondents' perceived trustworthiness of the person in the profile, a battery of three questions was adopted from the IMI-scale that measures different variables of relatedness. However, these resulted in an insufficient (Söderlund p. 136, 2018) Cronbach alpha level of 0.55, thus they could not be indexed. The first question in the battery, "I feel like I could really trust this person" was excluded as it showed no significance at any acceptable p value ($<.10$). The second question, "I prefer not to engage with this person in the future" was disregarded because it was intuitively the least related to trustworthiness of the three questions. Finally, "I don't feel like I could really trust this person" was used as it

showed high significance and is an intuitively acceptable measure of trustworthiness. Since the remaining question was a negative, we inverted the data in the analysis phase. The answers could be specified on a 10-point Likert scale with the endpoints being "not at all" and "very much".

Target attractiveness

To measure the perceived attractiveness of the person in the profile, the question "How attractive do you perceive the person in the profile to be?" was asked on a 10-point Likert scale with endpoints "not at all" and "very much".

3.1.3.2 Survey Distribution and Manipulation checks

Both surveys were distributed via Norstat Sverige AB, a company specialized in collecting responses to questionnaires in exchange for monetary compensation. In these circumstances it is not uncommon for the respondents to have incentives to complete the survey as quickly as possible, so the risk occurs that they do not pay enough attention to the instructions given. Therefore, in order to make sure we only included participants who understood what they were subjected to and to increase the reliability of the final dataset further (Oppenheimer et al., 2009), instructional manipulation checks were employed (Söderlund, p.96, 2018). The manipulation checks were designed as follows: "What was this survey about?"; A) Purchasing behavior, B) Dating apps, C) Movie reviews. In addition, an additional attention-related question was also included where the participants were asked to "check box 7" on a 1-10 interval scale. Wrong answers resulted in discarding the respondent in entirety from the final dataset. Nonetheless, using a panel with a sample of respondents that represents the total population strengthens the external validity of our study.

3.1.3.3 Stimuli

In this experiment, the fictitious Tinder profiles were manufactured as attractive and less attractive, but with similar facial features and photo angles as can be seen in **Figure 1**. The profiles were then also fitted with high and low ratings. This resulted in four different stimuli that our four different groups were exposed to. In short, the level of attractiveness was put against high and low rating scores as below:

1. "More attractive person" with high rating - five star rating
2. "More attractive person" with low rating - one star rating
3. "Less attractive person" with high rating - five star rating
4. "Less attractive person" with low rating - one star rating
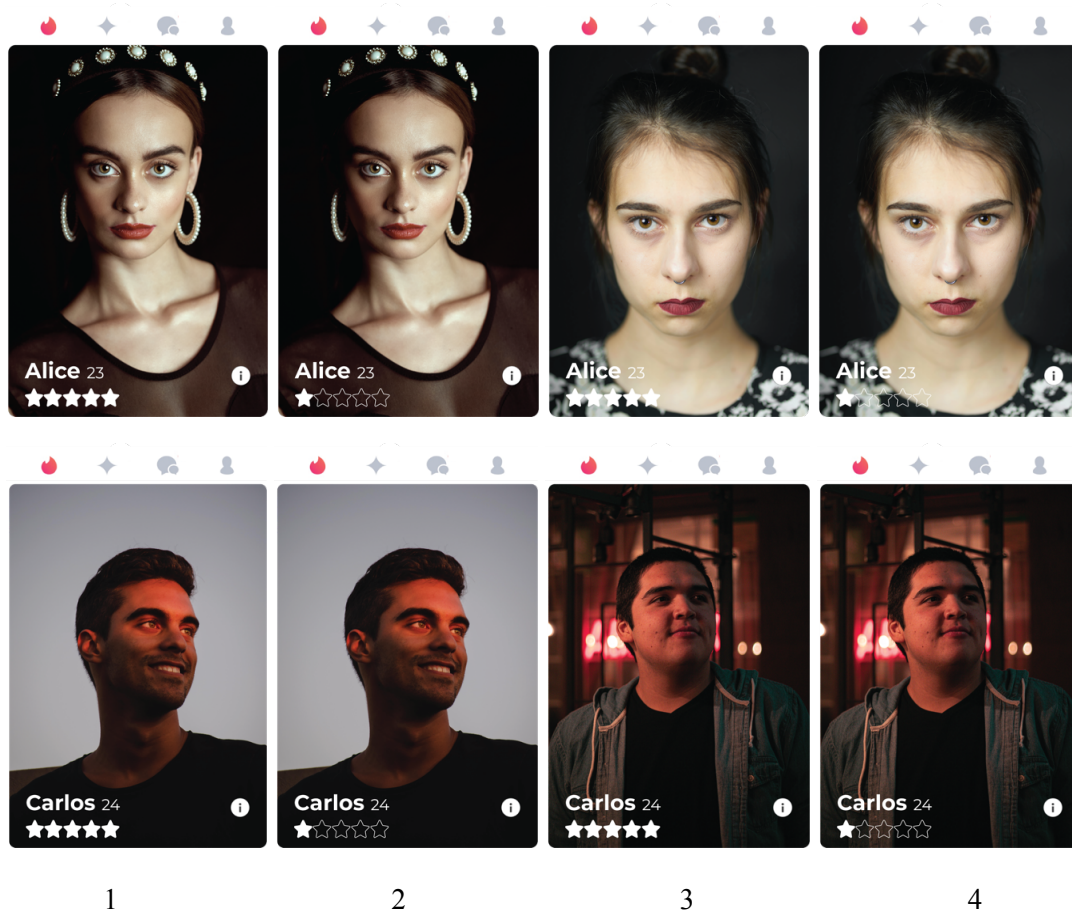

1   2   3   4

Figure 1: More and less attractive profiles with good and bad ratings.

3.1.4 Sampling

The survey was public between March 11th and March 26th. The dataset without any modification contained (n=374) participants. Respondents who failed to answer the control questions in the manner described in section 3.1.3.2 were deleted from the final dataset (n=54). This resulted in (n=320) valid responses with a mean age of 32,12 (SD=9,062). The distribution between genders were women (n=180), men (n=139) and "other" (n=1). Women requested to have profiles of men displayed to a greater extent (n=162) than profiles of women (n=18). Similarly, men wanted to be displayed profiles of women (n=124) to a greater extent than profiles of men (n=15). As intended there were both dating app users (n=151) and non-users (n=169) respondents.

3.2 Results

*In this section we present the results of testing hypotheses **H1-H3** about rating systems' effect on target trustworthiness and perceived attractiveness, but also how physical attractiveness affects trustworthiness.*

Trustworthiness

Table 2							
Independent samples t-test	High Rating		Low Rating				
	Mean	SD	Mean	SD		Sig	*t*
Trustworthy	6.67	2.31	6.02	2.34		0.006	2.53
Trustworthy$_{women}$	6.74	2.27	5.60	2.40		0.0005	3.27
Trustworthy$_{men}$	6.60	2.37	6.54	2.15		0.43	1.67

1=not at all, 10=very much   n=320, women n=180, men n=139

First, a one-sided independent t-test was made for the mean comparison between the high and low level rating treatment groups on perceived trustworthiness of the profile. As Table 1 indicates, a high rating resulted in a higher degree of trust in the person behind the profile to which they were exposed ($M_{high,rating}$=6.67, SD=2.31) and a low rating treatment resulted in lower trustworthiness ($M_{low,rating}$=6.02, SD=2.34) conditions; t(318)=2.53, p=0.006, **thus we accept H1.**

For exploratory reasons, we set out to investigate differences in trust depending on gender and found that there was a difference evoked as a result of high or low rating stimuli. Women perceive men with higher ratings ($M_{high,rating,women}$=6.74, SD=2.27) to be significantly more trustworthy than men with low ratings ($M_{low,rating,women}$=5.60, SD=2.40), conditions; t(178)=3.27, p<0.001 . Although the same direction was found for the male respondents, the effects of ratings showed no significant effect in trustworthiness on the female profiles ($M_{high,rating,men}$=6.60, SD=2.37; $M_{low,rating,men}$=6.54, SD=2.15), conditions; t(137)=0.17, p=0.43 as shown in Table 1.

Perceived attractiveness

Table 3

Independent samples t-test	High Rating		Low Rating			
	Mean	SD	Mean	SD	Sig	*t*
Percieved attractiveness	5.36	2.63	5.38	2.16	0.48	-0.06
Percieved attractiveness_woman	5.23	2.50	5.54	2.18	0.18	-0.91
Percieved attractiveness_man	5.53	2.78	5.13	2.16	0.17	0.95

1=not at all, 10=very much   n=320, women n=180, men n=139

To illustrate the effects of high and low ratings on perceived attractiveness of a profile we conducted another one-sided independent t-test. The attractiveness of the profile was not affected by higher or lower ratings ($M_{high,rating}$=5.36, SD=2.63; $M_{low,rating}$=5.38, SD=2.16), conditions; t(304)= -0.059, p=0.48, **so we reject H2.**

For the same reasons as in H1, we also investigated potential differences between gender. However, no significant differences were found for women ($M_{high,rating,women}$=5.23, SD=2.50; $M_{low,rating,women}$=5.54, SD=2.18), conditions; t(178)=, p=0.18 and no differences were found for men ($M_{high,rating,men}$=5.53, SD=2.78; $M_{low,rating,men}$=5.13, SD=2.14), conditions; t(137)= -0.91, p=0.17.

Trustworthiness as an effect of attractive profile picture

Table 4

Independent samples t-test	More attractive		Less attractive			
	Mean	SD	Mean	SD	Sig	*t*
Trustworthy	6.13	2.42	6.54	2.26	0.06	-1.56
Trustworthy_woman	5.56	2.39	6.66	2.32	0.003	-2.85
Trustworthy_man	6.75	2.33	6.41	2.18	0.18	0.91

1=not at all, 10=very much   n=320, women n=180, men n=139

To test whether the halo effect instead held true in the opposite direction, being that physical attractiveness is transmitted to other unrelated traits as trustworthiness, we once again compared the means with one-tailed independent t-tests. A significance level of .10 was used in the test. Unexpectedly, a more physically attractive profile picture resulted in a lower degree of trustworthiness ($M_{attractive,picture}$=6.13, SD=2.42), and a less attractive profile picture commanded less favorable assessments of the profile's trustworthiness ($M_{less,attractive,picture}$=6.54, SD=2.26), conditions; t(318)= -1.56, p=0.06. **Thus, we reject H3.**

For consistency reasons, we also explored the potential for gender differences. Once again, only women showed significant differences after being exposed to attractive or less attractive profile pictures ($M_{attractive,picture,women}$=5.56, SD=2.39; $M_{less,attractive,picture,women}$=6.66 SD=2.32); $t(178)$= -2.85, p=0.003. Men showed the opposite result, but insignificantly so. ($M_{attractive,picture,men}$=6.75, SD=2.33; $M_{less,attractive,picture,men}$=6.41, SD=2.18); $t(137)$=0.91 p=0.18.

In addition to testing our hypothesis, we measured how many no and yes answers each kind of profile received. As the graphs below (Figures 2 and 3) unveil, although only significant for the more attractive profiles, high ratings received more right swipes (yes) to a greater extent than profiles with low ratings.
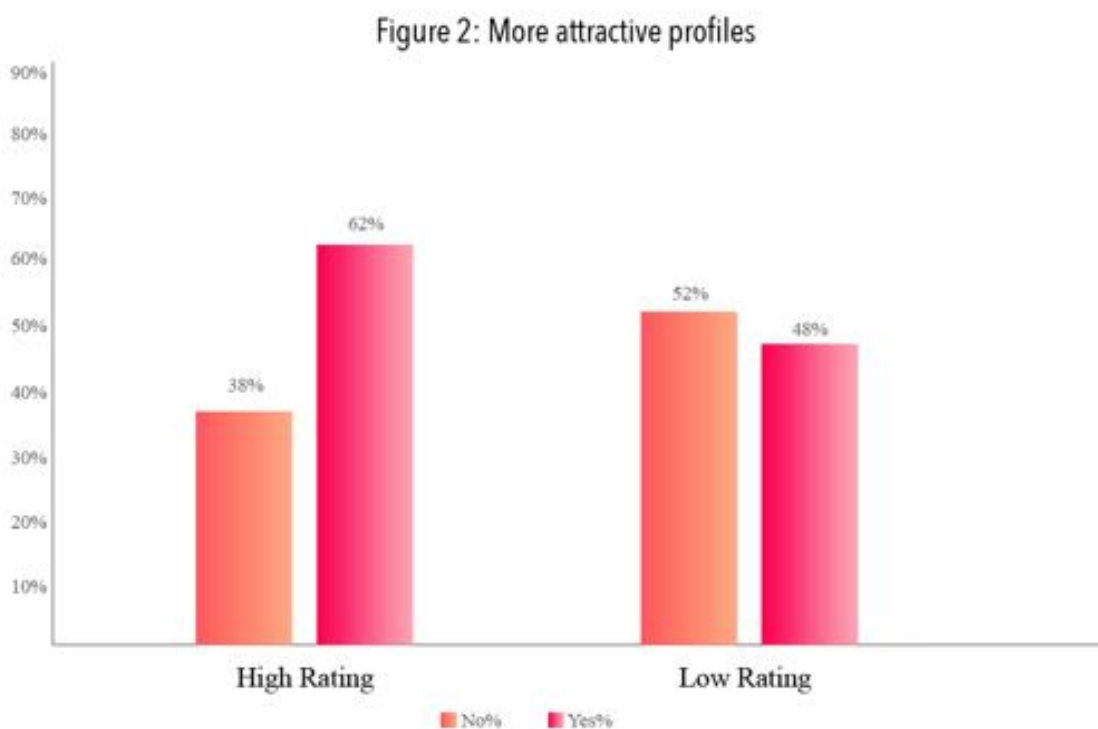

Figure 2: "More attractive profiles": The value of z is 1.6914. The value of p is .04551. Thus, the result is significant at p < .05
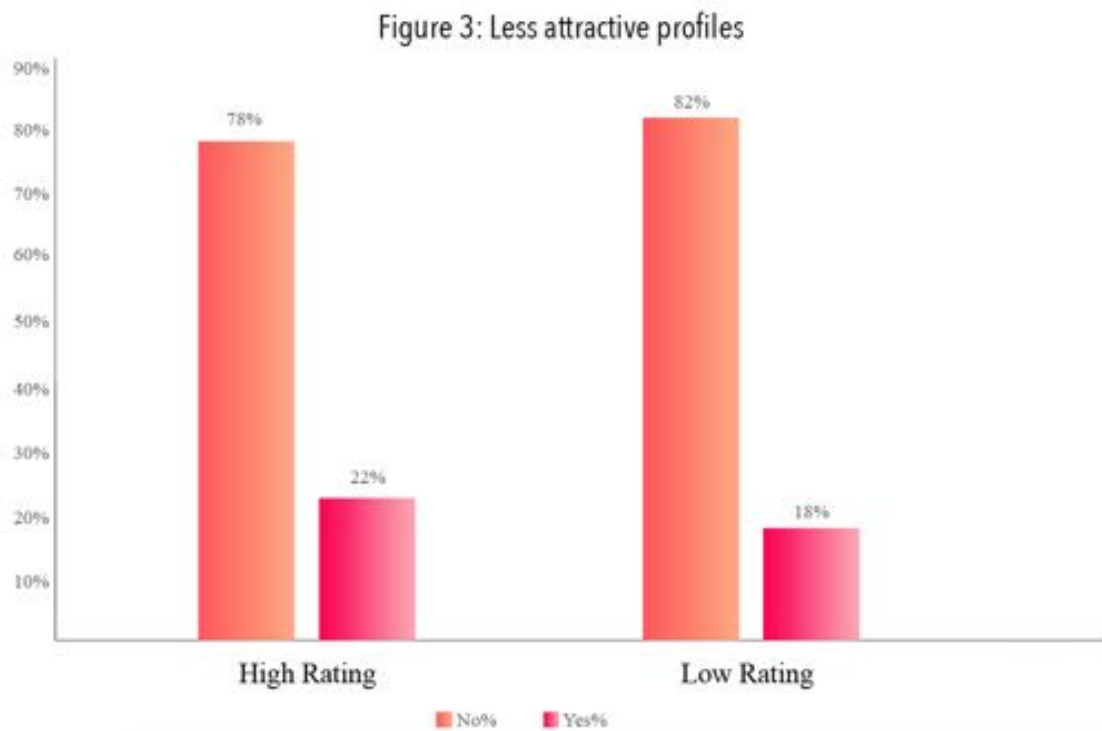
Figure 3: Less attractive profiles

Figure 3:"Less attractive profiles" : The value of z is 0.5508. The value of p is .29116. Thus, the result is not significant at p < .05.

3.3 Concluding Remarks

Study 1 shows support for H1, and thus we have found support to say that ratings in dating profiles can affect the trustworthiness of the person in the profile. However, there is no clear evidence that a higher or lower rating would affect perceived attractiveness, nor that more attractive profiles makes the target more trustworthy. See the summary of the results below:

	Hypothesis	
H1	A profile with a higher rating is perceived as more trustworthy than a profile with a lower rating	Supported
H2	A profile with a higher rating is perceived as more attractive than a profile with lower rating	Not supported
H3	A more attractive profile picture is perceived as more trustworthy than a less attractive profile picture	Not supported

Study 2

This study seeks to find the effects that a 5-star rating system has on intrinsic motivation and choice difficulty in a dating app context.

4.1 Methodology

In this section we motivate the choice of method for Study 2 and explain how we conducted the experiment and collected the data.

4.1.1 Study Design

Similarly to Study 1, this study adopted an experimental design as we sought to test the relationship between variables. In this study, however, the independent variable was the ratings (normally distributed), and the dependent variables were enjoyment, perceived work, and choice difficulty. Furthermore, this experiment followed a *post-test-only control group design* with two groups; one treatment group and one control group (Söderlund, p.43, 2018). The treatment group was exposed to multiple dating profiles where all profiles had a 5-star rating system present. The ratings that were assigned to the profiles ranged from 1 to 5 and were together normally distributed (see figure 4). The control group was exposed to the same profiles without any rating present.

This experiment was intended to test what happens to people's motivation when ratings are present in a dating activity and if people find it less difficult to choose between potential matches. Accordingly, the experiment imitated a real-life Tinder experience and was followed by a questionnaire distributed online in a survey format. Respondents were randomly allocated to one of two groups.

4.1.2 Preparatory work

Study 2's experimental design also required profile pictures, which were retrieved in the same manner as in Study 1. This time, however, we intended to create a real-life simulation for which we decided to include nine profiles for the experience, including the same age and gender possibilities as Study 1. This resulted in the final product of 108 fictitious profiles for the simulation, only nine of which each respondent would be shown. In this experiment the attractiveness of the profiles was irrelevant as we sought to reflect a real life dating experience, so we did not conduct any pre-study to ensure attractiveness.

4.1.3 Survey Design

Similarly to the first study, an introduction of the upcoming simulation was described. Essentially, the survey consisted of two parts: the fictitious Tinder simulation and a questionnaire. Prior to the simulation, respondents were asked about their dating preferences. Similarly to the first study, this was done with the purpose of allocation and to make the experience resemble a real Tinder experience, thus increasing the internal and external validity where you get to choose gender and age preference. Upon doing so, respondents were randomly allocated to one of the two Tinder simulations (treatment or control group). Furthermore, to imitate a real-life Tinder experience even more closely, we used multiple profiles and allowed the respondents to swipe either left or right on each profile before being exposed to the next one. The simulation was followed directly with a questionnaire measuring various dependent variables, and respondents' demographics. The survey included the same control questions as in Study 1.

To verify the reliability of the measures, questions were indexed given a Cronbach alpha of at least 0,7 or more, which is considered an acceptable level of reliability (Söderlund p. 136, 2018). Furthermore, indexing is especially suitable in this context since it is common practice when studying psychological reactions in experiments. This is due to the increased risk of sources of error with psychological reactions in comparison to, for instance, behavioral and physiological reactions (Söderlund p. 136, 2018). With this taken into account, we argue that the reliability of the measures in the studies is sufficient.

4.1.3.1 Dependent variables

Enjoyment and Perceived Work

To measure respondents' motivation after swiping and choosing between the profiles, we adopted the same scales as used in Etkin (2016) to measure motivation. First, we used a 5-item, 10-point Likert scale measuring enjoyment to capture overall intrinsic motivation. The five items that were used were "To what extent do you find the activity *enjoyable?*", "To what extent do you find the activity *boring?*", "To what extent do you find the activity *interesting?*", "To what extent do you find the activity *a waste of time?*","To what extent do you find the activity *fun?*" where the second and fourth were reversed in the analysis. The questions were then indexed with a Cronbach alpha level of 0.90. Second, we measured perceived work by asking respondents: "Do you perceive the dating activity to be *work* or

fun?" with 1 being "definitely work" and 10 "definitely fun". The scale was reversed in the data analysis phase in order for it to make more intuitive sense in the results section. A greater *definitely work* level is considered to be more extrinsically motivated whereas *definitely fun* is intrinsically motivated (Etkin, 2016).

Intrinsically Motivated

For precautionary and exploration reasons, in the event that respondents were *not* intrinsically motivated to perform the dating app simulation but rather extrinsically motivated by the monetary compensation of taking the survey, we included a question after the dating simulation was finished: "Would you have liked to continue?". As (Ryan & Deci, 2000) defines, an intrinsic activity that is one chosen to be performed due to the inherent satisfaction of doing it. The binary answer alternatives were "yes" or "no", and all "no" answers were excluded

Choice Difficulty

Perceived choice difficulty was measured using a 3-item scale including these questions: "Did you find it *difficult* to decide which profiles to like/unlike?", " Did you feel *frustrated* when making the choices?", and "Did you feel *tired* when you finished choosing among the profiles?". A 10-point likert scale was adopted with endpoints "not at all" and "very". With a Cronbach alpha of 0.74, all items were indexed.

4.1.3.2 Manipulation Checks

As this survey was distributed in the same way as the one in Study 1, we included the same control questions to avoid low-quality responses.

4.1.3.3 Stimuli

It should be noted that there is a common critique aimed at experiments based on role-playing because they are only a fictitious mirror of reality (Söderlund, p.179-180, 2018). To counteract this and increase the validity of the study, we have tried to replicate a real-life Tinder experience as closely as possible. Firstly, instead of only presenting the respondents with *one* artificial profile, we manufactured *nine* sequential profiles to imitate a real dating app scenario. Secondly, we allowed the respondents to choose both what gender and what age they would like to be exposed to prior to the experience. Thirdly, we simulated the swiping function.

The simulation investigated whether there is a difference in user *motivation* and *choice difficulty* in a dating-app context depending on the presence or absence of ratings on the profiles. Thus, the simulation consisted of two subject groups; one with ratings and the other without ratings. For the treatment group, the ratings on the profiles followed a 5-star rating system, where the lowest possible score was 1 and the highest 5. In contrast to Study 1, to ensure internal validity and to see whether it was the *presence* of ratings influencing the participants, and *not* the *rating score* itself, each profile in the simulation received a randomized rating score ranging from 1 to 5. The 9 different profiles had normally distributed rating scores. When designing the survey, we distributed the ratings to the profiles in a random order without any ascending or descending effect. For the control group, all profiles were identical to the treatment group except for the absence of the rating stimuli to ensure the cause and effect relationship. Furthermore, the order in which the profiles (including the rating score for the treatment group) were presented to the respondents was identical for all groups.
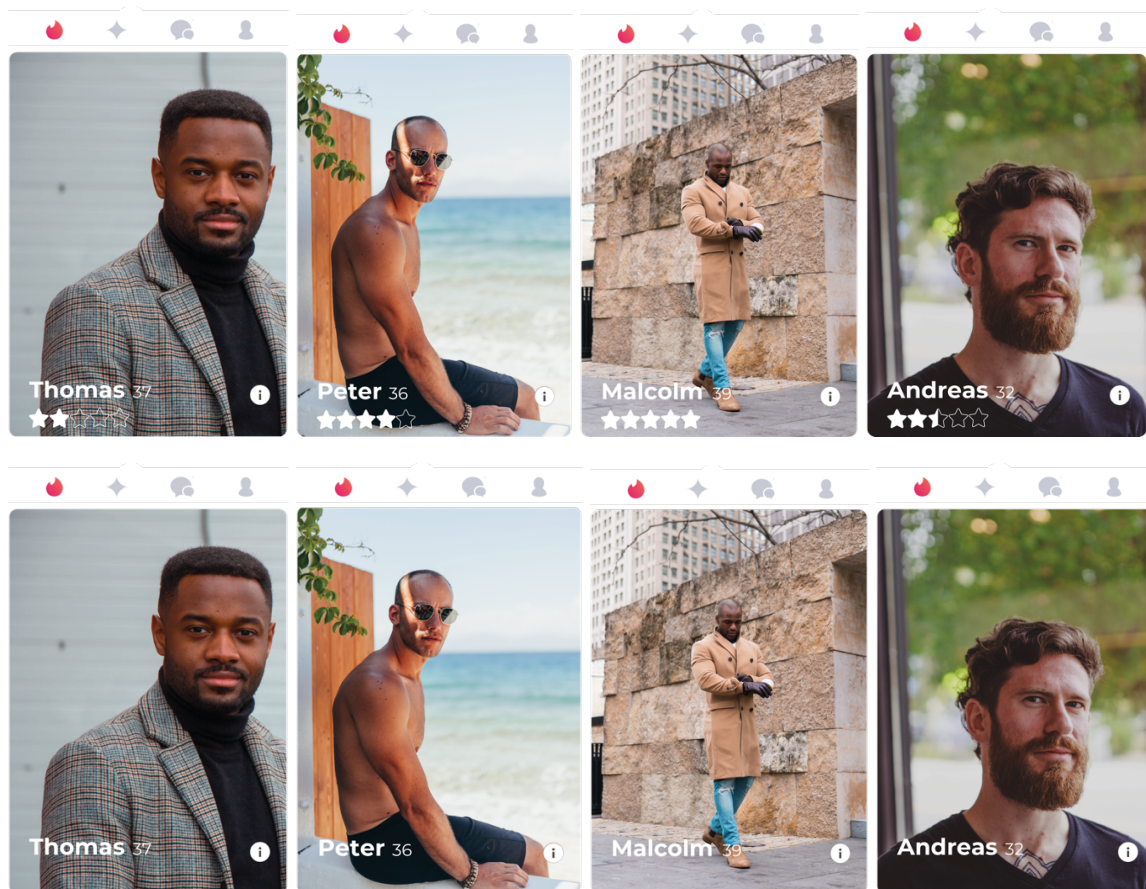

Figure 4: extract from simulation (men, 30-39 with and without ratings)

4.1.4 Sampling

The survey was published March 10th and stayed online until March 24th. During this time a total of (n=367) participated in the simulation and responded to the subsequent survey. After reviewing the data, we excluded the respondents who answered our two control questions wrong, (n=50). The final dataset which the analyses were based on were the remaining 317 answers. Out of these 317 respondents the distribution in gender was as follows: female (n=174), male (n=140) and "other" (n=3). Similarly to the first study, women requested to have profiles of men displayed to a greater extent (n=162) than profiles of women (n=12). Furthermore, men wanted to be displayed profiles of women (n=132) to a greater extent than profiles of men (n=8). The age of the respondents was targeted towards people 18-50 years old since the created simulations only included profiles with ages in the given range and because there is an age limit of 18 on most dating apps available. The mean age among participants was 33,76 (SD=9,229). As intended, the distribution between dating app users (n=160) and non-users (n=157) was equally weighted since we wanted to examine the psychological effects of ratings, and that is not limited to people who use dating apps. The concept and implications of "swiping" were clearly explained in the introduction of the survey, thus eliminating potential sources of error from non-users.

4.2 Results

In this part we test our hypotheses (H3-H6) connected to the 5-star rating system's effects on intrinsic motivation and choice difficulty.

Enjoyment

Table 5

Independent samples t-test	Ratings		No Ratings			
	Mean	SD	Mean	SD	Sig	*t*
Enjoyment	5.78	2.13	6.04	2.18	0.14	-1.07

1=not enjoyable, 10=very enjoyable   n=317

First, we compared the mean levels for two groups on a general level: ratings and non-ratings. This was done with a one-sided independent t-test. As presented in Table 1, the level of enjoyment was lower for the group exposed to ratings in the profiles (M_{rating}=5.78, SD=2.13) in comparison to those with no rating ($M_{no,rating}$=6.04, SD=2.18). However, the result was not statistically significant, (318)= -1.07, p=0.14, n=317, so **H4 was rejected**.

Nonetheless, the interaction showed a tendency in line with our hypothesis that was worth investigating further.

Enjoyment for intrinsically motivated

Independent samples t-test	Ratings		No Ratings			
	Mean	SD	Mean	SD	Sig	t
Enjoyment$_{motivated}$	6.78	1.45	7.08	1.46	0.10	-1.29
Enjoyment$_{motivated,women}$	6.40	1.34	6.94	1.46	0.049	-1.68
Enjoyment$_{motivated,men}$	7.20	1.46	7.20	1.45	0.49	-0.035

1=not enjoyable, 10=very enjoyable   n=162, women n=78, men n=84

As mentioned earlier, this study aimed at investigating whether ratings undermine the intrinsic motivation in the enjoyable activity of online dating. Thus, for exploratory reasons we only included respondents who were *intrinsically motivated,* in other words respondents who answered that they would have liked to continue with the activity. Again, an independent one-sided t-test was conducted. This yielded significant results ($t(160)=-1.29$, $p=.10$, $n=162$) where the treatment group found it less enjoyable ($M_{rating}=6.78$, $SD=1.45$) and the control group more enjoyable ($M_{no\_rating}=7.08$, $SD=1.46$). This was in line with previous research by Etkins (2016) where an intrinsically motivated activity is undermined by measurements. We also set out to investigate differences between genders. Interestingly, when searching for mean differences, only women indicated significant results ($t(76)=1.68$, $p=.049$, $n=78$) of lower enjoyment when exposed to ratings ($M_{rating,motivated,women}=6.40$, $SD=1.34$) in comparison to the no-rating group ($M_{no,rating,motivated,women}=6.94$, $SD=1.46$), as opposed to the men who were indifferent ($M_{rating,motivated,male}=7.20$, $SD=1.46$; $M_{no,rating,motivated,male}=7.20$, $SD=1.45$) conditions; $t(82)= -0.035$, $p=0.49$, $n=84$).

Perceived Work

Table 7

Independent samples t-test	Ratings		No Ratings			
	Mean	SD	Mean	SD	Sig	*t*
Percieved work	4.72	2.67	4.00	2.55	0.008	-2.45
Percieved work_women	5.07	2.91	4.12	2.65	0.013	-2.24
Percieved work_men	4.18	2.18	3.83	2.46	0.19	-0.87

1=definitely fun, 10=definitely work   n=317, women n=174, men n=140

As for perceived work, the one-sided independent t-test indicated that the perceived work was indeed higher when exposed to a simulation of profiles with ratings (M_{rating}=4.72, SD=2.67) than to those with no ratings: ($M_{no,rating}$=4.0, SD=2.55) conditions; t(315)= -2.45, p=0.008, n= 317. Thus, **we accept H5**.

Choice Difficulty

Table 8

Independent samples t-test	Ratings		No Ratings			
	Mean	SD	Mean	SD	Sig	*t*
Choice difficulty	2.74	1.70	2.70	1.75	0.42	0.21
Choice difficulty_women	2.91	1.88	2.55	1.72	0.09	1.34
Choice difficulty_men	2.55	1.44	2.90	1.79	0.10	-1.27

1=not difficult, 10=very difficult   n=317, women n=174, men n=140

For the choice difficulty, no significant (t(315)=0.21, p=.42, n=317) mean difference was found when comparing the groups on a general level (M_{rating}=2.74 SD=1.7, M_{no_rating}=2.7, SD=1.75). **Thus**, **we reject hypothesis H6.**

As before, we also investigated potential explanatory factors to our results and therefore investigated the effects on the different genders respectively. Surprisingly, there was a significant opposite-gender effect uncovered where women found it more difficult to choose between the profiles they wanted to match with when ratings were present ($M_{women,rating}$=2.91, SD=1.88; $M_{women,no,rating}$=2.55, SD=1.72) conditions; t(172)=1.34, p=.09, N=174, whereas men in contrast found it easier ($M_{men,rating}$=2.55, SD=1.44, $M_{men,no,rating}$=2.90; SD=1.79) conditions; t(135)= -1.27, p=.10, n=140.

4.3 Concluding Remarks

Study 2 suggested that an online dating activity seems more like work and is less fun with the presence of ratings. However, not enough support was found for decreased enjoyment as a consequence of ratings, thus we reject the hypothesis. Likewise, the study found no support for choice difficulty decreasing in the presence of ratings.

	Hypothesis	
H4	Ratings in dating apps reduce the enjoyment of performing the activity	Not supported
H5	The perceived work of performing the activity is greater with the presence of ratings	Supported
H6	Choice difficulty decreases with the presence of ratings in a dating-app context	Not supported

Discussion

In this final section we will reconnect to our purpose and discuss the results from both studies in detail. We proceed with presenting implications for managers and practitioners as well as suggesting limitations and implications for future research.

5.1 Studies 1 & 2

Do ratings in dating make it more trustworthy but less fun? Our two studies combined unveil that a 5-star rating system can indeed increase trustworthiness of a person in a dating profile. While intrinsic motivation was not affected for the average person, it was undermined for users who did indeed enjoy swiping, and it also increased the perceived work of swiping. We hereby add to the field of research of positive effects of reputation systems by bringing it to the context of dating. We also show that exposure to measurement of *others* has a tendency to undermine intrinsic motivation, which adds to previous research investigating the effects on motivation when receiving measurement.

Piecing it all together, a theme emerges where the results indicate that effects of ratings in a dating app context were more significant for women than for men in the majority of the tests. A probable cause for this phenomenon is the difference in risk associated with dating apps for the respective genders. It also relates to how attentive the genders are while swiping, being that women are more selective than men choosing between profiles. As we hypothesized, the implementation of ratings in a dating-app context is associated with both positive and negative outcomes for the user. Our assumption was that the use of dating apps was something people do for fun. However, as the data reveals, not all find it to be an enjoyable activity, or at least not our simulation. Therefore, it seems as if the negative effects on intrinsic motivation found for those who perceived the activity to be enjoyable do not carry as much importance as for the positive effects found on perceived trustworthiness.

5.1.1 Trustworthiness & Attractiveness
H1
Will a profile with a higher rating be perceived as more trustworthy than a profile with a lower rating? As can be derived from the results, higher ratings resulted in greater trustworthiness in the counterpart. This was in line with previous studies of ratings (Dahlén and Thorbjornsen, 2021) and with information seeking in order to reduce uncertainty of

interactions (Berger & Cabrese, 1975). However, after further exploration the effect could mostly be attributed to women. Once again, an explanation for this can be ascribed to the different risk implications for men and women respectively on dating apps (Lopes & Vogel, 2017; Whitty, 2008). As mentioned in section 1.2, women experience a greater degree of offensive behavior on dating apps than men and can expect a higher risk of physical harm. In contrast, the risk for men on dating apps shown in previous research relates to being disappointed regarding the appearance of a potential date as women lie about their physical appearance to a greater extent. One can thus argue that the gender-specific risk and level of uncertainty is greater on dating apps for women and entails bigger consequences if the match is either lying or being deceitful. This could be a contributing factor to *why* higher ratings evoked greater trust for women than for men. As previous studies have illustrated (Berger & Calabrese, 1975), the ambiguity of first encounters is what drives information seeking to make future interaction more predictable. Thus, if the level of ambiguity is in general larger for women on dating apps, it is reasonable to assume that this is a contributing factor for why ratings invoke higher degrees of trust for women in a dating app context.

H2

Study 1 showed no significance that high ratings would increase the perceived attractiveness of the person in the profile. Thus, a 5-star rating system where targets are rated high (or low) has no further implication on the targets' attractiveness in a dating app context. One explanation could be simply that attractiveness overrules any indicator about behavior. Previous research does emphasize the importance of physical attractiveness when choosing among potential partners (Shaffer, 1997; Walster et al., 1966). Our results from measuring how respondents chose to swipe (left or right) further demonstrated that, regardless of rating, people want to date attractive people. We previously reasoned that behavioral aspects represented by ratings could affect perceived attractiveness by virtue of the halo effect. That is, people who have good personalities are viewed as more attractive. The research supporting our hypothesis (Paunonen, 2006) studied the effect of biographies in online dating to illustrate people's personality and change the perceived attractiveness of people in the profiles, but our study used ratings. Therefore, the 5-star rating system does not seem to encompass sufficient conviction of personality traits for daters to take this into consideration in dating apps. In other words, rating systems do not illustrate enough information about the profile users' personalities to make them seem less or more attractive. This holds true for both men and women on dating apps.

H3.

The results show an opposite effect of what we hypothesised: people perceive a target to be less trustworthy if they are more attractive. Furthermore, this was especially true for women. This was surprising, as the halo effect suggests that attractive people are assumed to have positive personality traits. One explanation for this counterintuitive result could be that people believe that the attractive profile is not true to life and that the positive features are exaggerated. Research by Lo et al. (2013) found that attractive profiles seem less authentic than unattractive profiles and thus conclude that attractive profiles are less trusted in an online dating setting. Moreover, it is probable that the less attractive profiles in our study are considered more accurate and authentic and are thus trusted to a greater extent than the more attractive profiles. The difference in significance between the genders is an interesting finding as women tend to misrepresent their appearance to a greater extent than men on dating apps, which could imply a "takes one to know one" effect. In other words, knowing that you often enhance your own appearance in dating apps, you also believe that others do the same. Another explanation could be that women are more considerate and attentive when they are choosing whom to initiate contact with. This was found in Tyson et al. (2016) after analyzing half a million Tinder users' behavior, where the respective genders displayed very different trends in terms of selection between profiles when swiping. Women tended to be more selective of the profiles they liked while leaving men with fewer possible matches. Men tended to first like far more profiles without much consideration and *then* filter through their matches post match, meaning that they are not as attentive while swiping. This is a plausible explanation for the significance found in women but not for men.

5.1.2 Intrinsic motivation & Choice Difficulty

H4 & H5

Revisiting the scope of the second study, the theoretical framework of it was based on quantifying one's own performance and *receiving* positive or negative feedback. These aspects have only been investigated when the subject her/himself is the one being measured or given feedback to, whereas our study seeks to find if the underlying theoretical framework also holds true when being exposed to social approval feedback on *other* people. Our results do confirm this. We further confirm previous findings that measurements in enjoyable activities such as dating apps undermine intrinsic motivation, making it seem more like work and less fun (Etkin, 2016). Further, that the effects are not only limited to measuring oneself,

but also being exposed to social approval feedback on other people given by others. However, the ratings' effect on enjoyment, adopted from Etkin's framework, was only significant for intrinsically motivated respondents. Investigating even further between gender, only intrinsically motivated women experienced less enjoyment as a consequence of ratings in our context.

The insignificant results for men in regards to decreased enjoyment and thus intrinsic motivation can partly be explained with the help of a previous study conducted by Deci et al. (1973). The ratings in our simulation were mixed social approval feedback fictitiously given by other users ranging from scores of 1 star up to 5, so as to include both positive *and* negative feedback in our experiment. Deci found that verbal feedback (negative and positive) had negative effects on intrinsic motivation for women which is in line with our findings, but *only* negative feedback resulted in decreased intrinsic motivation for men whereas positive feedback had the opposite effect. The interpretation for Deci's finding was that *positive* feedback will strengthen the self-determination for men and thereby increase intrinsic motivation, whereas the *negative* feedback has the inverse effect. In this context where someone else is rated, men seem to also attribute the social approval feedback of others to their own self-determination. Furthermore, the reason for women finding the dating app activity less intrinsically motivating is that the controlling aspect tied to the *positive feedback* is more salient for women than for men, meaning that they do not feel as much self-determination when exposed to any feedback. It seems like this also holds true for social approval given by other people in profiles in the form of ratings. Therefore, a probable cause for the insignificant enjoyment result for men and the significant result for women was due to the nature of our stimuli given to the treatment groups, being both positive and negative feedback.

H6
Hypothesis 6 resulted in insignificant results on the entire group. Interestingly, as we explored further, the choice difficulty also showed different implications for men and women. Men found it easier to choose between profiles with ratings, which is in line with findings in Etkin (2016), being that measurement increases efficiency in performing activities and women experience an inverse effect. The inverse relationship can also be ascribed to Tyson et al's (2016) findings of dating profile selection trends between genders, namely that women were found to be more selective choosing between their matches. To them, additional

measurement of ratings might have contradicted their own view of the perceived authenticity of the profile picture, which was also found to be an important factor for their evaluation of mens' attractiveness (Fiore et al., 2008), thus making it more difficult to make decisions. In contrast, a plausible explanation for the decreased choice difficulty for the men in our study can be assigned to the fact that men were less attentive while swiping and thus found ratings to be a reliable filtering mechanism.

5.2 Implications

5.2.1 Managerial Implications

The findings in our thesis can be useful for dating app management and developers. Today, the anonymity pertaining to online dating as well as the power dynamics can be troublesome for many dating app users. The findings of this research paper suggest that rating systems can facilitate judgements of target trustworthiness and that with practical implementation, dating apps can overcome some of the issues they have today. The result of our research is important for management to consider as it can be utilized to compete with other dating apps that manage these issues in other ways (e.g. Bumble). Moreover, we suggest that implementing a 5-star rating system where users can rate their dates is a viable option to eliminate unserious and fraudulent users.

However, we also conclude that ratings can have negative effects on users as well, which could be worth considering before implementation. The question management and dating app developers will have to review is the importance of having intrinsically motivated users.

5.2.2 Implications for Practitioners

A good rating will make the profile seem more trustworthy. This favors users who usually behave in a decent manner on dating apps. Further, men with good ratings had the biggest differences in perceived trustworthiness depending on good or bad ratings, meaning that the implementation of ratings would benefit the majority of women on dating apps along with highly-rated men. The attractiveness of a user's profile was not significantly affected by higher or lower ratings. Although counterintuitive, less attractive men were seen as more trustworthy meaning that men should represent themselves as less attractive on dating apps *if* the prime goal is to be perceived as more trustworthy.

5.2.3 Implications for future research

Although this study found evidence that ratings in dating apps can increase target trustworthiness and thus ratings provide additional information for users that guide them in their choices, we encourage future research efforts that strengthen the relationship between ratings and trust. Evidently, the measure that we used for trust in this experiment proved to be somewhat inadequate as the internal consistency of the 3-item battery was too low. We therefore suggest future researchers find other alternatives to the IMI scale to measure trustworthiness.

This research paper also indicates a relationship between ratings and intrinsic motivation. However, the study is limited to one brief Tinder session and as motivation is arguably attained over a longer period of time, we would like to see future studies that measure the effects of ratings on intrinsic motivation in a longer experiment. In addition, we solely focused on the motivation of the user when being exposed to rated profiles. To get a more complete perspective, we encourage future studies to investigate the effects of being rated themselves.

.

References

AppMagic. (January 29, 2021). *Most popular dating apps worldwide as of January 2021, by number of monthly downloads (in 1,000s)* [Graph]. In Statista. Retrieved April 27, 2021, from https://www-statista-com.ez.hhs.se/statistics/1200234/most-popular-dating-apps-worldwide-by-number-of-downloads/

Bente, G., Baptist, O., & Leuschner, H. (2012). To buy or not to buy: influence of seller photos and reputation on buyer trust and purchase behavior. *International Journal of Human‑Computer Studies*, 70(1), 1–13. https://doi.org/10.1016/j.ijhcs.2011.08.005

Berger, C. R., & Calabrese, R. J. (1975). Some explorations in initial interaction and beyond: Toward a developmental theory of interpersonal communication. *Human Communication Research*, 1,99-112.

Dahlén, M & Thorbjornsen, H. (2021). Sifferdjur: hur siffror påverkar hur vi lever och beter oss, Volante förlag.

Dating Services - Worldwide. (n.d.). Retrieved April 27, 2021, from https://www-statista-com.ez.hhs.se/outlook/dmo/eservices/dating-services/worldwide

Deci, E. L. (1971). Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology*, 18(1), 105–115. https://doi.org/10.1037/h0030644

Deci, E. L., & Cascio, W. F. & Krussel, J. (1973). Sex differences, postive feedback and intrinsic motivation. *Rochester University., N. Y. Management Research Center*. Paper presented at the Eastern Psychological Association Convention (Washington, D.C. May 3-5, 1973)

Deci, E. L., & Ryan, R. M. (1985). Intrinsic motivation and self-determination in human behavior. New York: *Plenum.*

Deci, E. L., & Ryan, R. M. (2008). Self-determination theory: A macrotheory of human motivation, development, and health. *Canadian Psychology/Psychologie canadienne,* 49(3), 182–185. https://doi.org/10.1037/a0012801

Drouin, M., Miller, D., Wehle, S. M., & Hernandez, E. (2016). Why do people lie online?"Because everyone lies on the internet". *Computers in Human Behavior*, 64, 134-142.

Etkin, J. (2016). The Hidden Cost of Personal Quantification, *Journal of Consumer Research,* Volume 42, Issue 6, Pages 967–984, https://doi.org/10.1093/jcr/ucv095

Filieri, R. (2015). What makes online reviews helpful? A diagnosticity-adoption framework to explain informational and normative influences in e-WOM. *Journal of Business Research*, *68*(6), 1261-1270.

Fiore, A & Taylor, L & Mendelsohn, G. & Hearst, M.. (2008). Assessing Attractiveness in Online Dating Profiles. *Conference on Human Factors in Computing Systems - Proceedings.* 797-806. 10.1145/1357054.1357181.

Fletcher, E. (10 Febuary 2021). Romance scams take record dollars in 2020. *Federal Trade Commission.* Retrieved from https://www.ftc.gov/news-events/blogs/data-spotlight/2021/02/romance-scams-take-record-dollars-2020

Forgas, J. P., & Laham, S. M. (2016). Halo effects. Cognitive illusions: Intriguing phenomena in judgement, thinking and memory, 276-290.

Frost, J. H., Chance, Z., Norton, M. I., & Ariely, D. (2008). People are experience goods: Improving online dating with virtual dates. Journal of Interactive Marketing, 22(1), 51-61.

Gulas, C. S., & McKeage, K. (2000). Extending social comparison: An examination of the unintended consequences of idealized advertising imagery. *Journal of advertising*, *29*(2), 17-28.

Kellermann, Kathy & Reynolds, Rodney. (1990). When Ignorance Is Bliss The Role of Motivation to Reduce Uncertainty in Uncertainty Reduction Theory. *Human Communication Research - HUM* COMMUN RES. 17. 5-75. 10.1111/j.1468-2958.1990.tb00226.x.

Kramer, M. W. (1999). Motivation to reduce uncertainty: A reconceptualization of uncertainty reduction theory. *Management Communication Quarterly*, 13(2), 305-316.

Liesel L. Sharabi (2020): Online dating profiles, first-date interactions, and the enhancement of communication satisfaction and desires for future interaction, *Communication Monographs*, DOI: 10.1080/03637751.2020.1766094

Liu, Y., Nie, L., & Li, L. (2016). Homogeneity, trust, and reciprocity: Three keys to the sustainable hospitality exchange of couchsurfing. *Tourism Analysis*, 21(2-3), 145-157.

Lopes, M. R., & Vogel, C. (2017, July). Gender bias on Tinder: Transforming an exploratory qualitative survey into statistical data for contextualized interpretation. *In International Symposium on Qualitative Research* (pp. 225-236). Springer, Cham.

Lo, S. K., Hsieh, A. Y., & Chiu, Y. P. (2013). Contradictory deceptive behavior in online dating. Computers in Human Behavior, 29(4), 1755-1762.

Malinen, S., & Ojala, J. (2011). Perceptions of trust between online auction consumers. *International Journal of Web Portals* (IJWP), 3(4), 15-26.

McKnight, D. H., & Chervany, N. L. (2001). Trust and distrust definitions: One bite at a time. In *Trust in Cyber-societies* (pp. 27-54). Springer, Berlin, Heidelberg.

Norcie, G., De Cristofaro, E., & Bellotti, V. (2013). Bootstrapping trust in online dating: Social verification of online dating profiles. In *International Conference on Financial Cryptography and Data Security* (pp. 149-163). Springer, Berlin, Heidelberg.

Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology, 45*(4), 867–872. https://doi.org/10.1016/j.jesp.2009.03.009

Paunonen, S. V. (2006). You are honest, therefore I like you and find you attractive. *Journal of Research in Personality*, *40*(3), 237-249.

Rosenfeld, M. J., Thomas, R. J., & Hausen, S. (2019). Disintermediating your friends: How online dating in the United States displaces other ways of meeting. *Proceedings of the National Academy of Sciences*, *116*(36), 17753-17758.

Roumeiu, G. (12 November 2011). Love, Lies and What They Learned. *New York Times*. Retrieved from https://www.nytimes.com/2011/11/13/fashion/online-dating-as-scientific-research.html

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary educational psychology*, *25*(1), 54-67.

Sansone, C., & Harackiewicz, J. M. (Eds.). (2000). Intrinsic and extrinsic motivation: The search for optimal motivation and performance. *Academic Press*.

Shaffer, D. R., & Bazzini, D. G. (1997). What Do You Look for in a Prospective Date? Reexamining the Preferences of Men and Women Who Differ in Self-Monitoring Propensities. *Personality and Social Psychology Bulletin*, *23*(6), 605–616. https://doi.org/10.1177/0146167297236004

Söderlund, M. *Experiments in Marketing*. First edition. Lund: Studentlitteratur, 2018. Print.

Söderlund, M. Mätningar och mått : i marknadsundersökarens värld . 1. uppl. Malmö: Liber ekonomi, 2005. Print.

TV4 Nyheterna. (25 April 2021). *Tinder-bedragaren blåste Mari på och andra på 2 miljoner kronor* [Video]. TV4play. Retrieved from https://www.tv4play.se/program/nyheterna/tinder-bedragaren-bl%C3%A5ste-mari-och-andra-p%C3%A5-2-miljoner-kronor/13336411

Tyson, G., Perta, V. C., Haddadi, H., & Seto, M. C. (2016). A first look at user activity on tinder. In *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)* (pp. 461-466). IEEE.

Walster, E., Aronson, V., Abrahams, D., & Rottman, L. (1966). Importance of physical attractiveness in dating behavior. *Journal of Personality and Social Psychology, 4*(5), 508–516. https://doi.org/10.1037/h0021188

Wang, P., Zheng, H., Chen, D., & Ding, L. (2015). Exploring the critical factors influencing online lending intentions. *Financial Innovation*, 1(1), 1-11.

Whitty, M. T. (2008). Revealing the "real" me, searching for the "actual" you: Presentations of self on an internet dating site. *Computers in Human Behavior*, 24(4), 1707–1723. doi:10.1016/j.chb.2007.07.002

Appendices

6.1 Appendix 1 - Survey Study 1

6.1.1 Pre-study

This question was asked for all profiles in 6.1.2

6.1.2 Survey flow (limited to men 18-29)

6.1.3 Stimuli

Women 18-29

Men 18-29

Women 30-39

Men 30-39

Women 40-50

Men 40-50

6.1.4 Questionnaire

Introduction

Intro_sim

Du kommer nu att delta i en kort simulerad tinder-session där du ska välja om du vill matcha med personen eller inte. Profilen har en rating som är baserad på personens beteende. Titta noga på profilen som presenteras och gör sedan ditt val.

Ett grönt hjärta = jag vill matcha.
Ett rött "X" = jag vill inte matcha.

Sorting questions

Sex_choice ★

Vilket kön skulle du vilja se på profilen?

○ Man
○ Kvinna

Age_choice ★

I vilket åldersspann?

○ 18-29
○ 30-39
○ 40-50

Questions and manipulation checks

Attractivness iQ ★

	Inte alls (1)	2	3	4	5	6	7	8	9	Väldigt (10)
Hur attraktiv tyckte du att personen i profilen var?	○	○	○	○	○	○	○	○	○	○

Control_1 iQ ★

	1	2	3	4	5	6	7	8	9	10
Var vänlig välj svarsalternativ '7'	○	○	○	○	○	○	○	○	○	○

Relatedness iQ ★

Hur väl stämmer nedanstående påståenden överens med din uppfattning?

	Stämmer inte alls (1)	2	3	4	5	6	7	8	9	Stämmer helt (10)
Jag känner att jag verkligen skulle kunna lita på den här personen.	○	○	○	○	○	○	○	○	○	○
Jag föredrar att inte interagera med den här personen i framtiden	○	○	○	○	○	○	○	○	○	○
Jag känner inte att jag skulle kunna lita på den här personen.	○	○	○	○	○	○	○	○	○	○

Kön

Min könstillhörighet

○ Kvinna
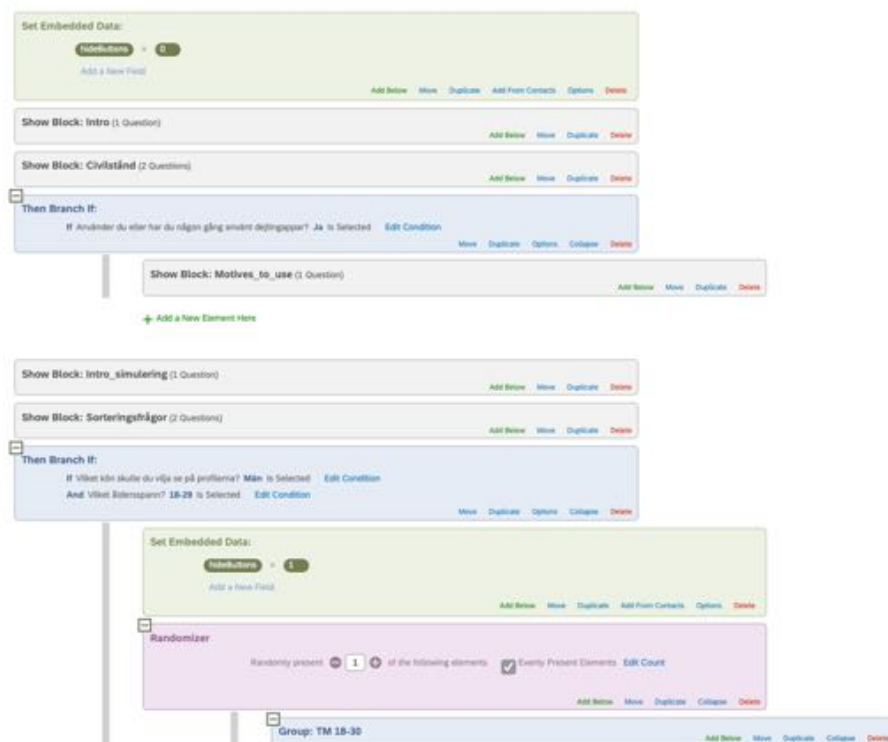
○ Man

○ Annat

*

Ålder

Ålder

*

Kontroll_2

Vad handlade den här studien om?

○ Köpbeteende

○ Dejtingappar

○ Filmrecensioner

*

6.2 Appendix 2 - Survey Study 2

6.2.1 Survey flow (limited to men 18-29)

Group: TM 18-30   Add Below Move Duplicate Collapse Delete

Show Block: 8. 18-30 (1 Question)   Add Below Move Duplicate Delete

Show Block: 4. 18-30 (1 Question)   Add Below Move Duplicate Delete

Show Block: 2. TM 18-30 (1 Question)   Add Below Move Duplicate Delete

Show Block: 7. 18-30 (1 Question)   Add Below Move Duplicate Delete

Show Block: 5. 18-30 (1 Question)   Add Below Move Duplicate Delete

Show Block: 1. TM 18-30 (1 Question)   Add Below Move Duplicate Delete

Show Block: 9. 18-30 (1 Question)   Add Below Move Duplicate Delete

Show Block: 6. TM 18-30 (1 Question)   Add Below Move Duplicate Delete

Show Block: 3. TM 18-30 (1 Question)   Add Below Move Duplicate Delete

+ Add a New Element Here

Group: TRM 18-30   Add Below Move Duplicate Collapse Delete

Show Block: 8. TRM 18-30 (1 Question)   Add Below Move Duplicate Delete

Show Block: 4. TRM 18-30 (1 Question)   Add Below Move Duplicate Delete

Show Block: 2. TRM 18-30 (1 Question)   Add Below Move Duplicate Delete

Show Block: 7. TRM 18-30 (1 Question)   Add Below Move Duplicate Delete

6.2.2 Stimuli in the Randomized Order

Women 18-29 with ratings

Women 18-29 without ratings

Men 18-29 with ratings

Men 18-29 without ratings

Women 30-39 with ratings

Women 30-39 without ratings

Men 30-39 with ratings

Men 30-39 without ratings

Women 40-50 with ratings

Women 40-50 without ratings

Men 40-50 with ratings

Men 40-50 without ratings

6.2.3 Questionnaire

Introduction

Intro_sim

Du kommer nu att delta i en simulerad tinder-session, där du ska välja om du vill matcha med några personer eller inte. Titta noga på profilerna som presenteras och gör sedan dina val.

Ett grönt hjärta = jag vill matcha.
Ett rött "X" = jag vill inte matcha.

Sorting questions

Sex_choice ★

Vilket kön skulle du vilja se på profilen?

○ Man
○ Kvinna

Age_choice ★

I vilket åldersspann?

○ 18-29
○ 30-39
○ 40-50

Questions and manipulation checks

Väl_du_fortsätta ★

Hade du velat fortsätta?

○ Ja
○ Nej

Enjoyment iQ ★

I vilken utsträckning finner du aktiviteten (simuleringen) som...

	Inte alls (1)	2	3	4	5	6	7	8	9	Väldigt (10)
Ett nöje	○	○	○	○	○	○	○	○	○	○
Tråkig	○	○	○	○	○	○	○	○	○	○
Intressant	○	○	○	○	○	○	○	○	○	○
Slöseri med tid	○	○	○	○	○	○	○	○	○	○
Kul	○	○	○	○	○	○	○	○	○	○

Choice_difficulty iQ ★

	Inte alls (1)	2	3	4	5	6	7	8	9	Väldigt (10)
Upplevde du det svårt att välja vilka profiler att gilla/ogilla?	○	○	○	○	○	○	○	○	○	○
Kände du dig frustrerad när du valde profiler att gilla/ogilla?	○	○	○	○	○	○	○	○	○	○
Kände du dig trött när du var klar med dina val om profilerna?	○	○	○	○	○	○	○	○	○	○

Perceived work iQ ★

	Definitivt som ett jobb (1)	2	3	4	5	6	7	8	9	Definitivt som något kul (10)
Tänk dig att tinder sessionen var verklig, kändes det som ett jobb eller som något kul?	○	○	○	○	○	○	○	○	○	○

Kontroll_1 iQ ★

	1	2	3	4	5	6	7	8	9	10
Var vänlig välj svarsalternativ '7'	○	○	○	○	○	○	○	○	○	○

Kön

Min könstillhörighet

○ Kvinna

○ Man

○ Annat

Ålder

Ålder

Kontroll_2

Vad handlade den här studien om?

○ Köpbeteende

○ Dejtingappar

○ Filmrecensioner