Measuring the “Why” of Interaction: Development and Validation of the User Motivation Inventory (UMI)

Florian Brühlmann, Beat Vollenwyder, Klaus Opwis, Elisa D. Mekler

Center for Cognitive Psychology and Methodology, Department of Psychology, University of Basel
{florian.bruehlmann, beat.vollenwyder, klaus.opwis, elisa.mekler}@unibas.ch

ABSTRACT

Motivation is a fundamental concept in understanding people’s experiences and behavior. Yet, motivation to engage with an interactive system has received only limited attention in HCI. We report the development and validation of the User Motivation Inventory (UMI). The UMI is an 18-item multidimensional measure of motivation, rooted in self-determination theory (SDT). It is designed to measure intrinsic motivation, integrated, identified, introjected, and external regulation, as well as amotivation. Results of two studies (total $N = 941$) confirm the six-factor structure of the UMI with high reliability, as well as convergent and discriminant validity of each subscale. Relationships with core concepts such as need satisfaction, vitality, and usability were studied. Additionally, the UMI was found to detect differences in motivation for people who consider abandoning a technology compared to those who do not question their use. The central role of motivation in users’ behavior and experience is discussed.

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous; J.4. Social and Behavioral Sciences: Psychology

Author Keywords

Motivation; user experience; technology use; self-determination theory; scale development; usability

Open data policy

All data, study materials, and main analysis code used for the development and validation of the UMI are available at https://osf.io/m3fbk/

INTRODUCTION

Motivation is a fundamental concept in our lives, as it drives all intentional behavior. This also holds true for technology use since the motivation to engage with a given interactive system is at the core of the formation of user experience [22]. The reasons why people engage with a technology affect how users perceive product qualities, what qualities are important, and how they affect the users’ experience. For instance, the pursuit of instrumental, task-directed goals renders usability problems more salient, which may in turn negatively influence users’ experience and retrospective product evaluation [20]. In contrast, non-utilitarian qualities are preferred when pursuing more exploratory [56] or experientially motivated behaviors involving technology use (e.g., to have fun, [20, 41]). Product beauty was found to be more important for leisurely rather than work-related technology use [52]. Similarly, users’ experiences with technology varied when pursued for eudaimonic (e.g., developing one’s personal potential) or for hedonic (e.g., pleasure) reasons [31]. In fact, depending on users’ motivation, the same technology-supported activities might be experienced very differently, say, when playing digital games for leisurely or for professional purposes [11].

Self-determination theory (SDT), an influential theory of human motivation [10, 43], differentiates the *what* (i.e., goal content) and the *why*, that is, the regulatory processes underlying goal pursuit [8]. According to SDT, people can satisfy the innate psychological needs for autonomy, competence, and relatedness [8] through a variety of behaviors. However, the quality of people’s behavior, the extent of need satisfaction, as well as the consequences on well-being depend on the motivational regulations underlying these behaviors [8]. Surprisingly, while need satisfaction – itself a key concept of SDT – was repeatedly found to be a defining characteristic of positive user experience (e.g., [19, 31]) and considered core to the understanding of what makes interaction good [22], the regulatory processes posited by organismic integration theory (OIT), a sub-theory of SDT, have so far received scant attention within HCI research – which may in part be due to the lack of a suitable measuring instrument. Distinguishing between different regulations might provide a more nuanced understanding of positive (and negative) experiences and their effects on need satisfaction (e.g., [58]). Additionally, motivation is a fundamental element to consider in studies concerning the effects of technology on well-being [31]. Hence, a multidimensional scale of motivation could extend existing models of user experience [27] and complement qualitative approaches to the "why" of interaction by providing a reliable tool that can be used to find generalizable and replicable results. A questionnaire for motivation can be applied to test theories and hypotheses and establish causal relationships in randomized controlled experiments.

In the present work, we describe the development and validation of the User Motivation Inventory (UMI). Our contribution is three-fold: First, the results of two validation studies (total sample size $N = 941$) indicate that the UMI has excellent psychometric properties, measuring six different types of mo-

CHI 2018 April 21–26, 2018, Montreal, QC, Canada

ACM ISBN 978-1-4503-5620-6/18/04.
DOI: https://doi.org/10.1145/3173574.3173680

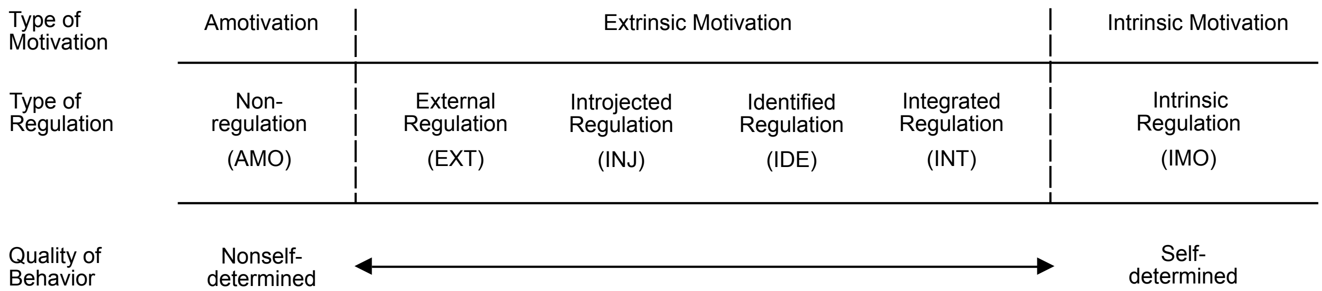

Figure 1. The different regulations of self-determination theory ranging from the least self-determined (amotivation) to the most self-determined regulation (intrinsic motivation). Figure adapted from [10], pp. 16.

tivational regulation across a wide range of technologies with high reliability as well as convergent and discriminant validity of each subscale. Second, we demonstrate how the different motivational regulations relate to core concepts in user experience research, such as need satisfaction and usability. Third, the UMI was found to detect differences in motivation for people who consider abandoning a technology compared to those who do not question their use. This initial test of criterion validity shows that the motivational regulations relate to different experiences. Taken together, the UMI represents a promising tool for assessing the motivation that users approach an interactive system with.

RELATED WORK

Self-determination theory in HCI

SDT is a theory of motivational states and processes as shaped by the social context and people's individual differences [10, 43]. SDT and its related concepts, most notably need satisfaction, have been applied to studying various areas within HCI, such as games [2, 11, 45] and user experience [19, 21]. Hassenzahl et al. [19], for instance, showed that need satisfaction is a key component of positive experiences with technology, a notion which was supported in several subsequent studies [31, 35, 52, 53]. Specifically, Hassenzahl [18, 19] argued that striving for need satisfaction constitutes one of the underlying reasons for "why" people choose to interact with technology. Making a phone call, for instance, is in itself not a meaningful action [18]. The same action (the "what"), however, becomes meaningful through striving for need satisfaction, such as calling one's spouse to satisfy the need for relatedness (the "why").

SDT defines the "what" and "why" of a behavior differently than Hassenzahl [18]. In SDT, need satisfaction is an outcome of goal pursuit (the "what") [8]. According to organismic integration theory, a sub-theory of SDT, the extent to which an activity supports need satisfaction is dependent on the underlying motivational regulation (the "why") [8]. For example, if the aforementioned phone call with one's spouse were extrinsically motivated to avoid feelings of guilt, it might result in less satisfaction of the need for relatedness than if it were driven by intrinsic motivation. We chose to base our approach on SDT, because the posited motivational regulations potentially offer a more nuanced understanding of how people's motivation for technology use affect the user experience, such as need satisfaction, than is provided in current HCI research.

Organismic integration theory

Organismic integration theory (OIT), a sub-theory of SDT [10, 43], differentiates three broad types of motivation to engage in an activity (see Figure 1): 1) **Amotivation** is characterized by a lack or absence of motivation; 2) **Extrinsic motivation** occurs when pursuit of a behavior is not completely self-determined, meaning it is controlled by factors outside of the self; 3) **Intrinsic motivation** is regarded as the most positive form of motivation, as behavior is completely self-determined and, in contrast to extrinsic motivation, not a means to an end but rather pursued for its own sake. Intrinsically motivated behavior is sustained by the experience of interest and enjoyment.

Activities that are not experienced as interesting or inherently enjoyable require extrinsic motivation. To initially engage in these activities, the perception of a relation between the activity and a desired outcome such as implicit approval or a reward is needed. In OIT, this is described as **external regulation** (a form of extrinsic motivation), which yields the least self-determined behavior and typically occurs in situations where people act to obtain a reward or avoid punishment (e.g., My friends would be angry with me if I quit using Facebook). However, when people take up values, attitudes, or regulatory structures, externally regulated behaviors may become internalized and then no longer require the presence of rewards or threats [10]. Specifically, SDT posits that the degree of internalization operates along a controlled-to-autonomous continuum (see Figure 1, from left to right): **Introjected regulation** describes an external regulation which has been partially internalized but not truly accepted as one's own. Such behaviors are pursued to avoid guilt or shame or to achieve feelings of self-worth (e.g., I would feel guilty if I quit using Fitbit). The more self-determined behavior of **identified regulation** follows from the conscious valuing of a behavioral goal. People whose behavior is regulated through identification accept the behavior as personally important (e.g., Using Excel to keep track of expenses). **Integrated regulation** is the most self-determined form of extrinsic motivation and results when an activity is congruent with personally endorsed values, goals, and needs that are already part of the self (e.g., I use LaTeX because I am a scientist, not because it is particularly enjoyable or interesting).

Differing consequences of motivational regulations

The different types of motivations on the self-determination continuum are associated with different behavioral, cognitive and emotional consequences. SDT postulates that the consequences are decreasingly positive the less self-determined the quality of a behavior is [10]. Specifically, intrinsic motivation is expected to lead to the most positive consequences, followed by integrated and identified regulation, which are forms of extrinsic motivation. Introjected and external regulation are presumed to lead to negative consequences, and amotivation to the most negative consequences [10]. Several studies have provided evidence that more self-determined types of motivation (intrinsic motivation, integrated regulation, and identified regulation) lead to the most positive behavioral (e.g. greater persistence), cognitive (e.g., enhanced concentration), and emotional (e.g., more positive emotions, greater well-being) outcomes when compared to nonself-determined types of motivation (introjected regulation, external regulation, and amotivation) [8, 10]. Support for this notion has been found for a wide variety of life domains such as academic achievement [17], sports [29, 33, 37], romantic relationships [3, 16], environmental protection [38], therapy motivation [39], and consumer behavior [58]. Zhang et al. [58], for instance, found that while experiential purchases (e.g., holidays) are typically regarded as more positively related to well-being than material purchases, this effect largely depended on people's motivational regulation. People who spent money on experiential purchases for autonomous reasons, meaning that they regarded them as an important part of their life, reported more need satisfaction, more flourishing, and vitality than people who spent money on these experiences for controlled reasons, such as for the recognition they got from others.

As with other life domains, technology use is likely motivated by different regulations. For instance, the notions of hedonic and eudaimonic motivation employed in the study of Mekler and Hornbæk [31] bear much semblance to intrinsic motivation and integrated regulation respectively, but do not account for less autonomous regulations. In another example, LaFrenière et al. [26] followed OIT when developing a scale for assessing gaming motivation and also found that more autonomous regulations (i.e., intrinsic motivation, integrated and identified regulation) were associated with need satisfaction, while the less self-determined regulations (i.e., introjected and external regulation, amotivation) were not. However, due to their instrument being specific to (arguably certain) games only (items include e.g., "I play video games to acquire powerful rare items"), it is not readily applicable to assessing people's motivation for using other interactive systems.

Yet given the great influence these different regulations might have on people's experience and use of interactive technology, a better understanding of users' motivation – the *why* of interaction – is imperative. To this end, we aim to measure different types of motivation for technology use. In the first study, a new measure of user motivation is developed. In the second study the underlying theoretical structure is verified and the impact of different types of motivation on usability, well-being and likelihood to recommend is investigated.

Development and validation strategy

The development and validation of the UMI followed best practices [12, 32]. In the first phase, we reviewed existing scales and adapted items to reflect the theoretical dimensions of motivational regulation in the context of technology use. This large item pool was subject to an item sort task and further refinement by the authors. This phase also included an independent expert review of content validity. In the second phase, the item pool was administered to a development sample in Study 1 to optimize scale length and identify the best items reflecting each of the six motivational regulations. In the third phase, we explored the dimensionality, reliability, convergent and discriminant validity with an independent validation sample in Study 2. To ensure construct validity, we also studied relations of these six motivational regulations to conceptually relevant measures from SDT and UX research. In the fourth phase, we investigated how motivational regulation differs in people who had thought about abandoning a technology compared to those who never thought about quitting.

ITEM POOL DEVELOPMENT AND REVIEW

Existing scales

In line with previous SDT research, the UMI was designed to measure the general motivation to engage with a specific technology. Existing scales on motivation were the basis of item development along with the definition of the different types of regulation described in the Handbook for Self-Determination Research [9]. The existing scales we used as item sources were developed for the areas of academic achievement (SIMS [17]), video games (GAMS, [26]), sports (BRSQ [28]; BREQ [34] and BREQ-2 [30]; SMS-6, [29]; SMS-II, [37]; PLOC-R, [57]), environmental protection (METS, [38]), romantic relationships (CMQ, [3]), therapy motivation (CMOTS, [39], school (PLOC, [42]), and well-being ([50]). While some motivational regulation scales do not include all dimensions posited by organismic integration theory (e.g., SIMS [17] or BREQ-2 [30]), we opted to cover all six dimensions to adequately represent the theoretical foundation and granularly differentiate between all regulations. Particular care was therefore taken to ensure that the UMI items have as little overlap as possible. Still, we expected that the items for these facets will correlate more strongly the closer to one another they are on the spectrum of motivation [42]. Based on these scales, an initial item pool of 249 items was created by the first author. Particular care was taken to adapt the wording of the items to reflect technology use. In a next step, all authors reviewed the item pool and removed or rephrased duplicates, near-duplicates, as well as items that were too specific. A pool of 150 items remained, which were, similar to the User Burden Scale [51], rephrased to include a placeholder for the technology in question (e.g., "I enjoy using [x]").

Item Sort Task

The first, second and last author independently conducted an item sort task [24] with all 150 items. Any items that did not receive a 100% agreement on the intended construct were removed, unless one of the authors involved at this stage vetoed the removal of an item. A total of 102 items remained.

Expert review

Two psychologists with expertise in self-determination theory, but who were not themselves involved in this research project, reviewed the pool in a 2-hour workshop. The goal was to review content validity, that is, ensure that all relevant aspects of motivation for technology use were covered. The experts rated each item on relevance, clarity and checked whether any aspects were missed. This review led to the removal of fifteen items and rewording of four items. Additionally, six items capturing integrated regulation were created. At this stage, the questionnaire consisted of 93 items. Amotivation was measured with 16 items, external regulation with 12, introjected regulation with 18, identified with 12, integrated regulation with 20, and intrinsic motivation with 15 items.

STUDY 1

The purpose of Study 1 was to test and reduce the UMI item pool to identify the best items measuring the proposed six regulations. To this end, we deployed a survey on Amazon Mechanical Turk and conducted an analysis with four steps: 1) psychometric analysis of all 93 items; 2) factor analysis for a subset of items; 3) selection of the best items; 4) factor analysis with these items for a preliminary structural validation. The number of items and the expected communalities determine the sample size required for factor analysis [12]. In general, a sample size of at least 200 participants is recommended [23]. We expected high communalities and good performance of the items as they were based on existing scales. Nevertheless, since the number of items under examination was large, we aimed for a sample size of over 450.

Procedure

After providing consent, participants were asked to fill in basic demographic information (gender, age, and experience with games). Next, they named an interactive technology that they used frequently. The focus was set on a frequently used technology to make sure that the UMI is applicable to widely used technologies. Participants were then asked to describe the technology and explain how they use it. The rationale behind this question was that if an uncommon website or technology was listed, we would have some information about what it is and how it can be used. We then also asked participants to report how frequently the technology was used and asked them to answer several scales in relation to this specific technology that will be discussed in the next paragraph. On the last page, participants were asked to indicate whether they answered the questions seriously (this served as a self-reported measure of data quality), they also had possibility to comment on the study and were given a completion code for Mechanical Turk. An instructed response item was included in the UMI items to filter out careless participants.

Participants

A total of 507 participants from the US completed the full questionnaire on Amazon Mechanical Turk. The survey took 13 minutes ($SD = 5.8$ minutes) to complete on average.

Data cleaning

Based on the recommendations by [7], 17 participants were excluded because they completed the survey in less than 4 minutes or not in one session. Two additional participants were excluded because they selected the same answer for more than 83.3% of the UMI items (5 out of 6). Seven participants were excluded because of a negative person-total correlation, which is an indicator for very unusual answering patterns [7].

Sample description

A total of 481 participants ($M_{age} = 38.31, SD = 12.61$, $range = 19-75$; 39.1% male, 1.5% non-binary or not specified) were included in the analysis. Participants could freely report on any interactive technology they used frequently. A majority of 33% chose Facebook, 11% a not further specified Smartphone, 10% iPhone, and 46% various other technologies, such as the Fitbit, other handheld devices such as Android OS, iPad, video game consoles such as the Playstation 4, as well as other social networking services such as Reddit and Twitter. With regards to the last 14 days, 45.9% of participants indicated that they used the interactive technology on average six times a day or more, 17.5% four to five times per day, 21.2% two to three times per day and 15.4% once a day or less.

Measures

In addition to the UMI items and demographic variables, five scales were included in Study 1. However, due to space concerns, only the measures relevant for the development of the UMI are reported here. The other scales were also included in Study 2 and are discussed in more detail in the Measures section of Study 2. Please note that all measures and data from *both* studies are available on https://osf.io/m3fbk/.

Type of technology

Participants could name any single technology. The statistical software R was used to semi-manually clean this data to ensure that typos and different spellings were associated with the correct technology name.

Frequency of use

A single-item measure captured frequency of use: *How frequently did you use this [referring to the technology they named above] interactive technology in the last 14 days?*.

UMI

The 93 items of the initial UMI item pool were distributed over four pages and displayed in random order. A 7-point Likert-type scale ranging from 1 (strongly disagree) to 7 (strongly agree) was used. Two attention check items were implemented to flag and exclude participants that answered carelessly: *This is a verification item, please select strongly disagree* and *I read instructions carefully. To show that you are reading these instructions, please leave this question blank.*

Results

The analysis was twofold: A psychometric item analysis was performed to remove problematic items, followed by an exploratory factor analysis to examine the structure and reduce the number of indicators to a scale with 3 items per construct.

Item analysis

Since the UMI was intended to measure six distinct (but possibly correlated) constructs based on SDT, it was decided to investigate descriptive statistics, difficulty indices, item variance, and discriminatory power for each construct separately. Please refer to the online materials for statistics on all 93 items. The item analysis followed recommendations by Moosbrugger and Kelava [32]. We removed one item with a variance of less than 1, because items with low variance are not suitable for differentiating between participants [32]. Discriminatory power describes how a single item's ability to differentiate between participants relates to the ability of the entire scale to differentiate between participants. Two items were removed because their discriminatory power was below the recommended value $< .30$ [4]. For each construct, inter-item correlations and average inter-item correlation (homogeneity) were investigated. Six items with an average inter-item correlation of less than .4 were removed, resulting in 84 items remaining for factor analysis.

Exploratory Factor Analysis (EFA)

Bartlett's test indicated factorability ($\chi^2_{df=3486} = 34439.59, p < .001$) as did the average Kaiser-Meyer-Olkin factor adequacy measure (Overall MSA = 0.96, none below .8). Parallel analysis suggested seven factors but the visual inspection of the Scree plot and our theoretical assumption suggested six factors. Since the goal was to reduce the number of items of the UMI and construct a scale that is consistent with the theoretical model, we conducted the EFA with six factors.

Data were tested for multivariate normality with Mardia Tests ($\chi^2_s = 181720.4, p < .001$; $Z_k = 127.4, p < .001$). Both tests indicated highly non-normal data. Hence, we chose to use principal axis factoring. The six factors had eigenvalues greater than 1.99 and explained 59% of the total variance. The factors were rotated using the Oblimin method, because following theory some of the factors were expected to be correlated. Results revealed factor correlations ranging from -.31 to .56. The factor loadings and communalities of all 84 items are reported in the online materials. The communality of one item was below the recommended threshold of .30 and subsequently removed. To further reduce the item pool and improve the measurement model, items were removed based on two criteria. First, items that showed substantial (>.30) cross-loadings were removed. Second, items with the highest loadings on the primary factor were retained unless there was already an item selected with very similar wording. The goal was to balance psychometric and theoretical grounds for item selection and optimize scale length [12]. Based on these criteria, the best eighteen items with three items per factor were identified (item wording depcited in Table 1).

These eighteen items were then subject to a second analysis. Bartlett's test was significant ($\chi^2_{df=153} = 5238.14, p < .001$) and the average Kaiser-Meyer-Olkin measure was .83 with none below .75. Parallel analysis and inspection of the Scree plot suggested six factors and data were non-normal (Mardia tests: $\chi^2_s = 6270.91, p < .001$; $Z_k = 59.75, p < .001$).

Principal axis factoring (Oblimin rotation) extracted six factors with eigenvalues greater than 0.89 explaining 71% of the total variance. Correlations of the six factors and internal consistencies are shown in Table 2. Factor loadings and communalities of the final 18 items are depicted in Table 1. All eighteen items loaded substantially on their designated factor without any notable cross-loadings (none above .3), reflecting the theoretically assumed six-factor structure.

Discussion

In Study 1, we identified eighteen items measuring six related constructs. In line with OIT, the correlations between the factors show that conceptually close regulations such as intrinsic motivation and integrated regulation correlate more strongly than intrinsic motivation and introjected regulation. The results of the second factor analysis support a six dimensional measure with high reliability and good psychometric properties. This structure was put to test and investigated in relation to other measures in Study 2. As with Study 1, we aimed at a sample size of over 450 participants.

STUDY 2

The goal of Study 2 was to validate the measurement model with a different sample applying confirmatory factor analysis (CFA). To test construct validity, we first investigated fit measures of the proposed model and compared them to alternative models. Second, to ensure convergent and divergent validity, we studied how the different UMI dimensions relate to other relevant constructs from SDT and user experience research. Third, motivational patterns between participants who questioned their use and those who never did so were investigated to gain a deeper understanding of the interplay between users' intentions, behavior and motivation.

Procedure

Study 2 was largely patterned after Study 1. After providing consent and basic demographic information, participants were again asked to name an interactive technology that they used frequently, followed by the eighteen UMI items, four additional open questions and the remaining measures.

Participants

A total of 498 participants from the United States completed the full survey on Amazon Mechanical Turk. The survey took 8 minutes ($SD = 4.4$ minutes) to complete on average. None had previously partaken in Study 1.

Data cleaning

Data were cleaned following the same procedure from Study 1. Five participants were excluded because they listed more than one technology or did not comply with the task. One participant indicated that we should not use their data. Four participants completed the survey in less than 3 minutes or not in one single session. One participant selected the same answer for all 18 items of the UMI and 27 participants had a negative person-total correlation.

Sample description

After data cleaning, a total of 460 participants ($M_{age} = 37.38, SD = 11.74$, $range = 18 - 76$; 40.4% male, less than 1% non-binary or not specified) were included in the analysis. Again, participants could freely choose any interactive technology they used frequently. A majority of 42% chose Facebook, 5% Instagram, 5% Twitter, 4% Fitbit, and 44% various other

Subscale	Item	Factor						
		AMO	EXT	INJ	IDE	INT	IMO	$h2$
Amotivation	1. I use [X], but I question why I continue to use it	.864						.739
	2. I use [X], but I wonder what is the point in using it	.854						.739
	3. I use [X], but I don't see why I should keep on bothering with it	.830						.711
External regulation	1. Other people will be upset if I don't use [X]		.799					.571
	2. I use [X] because others will not be pleased with me if I don't		.836					.727
	3. I feel under pressure from others to use [X]		.736					.681
Introjected regulation	1. I would feel bad about myself if I quit [X]			.943				.855
	2. I would feel guilty if I quit using [X]			.840				.740
	3. I would feel like a failure if I quit using [X]			.827				.728
Identified regulation	1. Using [X] is a sensible thing to do				.546			.458
	2. The benefits of using [X] are important to me				.742			.601
	3. Using [X] is a good way to achieve what I need right now				.872			.744
Integrated regulation	1. I use [X] because it reflects the essence of who I am					.795		.714
	2. Using [X] is consistent with my deepest principles					.773		.695
	3. I use [X] because it expresses my values					.932		.818
Intrinsic motivation	1. I use [X] because it is enjoyable						.849	.734
	2. I think using [X] is an interesting activity						.760	.637
	3. Using [X] is fun						.929	.843

Table 1. Pattern matrix from the EFA in Study 1 ($N = 460$) with the final version of the UMI. Loadings of all 18 items on the six factors are depicted, loadings below .3 are not shown. $h2$ = Communalities. [X] is a placeholder for the technology chosen by the participants.

	1.	2.	3.	4.	5.	α
1. AMO	-					.89
2. EXT	.30	-				.84
3. INJ	.14	.44	-			.91
4. IDE	−.21	.11	.33	-		.80
5. INT	.02	.19	.50	.47	-	.89
6. IMO	−.33	−.20	.06	.15	.40	.89

Table 2. Factor correlations and internal consistency (Cronbach's α) for Study 1 with 18 items.

technologies. Among them were productivity software such as MS Word or Excel, other social media networks such as YouTube and Reddit, as well technologies such as Amazon Echo or Android OS. Over the last 14 days, 38.3% of the participants indicated that they used the interactive technology on average six times a day or more, 19.8% four to five times per day, 22.2% two to three times per day and 19.8% once a day or less.

Open questions about technology use

After answering the UMI, participants were asked to describe in their own words why they use the technology in question (1). For illustration purposes, two contrasting answers from two different participants about the same technology are presented:

> *I use Facebook because it is a way to connect to people [...] using Facebook allows me to see family photos, to reach out to other loved ones and to see what is going on in the lives of those that I really care about. It's just a great way to keep the connection going. I really do value the technology.* [P78, M, 45, Facebook]

> *I signed up with Facebook about 2 years ago. I started because my child's sport was keeping parents informed about sport related information. Now I mostly use it to stalk people. I hate that I look at it all the time. It feels like a time-suck. I'm looking at it and I don't even know why I keep scrolling through items, but I do. I'm trying to limit it to just using it to post garage sale related items [...]* [P85, F, 35, Facebook]

As follow-up questions, we asked participants who or what had brought them to use the technology in question (2), why they think they continue using it (3) and whether they had ever thought about quitting using this technology (4). The last question (4) was later used to create groups of participants questioning their technology use versus those that did not. A systematic qualitative analysis of the open questions was beyond the scope of this paper. However, all answers are available in the online materials.

Measures

All measures consisted of 7-point Likert-type scales, unless otherwise noted. The items were presented in randomized order for each measure.

Construct validity was examined by exploring the relationship of the UMI with several established measures from SDT and user experience research. In SDT, the positive effects of need satisfaction on well-being are thought to be mediated by motivation [54]. Need satisfaction and vitality were expected to be in general more positively related to the self-determined types of regulation. The same general pattern was also expected for usability and likelihood to recommend. Satisfaction with life was included as a very global measure of well-being that is distant from technology use. Thus, it should ideally be not or only weakly related to the other measures.

Need Satisfaction

Need satisfaction is an essential aspect of positive user experiences with interactive technology [18, 31]. Satisfaction of the needs for autonomy (Cronbach's $\alpha = .83$), competence ($\alpha = .76$), and relatedness ($\alpha = .91$) were measured with three items each, taken from Sheldon's need satisfaction scale [49].

The introductory question was adapted to reflect the use of technology: *How do you feel when you use [X]?*. Perceived need satisfaction for autonomy ($M = 5.1$, $SD = 1.4$), competence ($M = 4.14$, $SD = 1.59$), and relatedness ($M = 4.36$, $SD = 1.87$) were around the middle of the scale. An overall need satisfaction score ($\alpha = .8$) aggregated over all three needs was calculated with an average of $M = 4.53$, $SD = 1.16$.

Vitality
To gain an understanding of how different motivations affect well-being, state level vitality ($\alpha = .92$) was measured with seven items developed by [44]. Item wording was slightly adapted to include the technology (e.g., *When I use [X], I feel alive and vital.*). Descriptive statistics showed that participants tended to answer this scale around the midpoint of the scale ($M = 4.25$, $SD = 1.4$).

Satisfaction with life scale (SWLS)
To measure a construct that was not related to use of technology directly, but might be related to feelings of need satisfaction and vitality, general life-satisfaction was measured with the five items developed by Diener et al. [13]. Internal consistency was high ($\alpha = .91$) and agreement was moderate ($M = 4.62$, $SD = 1.49$).

Usability Metric for User Experience (UMUX)
The four items of the UMUX developed by Finstad [14] were employed to measure perceived usability of the reported technology. Internal consistency was acceptable with $\alpha = .69$. Overall, perceived usability was high ($M = 6.12$, $SD = 0.86$).

Likelihood to recommend (LTR)
LTR is a measure of engagement and satisfaction that is distinct but related to usability [46]. LTR was assessed with the question used to calculate the Net Promoter Score [40] on a scale from 0 to 10. Average LTR was above 8 ($M = 8.47$, $SD = 1.94$).

Results
Item analysis
Descriptive statistics of the UMI items were in the same range as in Study 1 (see Table 3). Average agreement to the statement was higher and rather left-skewed for the more self-determined types of regulation than for amotivation, external regulation and introjected regulation. Inter-item correlations were high within the factors, but low to moderate between the different factors (see also additional Tables in the online materials).

Confirmatory factor analysis (measurement model)
To test the multidimensional factor structure of the UMI, a confirmatory factor analysis (CFA) was conducted. All items were specified to load on their designated factor, and the loading of the first item was constrained to one. Multivariate normality was not given (Mardia tests: $\chi^2_s = 3888.9, p < .001$; $Z_k = 32.278, p < .001$), therefore we used a robust maximum likelihood estimation method with Huber-White standard errors and a Yuan-Bentler based scaled test statistic. Results of the CFA suggest that the proposed model fits the data well [$\chi^2_{120} = 237.53$, $p < .001$, $\chi^2/df = 1.98$, $CFI = .966$, $SRMR = .046$, $RMSEA = .046$, $PCLOSE = .771$]. The measurement model is depicted in Figure 2.

	#	*M*	*SD*	*S*	*K*	*pv*
Amotivation	1.	2.19	1.60	1.25	0.50	31.3
	2.	2.23	1.60	1.23	0.55	31.8
	3.	2.11	1.49	1.30	0.76	30.1
External regulation	1.	2.28	1.73	1.19	0.20	32.6
	2.	1.98	1.53	1.66	1.90	28.3
	3.	2.07	1.54	1.40	0.88	29.6
Introjected regulation	1.	2.59	1.89	0.97	−0.32	37.0
	2.	2.49	1.86	1.02	−0.26	35.5
	3.	2.00	1.61	1.61	1.47	28.6
Identified regulation	1.	4.89	1.66	−0.48	−0.40	69.9
	2.	5.33	1.62	−0.84	−0.09	76.2
	3.	5.06	1.77	−0.60	−0.65	72.3
Integrated regulation	1.	3.52	1.90	0.13	−1.14	50.2
	2.	3.38	1.88	0.29	−0.99	48.3
	3.	3.61	1.87	0.11	−1.03	51.6
Intrinsic motivation	1.	5.82	1.29	−1.13	0.91	83.1
	2.	5.79	1.22	−1.05	0.96	82.8
	3.	5.82	1.28	−1.22	1.41	83.1

Table 3. Descriptive statistics of all items including all participants ($N = 460$). S = Skewness. K = Kurtosis. pv = Difficulty index. Higher difficulty values indicate that people on average agree with this item, while lower values indicate the opposite.

A model with two factors (amotivation and extrinsic-intrinsic spectrum) and a model with 3 factors (amovation, controlled regulation consisting of external and introjected regulations, and autonomous regulation consisting of identified and integrated regulations as well as intrinsic motivation) were tested. Results show that the fit for the alternative models were significantly worse than for the proposed model (refer to the online materials for detailed information on model comparison).

	M	*SD*	M_{tr}	ρ_C	α	AVE	MSV
AMO	2.18	1.42	1.77	.90	.90	.74	.36
EXT	2.11	1.36	1.71	.82	.81	.61	.22
INJ	2.36	1.58	1.92	.86	.85	.68	.27
IDE	5.10	1.44	5.21	.82	.81	.60	.36
INT	3.50	1.63	3.44	.84	.83	.63	.35
IMO	5.81	1.13	6.00	.88	.87	.71	.20

Table 4. Means, standard deviations and trimmed means (20%), Congeneric reliability (ρ_C), Cronbach's α, Average Variance Extracted (AVE) and Maximum Shared Variance (MSV) for the UMI in Study 2.

Reliability, convergent and discriminant validity
As seen in Table 4, congeneric reliablity and internal consistency were high ($\rho_C > .7$, Cronbach's $\alpha > .8$), indicating high reliability. For all subscales Average Variance Extracted (AVE) was above the threshold of .5, suggesting high convergent validity. Maximum Shared Variance (MSV) values were lower than the corresponding AVE scores, indicating high discriminant validity.

Motivation and Related Measures
To investigate the relationship and construct validity of the six types of regulation with other constructs, we calculated the mean for each UMI subscale for each participant. Descriptive statistics of the six motivations are depicted in Table 4. Because most of the measures were not normally distributed, we calculated Pearson correlations with bootstrapping (1000

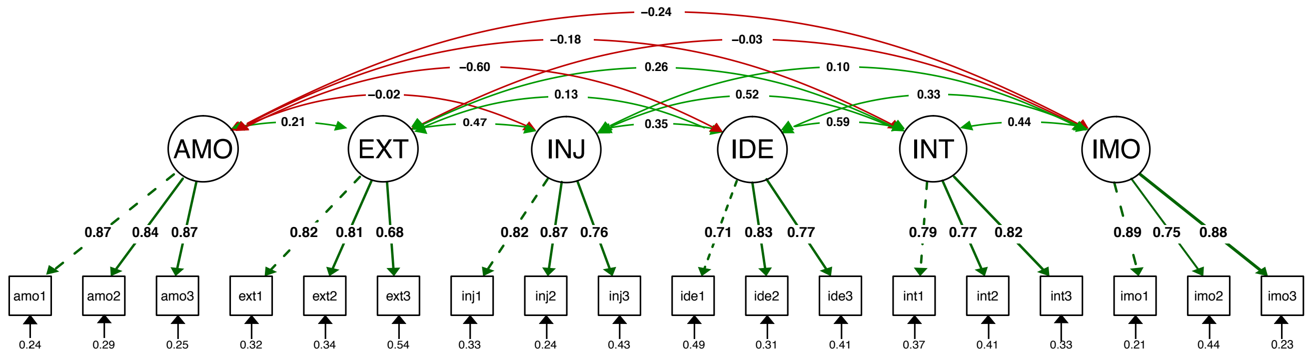

Figure 2. Measurement model of the UMI with standardized loadings. Dotted lines indicate loadings that were constrained to one. [$\chi^2_{120} = 237.53$, $p < .001$, $\chi^2/df = 1.98$, $CFI = .966$, $SRMR = .046$, $RMSEA = .046$, $PCLOSE = .771$]

iterations). Results are presented in Table 5. As posited by SDT, conceptually close regulations on the spectrum of motivation are more strongly correlated with each other than more distant regulations (termed 'simplex-like structure' in SDT literature). For the SDT related measures, need satisfaction and vitality were expected to be in general more strongly associated with the more self-determined regulations. The same pattern was also expected for usability and LTR since the more autonomous regulations are supposed to lead to the most positive outcomes. SWLS as a global measure of well-being was not expected to correlate substantially with technology related measures.

Need satisfaction

The most self-determined forms of regulation, integrated and intrinsic motivation, were positively associated with all need satisfaction measures. Relatedness was found to be positively associated with very self-determined regulations, but not with amotivation, identified or introjected regulations. Notably, relatedness was positively correlated with external regulation, which in SDT is not typically associated with positive outcomes on long-term motivation and well-being.

Vitality and Satisfaction with Life

Vitality showed positive correlations with all regulations except amotivation, which expectedly correlated negatively. In line with previous research within SDT (e.g., [58]), external regulation and vitality did not correlate significantly. Satisfaction with life was not directly related to any type of motivation, but was slightly positively correlated with the need for relatedness and feelings of vitality.

Usability

Amotivation and external regulation were negatively associated with perceived usability, whereas identified regulation and intrinsic motivation were positively correlated with usability. Neither introjected nor integrated regulation were significantly correlated with usability.

Likelihood to recommend

LTR was negatively correlated with amotivation, suggesting that users who do not know why they use a particular technology are less likely to recommend it to others. LTR was not significantly related to external motivation, but positively associated with all remaining types of motivation. As expected, usability and LTR correlated positively.

Motivation and Questioning Technology Use

As a test of criterion validity, we investigated whether the UMI is able to detect differences between groups that we expected to differ in their motivation. The majority of the participants in Study 2 (n = 297) had never questioned their technology use, but 163 participants indicated that they had at some point questioned their use and thought about abandoning the technology, even though they were presently still using it. The two groups were compared with regards to their UMI ratings. Due to the non-equal group sizes and data featuring non-normal distribution, outliers and unequal variances, we applied robust Yuen-Welch tests to check for significant differences in trimmed means (as recommended by [1]). Results in Table 6 show that except for introjected and external regulation, all differences for the motivational regulations were statistically significant with effect sizes ranging from small (integrated regulation), over medium (intrinsic motivation, identified regulation) to large (amotivation). Participants who had questioned their technology use reported higher levels of amotivation, as well as lower levels of the more autonomous regulations and intrinsic motivation. This suggests that users who question their use have different regulations and might be more likely to abandon the technology in the future.

GENERAL DISCUSSION

The aim of the present work was to develop and validate a measure of motivation in the context of technology use based on self-determination theory. Our study results support the proposed factor structure of the UMI and show that it is a reliable and valid measure of users' motivation. The scale was found to have excellent psychometric properties measuring users' motivation across a wide range of technologies. Correlations of the UMI with related measures, most importantly need satisfaction, are in line with previous SDT research. We could show how different motivations relate to need satisfaction, usability and how they might affect consequences of technology use, such as well-being and likelihood to recommend. Moreover, the UMI was found to be sensitive to users who think about abandoning a technology. Lastly, by making all data and statistical scripts used in our analysis available

		1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.
1.	Amotivation													
2.	External regulation	.20***												
3.	Introjected regulation	−.02	.37***											
4.	Identified regulation	−.51***	.09*	.30***										
5.	Integrated regulation	−.17***	.21***	.45***	.50***									
6.	Intrinsic motivation	−.23***	−.02	.11*	.29***	.40***								
7.	Autonomy	−.22***	−.05	.14**	.37***	.49***	.46***							
8.	Competence	−.29***	.05	.35***	.56***	.50***	.24***	.44***						
9.	Relatedness	.05	.31***	.04	.01	.29***	.22***	.33***	.07					
10.	Need satisfaction†	−.20***	.17***	.23***	.41***	.58***	.42***	.78***	.67***	.71***				
11.	Vitality	−.32***	.03	.32***	.46***	.58***	.47***	.51***	.63***	.28***	.64***			
12.	SWLS‡	−.05	.00	.02	−.02	.08	.03	.06	.08	.18***	.16***	.15**		
13.	Usability	−.43***	−.23***	−.06	.26***	.09	.33***	.32***	.14**	.03	.20***	.23***	.05	
14.	LTR§	−.40***	.00	.14***	.46***	.34***	.44***	.43***	.38***	.10*	.40***	.43***	−.02	.46***

Table 5. Pearson correlations with bootstrapping (1000 iterations) of the measures used in Study 2 ($N = 460$). *Note.* †Average of autonomy, competence, and relatedness. ‡Satisfaction with life scale. §Likelihood to recommend. * $p < .05$; ** $p < .01$; *** $p < .001$.

	Use never questioned ($n = 297$)			Use questioned ($n = 163$)			Yuen-Welch test			
	M	SD	M_{tr}	M	SD	M_{tr}	t	df	p	ξ
AMO	1.74	1.054	1.38	2.98	1.649	2.79	−8.496	122.9	< .001	0.583
EXT	2.05	1.353	1.62	2.23	1.381	1.89	−1.831	178.6	.069	0.124
INJ	2.46	1.670	2.00	2.18	1.375	1.82	1.081	244.7	.281	0.085
IDE	5.38	1.342	5.54	4.58	1.473	4.58	5.849	180.2	< .001	0.404
INT	3.69	1.665	3.67	3.15	1.512	3.05	3.567	223.0	< .001	0.250
IMO	5.90	1.176	6.16	5.65	1.030	5.72	3.526	208.8	.001	0.257

Table 6. Comparison of participants who never questioned their use and participants who thought about quitting. M_{tr} = 20% trimmed means used for the Yuen-Welch test. ξ = Explanatory measure of effect size; interpretation: 0.10 small, 0.30 medium, 0.50 large.

for the CHI community [25], we hope our paper provides a helpful and transparent template for future scale development and validation endeavors in HCI. In the following, we discuss the theoretical implications of our findings, We also report limitations of the present work, outline future research directions, and provide practical instructions for applying the UMI.

Participants reported relatively high levels of the self-determined and autonomous regulations (i.e., identified, integrated, and intrinsic motivation) and low levels of nonself-determined and controlled regulations (i.e., amotivation, external and introjected motivation). This suggests that the motivation to use a specific technology was relatively self-determined on average, perhaps reflecting the presumably more self-determined use of technology in users' spare time. However, about one third of participants had at some point thought about quitting the technology reported. Notably, they reported significantly lower levels of intrinsic motivation, integrated regulation, and identified regulation as well as higher levels of amotivation. This pattern is in line with the findings of Pelletier et al. [36], who showed that the more self-determined types of regulation were positively associated, whereas amotivation was consistently negatively associated with persistence. Moreover, while thinking about quitting a technology does not readily correspond to actually abandoning or even just intending to quit a technology, it is a step that may eventually lead to such a change in behavior. For instance, research on the motivation of high school students showed that less self-determined motivation correlated with higher levels of drop-out intentions, which was associated with actual drop out one year later [55].

Research in SDT also emphasized the importance of autonomy support as a predecessor of self-determined motivation and behavioral persistence [36, 55]. Correlations of the UMI with autonomy need satisfaction support this notion in the context of technology use, as autonomy was more strongly associated with self-determined regulations. In the SDT framework, these regulations are thought to link need satisfaction and affective, cognitive and behavioral consequences. With the UMI, researchers have a tool to investigate why people interact with a technology and possibly explain why a specific technology can have positive as well as negative effects on users' experience, well-being and behavior. For instance, results of Study 2 show that users who reported higher levels of the self-determined regulations also indicated stronger feelings of vitality after technology use – a measure of well-being. In general, this pattern of positive effects of self-determined regulation was also found for usability and likelihood to recommend.

Interestingly, among the more self-determined types of regulation, integrated regulation was not related to usability, perhaps suggesting that when the use of a technology aligns with one's values and core principles, usability might not be as important. In contrast, Mekler and Hornbæk [31] found that users who reported eudaimonically motivated experiences often mentioned instrumental qualities of a technology. However, note that Mekler and Hornbæk studied single experiences, whereas participants in the present work were asked to report on a frequently used technology and not a specific experience episode. Following previous SDT research, the UMI is a measure of general technology use, but evidence from other domains (e.g., academic achievement [17]) suggests that the UMI may also be applicable to single experience episodes. However, in specific situations motivated by one's personal values, usability might indeed be important (e.g., setting up a personal website). Taken together, it would be interesting to study how users' motivation shapes single experience episodes, and how experience episodes in turn influence motivation to use a technology. However, drawing from results in other domains (e.g., academic achievement [17]), situational motivational regulation can be expected to show a similar pattern as with general use. The UMI should be applied to examine motivation in single episodes to further test this notion.

SDT postulates that if people experience need satisfaction they can internalize an initially extrinsically motivated behavior (e.g., using a software because it is mandatory at work), shifting their motivational regulation from external towards integrated regulation over time. This means that over time people can feel effective in undertaking nonself-determined behaviors and they are more likely to personally endorse these actions. Autonomy supportive design [5], for instance, aims to design technologies that foster autonomy and self-determination over time. The UMI may thus be used to evaluate effects of different designs on motivation. Additionally, the UMI may possibly explain under what circumstances and for whom autonomy supportive technology is particularly effective. For instance, future research could study whether users with a controlled motivation or users with a more self-determined motivation benefit more from autonomy support. Although an in-depth investigation of these relationships was beyond the scope of the present work, our findings show that motivational regulations are associated differently with these constructs and provide a useful lens towards understanding user experience and outcomes of technology use.

With the UMI, researchers have a theory-based instrument to investigate the *why* of interaction. One advantage of approaching the why of interaction from the perspective of a well-researched theory is that one may draw from the large body of evidence on SDT to formulate and test hypotheses, and investigate this theory's applicability, limits, and predictive power for HCI [22]. The UMI may potentially help predict how likely users are to abandon a technology if given the opportunity. It may also serve to extend existing models of user experience that have already incorporated need satisfaction (e.g., [19, 22]), as different motivational patterns can explain why people have different experiences (and consequences) when using a particular technology. Specifically, the UMI may help to better understand the role of motivation in shaping need satisfaction, as well as how need satisfaction influences motivation to interact with a technology in the long term.

Limitations and Future Directions

The two studies reported here entail an initial thorough validation of the UMI. Although participants' age was distributed over a wide range, and a diverse set of technologies has been reported, the UMI needs to be further tested in-depth with users outside of Mechanical Turk and North America. The distinction between different types of regulation and their relation to well-being has been found to hold true in various languages and cultures [6, 15]. The structural validity of the UMI may be tested with other cultures to examine whether this is also applicable for motivation in technology use. Most technologies reported by the participants were leisure-oriented. Thus, the structure of the UMI needs to be tested in the work-related technology use as well. However, existing evidence for differentiating motivational regulations in the work domain [15] is encouraging.

Additionally, the types of technology could be specified to different types of domains or even specific technologies. For instance, user's experience and behavior has been found to vary in the domains of fitness technologies [47], games [11], and Facebook [48], depending on whether they engaged with the technology to get recognition from others or because they had personally endorsed it. The UMI might allow for more nuanced insights into the motivational processes underlying users' experience and technology use.

In the present study, we examined only technology that users reported to use frequently, therefore limiting us to technology that has not (yet) been abandoned. In a next step, known-groups validity should be investigated, for instance by examining how users perceived abandoned technology, similar as in the validation of the User Burden Scale [51]. Since the UMI was found to be sensitive to users who think about abandoning a technology, it would be interesting to test whether the UMI is predictive for abandoning a technology. Finally, a promising avenue for future research is examining whether the UMI relates to behavioral intentions and, most importantly, actual behavior.

UMI Guidelines for Use

While the present studies employed a placeholder [X] for the 18 UMI items, this may be replaced with the technology under investigation (e.g., "Using Facebook is fun"). To reduce sequence effects, it is generally advisable to randomize the order of the items. We used a 7-point Likert-type agreement scale from 1 (strongly disagree) to 7 (strongly agree) and recommend using the same answering scale to ensure comparability. Researchers can calculate a score for each regulation separately by averaging the three items corresponding to the subscale.

CONCLUSION

We present the development and validation of a multidimensional measurement tool rooted in self-determination theory that helps to deepen our understanding of *why* users interact with a technology. The development and validation followed best practices and all data collected in the two studies together with the materials and analysis code is available online. The UMI has been tested with over 900 participants and shows promising psychometric properties, high reliability, convergent and discriminant validity. The UMI has implications for theory and practice and opens up opportunities for future research on motivation and user experience.

ACKNOWLEDGEMENTS

Special thanks to Kasper Hornbæk, Lena Aeschbach, Seamus Forde, Livia Müller, and Serge Petralito. This work has been approved by the Institutional Review Board of the Faculty of Psychology, University of Basel under the number D-003-17.

REFERENCES

1. Marjan Bakker and Jelte M. Wicherts. 2014. Outlier removal, sum scores, and the inflation of the type I error rate in independent samples t tests: The power of alternatives and recommendations. *Psychological Methods* 19, 3 (2014), 409–427. DOI: http://dx.doi.org/10.1037/met0000014
2. Max V. Birk, Cheralyn Atkins, Jason T. Bowey, and Regan L. Mandryk. 2016. Fostering Intrinsic Motivation Through Avatar Identification in Digital Games. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 2982–2995. DOI: http://dx.doi.org/10.1145/2858036.2858062
3. Marc R. Blais, Stéphane Sabourin, Colette Boucher, and Robert J. Vallerand. 1990. Toward a Motivational Model of Couple Happiness. *Journal of Personality and Social Psychology* 59, 5 (1990), 1021–1031. DOI: http://dx.doi.org/10.1037/0022-3514.59.5.1021
4. Ingwer Borg and Patrick J. F. Groenen. 2005. *Modern multidimensional scaling: Theory and applications*. Berlin: Springer Science & Business Media. DOI: http://dx.doi.org/10.1007/0-387-28981-X
5. Rafael A. Calvo, Dorian Peters, Daniel Johnson, and Yvonne Rogers. 2014. Autonomy in technology design. In *Proceedings of the extended abstracts of the 32nd annual ACM conference on Human factors in computing systems - CHI EA '14*. ACM, 37–40. DOI: http://dx.doi.org/10.1145/2559206.2560468
6. Valery Chirkov, Richard M. Ryan, Youngmee Kim, and Ulas Kaplan. 2003. Differentiating autonomy from individualism and independence: A self-determination theory perspective on internalization of cultural orientations and well-being. *Journal of Personality and Social Psychology* 84, 1 (2003), 97–110. DOI: http://dx.doi.org/10.1037/0022-3514.84.1.97
7. Paul G. Curran. 2016. Methods for the detection of carelessly invalid responses in survey data. *Journal of Experimental Social Psychology* 66 (2016), 4–19. DOI: http://dx.doi.org/10.1016/j.jesp.2015.07.006
8. Edward L. Deci and Richard M. Ryan. 2000. The "What" and "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behavior. *Psychological Inquiry* 11, 4 (2000), 227–268. DOI: http://dx.doi.org/10.1207/S15327965PLI1104_01
9. Edward L. Deci and Richard M. Ryan. 2002a. *Handbook of self-determination research*. Rochester, NY: University Rochester Press.
10. Edward L. Deci and Richard M. Ryan. 2002b. Overview of Self-Determination Theory. In *Handbook of self-determination research*, Edward L. Deci and Richard M. Ryan (Eds.). Rochester, NY: University Rochester Press, Chapter 1, 3–33.
11. Sebastian Deterding. 2016. Contextual Autonomy Support in Video Game Play : A Grounded Theory. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 3931–3943. DOI: http://dx.doi.org/10.1145/2858036.2858395
12. Robert F. DeVellis. 2017. *Scale Development: Theory and Applications*. Vol. 26. Thousand Oaks, CA: Sage publications.
13. Ed Diener, Robert A. Emmons, Randy J. Larsen, and Sharon Griffin. 1985. The Satisfaction With Life Scale. *Journal of Personality Assessment* 49, 1 (1985), 71–75. DOI:http://dx.doi.org/10.1207/s15327752jpa4901_13
14. Kraig Finstad. 2010. The Usability Metric for User Experience. *Interacting with Computers* 22, 5 (2010), 323–327. DOI: http://dx.doi.org/10.1016/j.intcom.2010.04.004
15. Marylène Gagné, Jacques Forest, Maarten Vansteenkiste, Laurence Crevier-Braud, Anja van den Broeck, Ann Kristin Aspeli, Jenny Bellerose, Charles Benabou, Emanuela Chemolli, Stefan Tomas Güntert, Hallgeir Halvari, Devani Laksmi Indiyastuti, Peter A. Johnson, Marianne Hauan Molstad, Mathias Naudin, Assane Ndao, Anja Hagen Olafsen, Patrice Roussel, Zheni Wang, and Cathrine Westbye. 2015. The Multidimensional Work Motivation Scale: Validation evidence in seven languages and nine countries. *European Journal of Work and Organizational Psychology* 24, 2 (2015), 178–196. DOI: http://dx.doi.org/10.1080/1359432X.2013.877892
16. Graham S. Gaine and Jennifer G. La Guardia. 2009. The unique contributions of motivations to maintain a relationship and motivations toward relational activities to relationship well-being. *Motivation and Emotion* 33, 2 (2009), 184–202. DOI: http://dx.doi.org/10.1007/s11031-009-9120-x
17. Frédéric Guay, Robert J. Vallerand, and Céline Blanchard. 2000. On the Assessment of Situational Intrinsic and Extrinsic Motivation: The Situational Motivation Scale (SIMS). *Motivation and Emotion* 24, 3 (2000), 175–213. DOI:http://dx.doi.org/10.1023/A:1005614228250
18. Marc Hassenzahl. 2010. Experience Design: Technology for All the Right Reasons. *Synthesis Lectures on Human-Centered Informatics* 3, 1 (2010), 1–95. DOI: http://dx.doi.org/10.2200/S00261ED1V01Y201003HCI008
19. Marc Hassenzahl, Sarah Diefenbach, and Anja Göritz. 2010. Needs, affect, and interactive products – Facets of user experience. *Interacting with Computers* 22, 5 (2010), 353–362. DOI: http://dx.doi.org/10.1016/j.intcom.2010.04.002
20. Marc Hassenzahl and Daniel Ullrich. 2007. To do or not to do: Differences in user experience and retrospective judgments depending on the presence or absence of instrumental goals. *Interacting with Computers* 19, 4 (2007), 429–437. DOI: http://dx.doi.org/10.1016/j.intcom.2007.05.001

21. Kasper Hornbæk and Morten Hertzum. 2017. Technology Acceptance and User Experience: A Review of the Experiential Component in HCI. *ACM Transactions on Computer-Human Interaction* 24, 5, Article 33 (Oct. 2017), 30 pages. DOI:http://dx.doi.org/10.1145/3127358

22. Kasper Hornbæk and Antti Oulasvirta. 2017. What Is Interaction?. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 5040–5052. DOI: http://dx.doi.org/10.1145/3025453.3025765

23. Matt C. Howard. 2016. A Review of Exploratory Factor Analysis Decisions and Overview of Current Practices: What We Are Doing and How Can We Improve? *International Journal of Human-Computer Interaction* 32, 1 (2016), 51–62. DOI: http://dx.doi.org/10.1080/10447318.2015.1087664

24. Matt C. Howard and Robert C. Melloy. 2016. Evaluating Item-Sort Task Methods: The Presentation of a New Statistical Significance Formula and Methodological Best Practices. *Journal of Business and Psychology* 31, 1 (2016), 173–186. DOI: http://dx.doi.org/10.1007/s10869-015-9404-y

25. Matthew Kay, Steve Haroz, Shion Guha, Pierre Dragicevic, and Chat Wacharamanotham. 2017. Moving Transparent Statistics Forward at CHI. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '17 (CHI EA '17)*. ACM, New York, NY, USA, 534–541. DOI: http://dx.doi.org/10.1145/3027063.3027084

26. Marc-André K. Lafrenière, Jérémie Verner-Filion, and Robert J. Vallerand. 2012. Development and validation of the Gaming Motivation Scale (GAMS). *Personality and Individual Differences* 53, 7 (2012), 827 – 831. DOI: http://dx.doi.org/10.1016/j.paid.2012.06.013

27. Effie Lai-Chong Law and Paul van Schaik. 2010. Modelling user experience – An agenda for research and practice. *Interacting with Computers* 22, 5 (2010), 313–322. DOI: http://dx.doi.org/10.1016/j.intcom.2010.04.006

28. Chris Lonsdale, Ken Hodge, and Elaine A. Rose. 2008. The behavioral regulation in sport questionnaire (BRSQ): instrument development and initial validity evidence. *Journal of sport & exercise psychology* 30, 3 (2008), 323–55. DOI:http://dx.doi.org/10.1123/jsep.30.3.323

29. Clifford Mallett, Masato Kawabata, Peter Newcombe, Andrés Otero-Forero, and Susan Jackson. 2007. Sport motivation scale-6 (SMS-6): A revised six-factor sport motivation scale. *Psychology of Sport and Exercise* 8, 5 (2007), 600–614. DOI: http://dx.doi.org/10.1016/j.psychsport.2006.12.005

30. David Markland and Vannessa Tobin. 2004. A Modification to the Behavioural Regulation in Exercise Questionnaire to Include an Assessment of Amotivation. *Journal of Sport and Exercise Psychology* 26, 2 (2004), 191–196. DOI:http://dx.doi.org/10.1123/jsep.26.2.191

31. Elisa D. Mekler and Kasper Hornbæk. 2016. Momentary Pleasure or Lasting Meaning?. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16 (CHI '16)*. ACM, New York, NY, USA, 4509–4520. DOI: http://dx.doi.org/10.1145/2858036.2858225

32. Helfried Moosbrugger and Augustin Kelava. 2007. *Testtheorie und Fragebogenkonstruktion*. Berlin: Springer. DOI:http://dx.doi.org/10.1007/978-3-642-20072-4

33. Frederiki C. Moustaka, Symeon P. Vlachopoulos, Spyridoula Vazou, Maria Kaperoni, and David A. Markland. 2010. Initial validity evidence for the Behavioral Regulation in Exercise Questionnaire-2 among Greek exercise participants. *European Journal of Psychological Assessment* 26, 4 (2010), 269–276. DOI: http://dx.doi.org/10.1027/1015-5759/a000036

34. Elaine Mullan, David Markland, and David K. Ingledew. 1997. A graded conceptualisation of self-determination in the regulation of exercise behaviour: development of a measure using confirmatory factor analytic procedures. *Personality and Individual Differences* 23, 5 (1997), 745–752. DOI: http://dx.doi.org/10.1016/S0191-8869(97)00107-4

35. Timo Partala and Aleksi Kallinen. 2012. Understanding the most satisfying and unsatisfying user experiences: Emotions, psychological needs, and context. *Interacting with Computers* 24, 1 (2012), 25–34. DOI: http://dx.doi.org/10.1016/j.intcom.2011.10.001

36. Luc G. Pelletier, M. S. Fortier, Robert J. Vallerand, and N. M. Brierè. 2001. Associatons among perceived autonomy support, forms of self-regulations, and persistence: A prospective study. *Motivation and Emotion* 25, 4 (2001), 279–306. DOI: http://dx.doi.org/10.1023/A:1014805132406

37. Luc G. Pelletier, Meredith A. Rocchi, Robert J. Vallerand, Edward L. Deci, and Richard M. Ryan. 2013. Validation of the revised sport motivation scale (SMS-II). *Psychology of Sport and Exercise* 14, 3 (2013), 329–341. DOI: http://dx.doi.org/10.1016/j.psychsport.2012.12.002

38. Luc G. Pelletier, Kim M. Tuson, Isabelle Green-Demers, Kimberley Noels, and Ann M. Beaton. 1998. Why Are You Doing Things for the Environment? The Motivation Toward the Environment Scale (MTES)1. *Journal of Applied Social Psychology* 28, 5 (1998), 437–468. DOI: http://dx.doi.org/10.1111/j.1559-1816.1998.tb01714.x

39. Luc G. Pelletier, Kim M. Tuson, and Najwa K. Haddad. 1997. Client Motivation for Therapy Scale: a measure of intrinsic motivation, extrinsic motivation, and amotivation for therapy. *Journal of personality assessment* 68, 2 (1997), 414–35. DOI: http://dx.doi.org/10.1207/s15327752jpa6802_11

40. Frederick F. Reichheld. 2003. The one number you need to grow. *Harvard business review* 81, 12 (2003), 46–55.

41. Marco M. C. Rozendaal, David V. Keyson, and Huib de Ridder. 2007. Product Behavior and Appearance Effects on Experienced Engagement during Experiential and Goal-directed Tasks. In *Proceedings of the 2007 Conference on Designing Pleasurable Products and Interfaces*. ACM, 181–193. DOI: http://dx.doi.org/10.1145/1314161.1314178

42. Richard M. Ryan and James P. Connell. 1989. Perceived locus of causality and internalization: Examining reasons for acting in two domains. *Journal of Personality and Social Psychology* 57, 5 (1989), 749–761. DOI: http://dx.doi.org/10.1037/0022-3514.57.5.749

43. Richard M Ryan and Edward L Deci. 2000. Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology* 25, 1 (2000), 54–67. DOI: http://dx.doi.org/10.1006/ceps.1999.1020

44. Richard M. Ryan and Christina Frederick. 1997. On Energy, Personality, and Health: Subjective Vitality as a Dynamic Reflection of Well-Being. *Journal of Personality* 65, 3 (1997), 529–565. DOI: http://dx.doi.org/10.1111/j.1467-6494.1997.tb00326.x

45. Richard M. Ryan, C. Scott Rigby, and Andrew Przybylski. 2006. The Motivational Pull of Video Games: A Self-Determination Theory Approach. *Motivation and Emotion* 30, 4 (01 Dec 2006), 344–360. DOI: http://dx.doi.org/10.1007/s11031-006-9051-8

46. Jeff Sauro and James R. Lewis. 2012. *Quantifying the User Experience: Practical Satistics for User Research*. Cambridge, MA: Morgan Kaufmann.

47. Hanna Schneider, Kilian Moser, Andreas Butz, and Florian Alt. 2016. Understanding the Mechanics of Persuasive System Design: A Mixed-Method Theory-driven Analysis of Freeletics. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 309–320. DOI: http://dx.doi.org/10.1145/2858036.2858290

48. Kennon M. Sheldon, Neetu Abad, and Christian Hinsch. 2011. A two-process view of Facebook use and relatedness need-satisfaction: Disconnection drives use, and connection rewards it. *Journal of Personality and Social Psychology* 100, 4 (2011), 766–775. DOI: http://dx.doi.org/10.1037/a0022407

49. Kennon M. Sheldon, Andrew J. Elliot, Youngmee Kim, and Tim Kasser. 2001. What is satisfying about satisfying events? Testing 10 candidate psychological needs. *Journal of Personality and Social Psychology* 80, 2 (2001), 325–339. DOI: http://dx.doi.org/10.1037/0022-3514.80.2.325

50. Kennon M. Sheldon, Richard M. Ryan, Edward L. Deci, and Tim Kasser. 2004. The Independent Effects of Goal Contents and Motives on Well-Being: It's Both What You Pursue and Why You Pursue It. *Personality and Social Psychology Bulletin* 30, 4 (2004), 475–486. DOI: http://dx.doi.org/10.1177/0146167203261883

51. Hyewon Suh, Nina Shahriaree, Eric B. Hekler, and Julie A. Kientz. 2016. Developing and Validating the User Burden Scale: A Tool for Assessing User Burden in Computing Systems. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 3988–3999. DOI: http://dx.doi.org/10.1145/2858036.2858448

52. Alexandre N Tuch, Paul Van Schaik, and Kasper Hornbæk. 2016. Leisure and Work, Good and Bad: The Role of Activity Domain and Valence in Modeling User Experience. *ACM Transactions on Computer-Human Interaction (TOCHI)* 23, 6 (2016), 35.

53. Alexandre N. Tuch, Rune Trusell, and Kasper Hornbæk. 2013. Analyzing Users' Narratives to Understand Experience with Interactive Products. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 2079–2088. DOI: http://dx.doi.org/10.1145/2470654.2481285

54. Robert J. Vallerand. 1997. Toward A Hierarchical Model of Intrinsic and Extrinsic Motivation. Advances in Experimental Social Psychology, Vol. 29. Academic Press, 271 – 360. DOI: http://dx.doi.org/10.1016/S0065-2601(08)60019-2

55. Robert J. Vallerand, Michelle S. Fortier, and Frédéric Guay. 1997. Self-determination and persistence in a real-life setting: Toward a motivational model of high school dropout. *Journal of Personality and Social Psychology* 72, 5 (1997), 1161–1176. DOI: http://dx.doi.org/10.1037/0022-3514.72.5.1161

56. Paul van Schaik and Jonathan Ling. 2009. The role of context in perceptions of the aesthetics of web pages over time. *International Journal of Human-Computer Studies* 67, 1 (2009), 79 – 89. DOI: http://dx.doi.org/10.1016/j.ijhcs.2008.09.012

57. Symeon P. Vlachopoulos, Ermioni S. Katartzi, Maria G. Kontou, Frederiki C. Moustaka, and Marios Goudas. 2011. The revised perceived locus of causality in physical education scale: Psychometric evaluation among youth. *Psychology of Sport and Exercise* 12, 6 (2011), 583–592. DOI: http://dx.doi.org/10.1016/j.psychsport.2011.07.003

58. Jia Wei Zhang, Ryan T. Howell, and Peter A. Caprariello. 2013. Buying Life Experiences for the "Right" Reasons: A Validation of the Motivations for Experiential Buying Scale. *Journal of Happiness Studies* 14, 3 (2013), 817–842. DOI: http://dx.doi.org/10.1007/s10902-012-9357-z