Converging coolness and investigating its relation to user experience

Raptis, Dimitrios; Bruun, Anders; Kjeldskov, Jesper; Skov, Mikael

Published in:
Behaviour and Information Technology

DOI (link to publication from Publisher):
10.1080/0144929X.2016.1232753

Publication date:
2017

Document Version
Early version, also known as pre-print

Converging Coolness and Investigating its Relation to User Experience.

Dimitrios Raptis (raptis@cs.aau.dk)[1], Anders Bruun (bruun@cs.aau.dk)*[1], Jesper Kjeldskov (jesper@cs.aau.dk)[1] and Mikael B. Skov (dubois@cs.aau.dk)[1]

*: Corresponding author

[1]: Aalborg University, Dept. of Computer Science

Selma Lagerlöfs Vej 300, DK-9220 Aalborg Oest

Recently a number of studies appeared that operationalized coolness and explored its relation to digital products. Literature suggests that perceived coolness is another factor of user experience, and this adds to an existing explosion of dimensions related to aesthetics, hedonic quality, pragmatic quality, attractiveness, etc. A critical challenge highlighted in prior research is to study the relationships among those factors and so far, no studies have empirically examined the relationship between coolness and other established user experience factors. In this paper, we address this challenge by presenting two studies, one that focuses on factors from two cool questionnaires, and one that compares them against existing user experience factors. Our findings show that factors from the two cool questionnaires converge and they also converge to existing, established UX factors. Thus, eleven distinct cool and user experience factors converge into five for the case of mobile devices. Our findings are important for researchers, as we demonstrate through a validated model that coolness is part of user experience research, as well as for practitioners, by developing a questionnaire that can reliably measure both perceived inner and outer coolness as well as the overall coolness judgment based on five factors and 21 items.

Keywords: coolness, inner cool, outer cool, user experience, dimensionality explosion, questionnaires.

1 Introduction

For more than a decade User eXperience (UX) has been applied as a broad notion to describe experienced qualities of interactive products and user experience research focuses on exploring the experiential, affective, meaningful and valuable aspects of product use (Vermeeren et al., 2010). UX goes beyond the instrumental emphasis of usability (Bargas-Avila and Hornbæk, 2011) and although the satisfaction part of usability is considered as a relevant dimension for UX (Law et al., 2009), UX qualities are not limited to that. In this paper we focus on factors that, according to literature, contribute to coolness and we study, first how they converge and shape the cool perception and secondly how they relate to other subjective, measurable UX factors, such as affect, enjoyment, fun, aesthetics, appeal, attractiveness, hedonic quality, engagement, flow, enchantment and frustration (Bargas-Avila and Hornbæk, 2011).

The challenge for UX research is that the sheer volume of factors has reached such a large number where it is critical to start discussing the extent these are converging. Bargas-Avila and Hornbæk (2011) apply the term *"dimensionality explosion"* to denote this phenomenon within UX research. There is a need for the UX research community to study this explosion, for example, the relation between hedonic quality and attractiveness. Are we referring to the same or similar factor with two different names? In every context? For every product? Bargas-Avila and Hornbæk (2011) suggest that dimensionality explosion occurs firstly because many of these factors are not established as they have not been tested for their reliability and validity (*many researchers use self-made items without providing them*), and secondly because several proposed factors are not positioned in relation to the rest (*the main problem of this dimensionality explosion is that the relation to established constructs is rarely made clear)*. For example, no one to our knowledge has compared hedonic quality and attractiveness and produced specific results on if (or how) they converge into one factor. Besides this research challenge, this situation also creates significant problems for practitioners too, as there is no agreement on which questionnaires to use, and under which conditions.

A recent example of an emerging new perceived UX factor is coolness. In the past five years the HCI community has increased its focus on determining coolness of digital products, and “designing for cool” is becoming an essential criterion when developing new applications, interfaces and devices (Sundar et al., 2014). The main driving force behind this research effort was the fact that the term coolness has often been used by people to positively describe their experiences with various products such as cars, home appliances, mobile phones, etc. (Raptis et al., 2013).

Until recently, relatively few studies investigated coolness within HCI. Holtzblatt et al. (2011) discussed the concept of cool and emphasized that coolness contributes to our personal feelings of accomplishment, connection with others, identity, and delightful experiences. Read et al. (2012) developed a framework on “Being Cool”, by “Doing Cool Things” and by “Having Cool Stuff”. The aim of that framework is to facilitate the design of cool products for teenagers based on factors such as being rebellious, antisocial, retro, authentic, rich, and innovative. Culén and Gasparini (2012) argued that product coolness is related to fun, mastery, adding value, useful, successful, self-presentation and innovation. A more comprehensive overview on coolness can be found in Raptis et al. (2013).

In the above mentioned studies, the identified coolness characteristics were derived on the basis of literature reviews, which have made considerable contributions in shaping our initial understanding of coolness. McCrickard et al. (2012) moved a step further and explicated the need for a “cool engineering” approach to support designing for coolness. The aim of such an approach would be to understand how target users perceive coolness of products in various contexts of use. In order to define a cool engineering approach first there is a need to produce reliable tools and techniques for measuring coolness. Towards this end, two similar studies recently focused on breaking down the concept of cool into smaller entities and produced questionnaires that reliably measure perceived coolness (Sundar et al., 2014; Bruun et al., 2016). Sundar and colleagues (2014) produced a questionnaire with 15 items that measures coolness through factors related to subculture, attractiveness and originality. The COOL Questionnaire proposed by Bruun et al. (2016) consists of 16 items and it builds on a distinction between inner and outer coolness. Bruun et al. (2016) measure the perceived inner coolness of interactive products through factors related to usability, desirability and rebelliousness and they also suggest that perceived outer coolness is related to attractiveness and aesthetics.

Consequently, at this moment there are two questionnaires that measure coolness through differently labelled factors. Since some of these factors are seemingly comparable (e.g. subculture and rebelliousness) and by taking into consideration the problem of the dimensionality explosion, there is a need both to examine the extent of convergence between proposed coolness factors, but also to examine how coolness and existing UX factors converge. Thus, the aim of our paper is twofold:

1. To make a systematic comparison of the underlying factors of the two recently proposed coolness questionnaires,
2. To compare these coolness factors against established UX factors.

In this paper we report two studies, one for each aim. In the following we provide an overview of existing questionnaires for measuring perceived coolness and other UX related factors (sections 2 and 3). In section 4 we outline a set of hypotheses on how we expect the coolness factors to converge and how we expect them to converge with other UX factors. Section 5 describes the method, and then sections 6 and 7 highlight our results. In section 8 we discuss our findings against our research hypotheses as well as their implications for research and practice. Finally, in section 9 we conclude our paper by highlighting the most important findings.

2 Established Questionnaires for Measuring Coolness

In this section we present in detail two recently published questionnaires for measuring perceived coolness and their underlying factors.

2.1 "Capturing Cool"

In 2014 the paper "Capturing cool: Measures for assessing perceived coolness of technological products" was published by Sundar et al. The aim of that study was to produce a questionnaire for measuring coolness. Through literature, Sundar et al. (2014) identified a set of four factors that characterized coolness. The first factor is based on the work by Kerner et al. (2007) and Levy (2006) and relates to the uniqueness of a product over competing products. A second factor is related to product attractiveness, which has its roots within aesthetics. Its theoretical underpinning was based on the work of Levy (2006) and Tractinsky (1997). According to Sundar et al., attractiveness encompasses the externally visible aesthetic appeal, but is also related to the social acceptability of a given style, e.g. whether or not a product makes the owner look good in relation to others. Their third factor deals with the subcultural aspects of products. Based on the work by Dar-Nimrod et al. (2012) and Horton et al. (2012) the authors suggest that subculture includes an element of rebelliousness, e.g. that a product appeals to a minority (and not the mainstream consumer group) by being "edgy". According to Sundar et al. (2014), subculture also deals with the utility of a product for a particular group of persons. So, a product can be considered cool if it is useful for a specific group while indicating one's affiliation to this particular group. Their fourth and final factor relates to genuineness. Based on Conan (2008), Kerner et al. (2007) and Levy (2006) this factor is about authenticity and the sincere nature of a product. Thus, a product, of which the underlying intentions of its designer is to really improve the lives of its users, is perceived as genuine.

Sundar et al. applied these four factors as an offset to create 35 evaluative statements of coolness, e.g. “The designers of this product primarily want to create better products” (related to genuineness). An exploratory study with 315 participants was conducted and key factors within the 35 question items were extracted. This was followed by an additional two studies based on 1150 respondents assessing the coolness of various products such as USB drives, Nintendo Wii, Prezi, Warcraft etc. The statistical techniques of exploratory and confirmatory factor analyses were applied throughout the studies and Figure 1 shows their final three-factor structure of perceived coolness, which can be measured through the factors of subculture, attractiveness and originality and 15 items. Each item is represented on nine-item Likert scale (1=Strongly disagree, 9=Strongly agree). Furthermore, they demonstrated that these three factors do contribute to the overall coolness judgment for a product by comparing them to a set of 9 items that measured overall coolness.

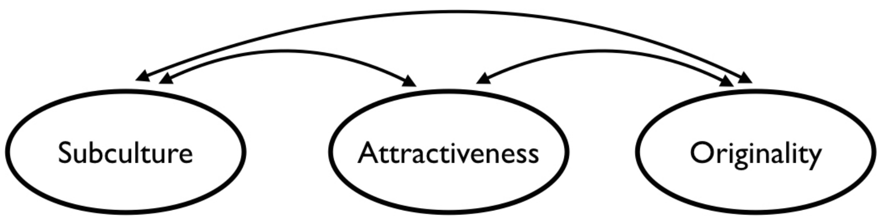

Figure 1: Three factor structure of cool (Sundar et al., 2014).

Subculture	Attractiveness	Originality
This device makes people who use it different from other people If I used this device, it would make me stand apart from others This device helps people who use it stand apart from the crowd People who use this product are unique People who use this device would be considered leaders rather than followers	This device is stylish This device is hip This device is sexy This device is hot This device is on the cutting edge	This device is original This device is unique This product is novel This device is out of the ordinary This product stands apart from similar products

Table 1: Factors and items from the "Capturing cool" questionnaire (Sundar et al., 2014).

2.2 "The COOL Questionnaire"

Bruun et al. (2016) similarly focused on breaking down the concept of cool into smaller entities and used them as building blocks to produce the "COOL Questionnaire"[1] (Bruun et al., 2016). The process is similar to what Sundar et al. (2014) used to create their "Capturing Cool" questionnaire, where factors are derived on a theoretical basis followed by a set of statistical studies.

Bruun et al. (2016) differ from the Sundar et al. (2014) study as they propose that coolness is decomposed to inner cool and outer cool. They base this distinction on a literature review presented in Raptis et al. (2013). According to Nancarrow et al. (2001) and MacAdams (2001), inner coolness deals with the personality of someone, i.e. how others perceive intra-person characteristics. As an example, a person can be perceived as cool if she keeps her calm under pressure. Inner cool in relation to products refers to the perceived personality traits, which are assigned to products by users, e.g. a product can be considered as cute or tough (Janler and Stolterman, 1997; Jordan 1997). Outer cool relates to how something or someone is presented through a certain style in physical appearance (Gioia 2009). For products this is a matter of aesthetic design, e.g. physical shape, materials, colors and so on.

[1] Tools for deploying the COOL questionnaire as well as for analyzing collected data can be found in: http://thecoolquestionnaire.weebly.com/

The questionnaire presented in Bruun et al. (2016) measures the perceived inner coolness of products, but not outer coolness. Authors suggest that outer cool is directly related to aesthetic attributes and thus it can be measured by existing UX factors, for example by using questionnaires that measure attractiveness or aesthetics such as the one proposed in (Lavie and Tractinsky, 2004). In a similar matter as the Sundar et al. (2014) study, Bruun et al. (2016) start the process of creating their questionnaire by identifying relevant characteristics that contribute to inner coolness. Informed by a literature review (Raptis et al., 2013) they identified eleven characteristics that contribute to inner coolness, namely being rebellious and antisocial, embracing authenticity and innovation, seeking exclusivity, pleasure and personal development, being/appearing in control, making hard things appear easy, being detached/emotionally neutral, and being strongly tight to a group. The majority of the eleven characteristics emerged from the work of Pountain and Robbins (2000), MacAdams (2001) and Nancarrow et al. (2001). For more details, we refer to Raptis et al. (2013).

Bruun et al. derived their questionnaire using an initial pool of 143 items related to the eleven inner cool characteristics. Through an iterative process with a total of 2236 respondents and by repeatedly applying the statistical techniques of exploratory and confirmatory factor analyses they ended up having a questionnaire with 16 items (Table 2) distributed to three factors of desirability, rebelliousness and perceived usability (Figure 2). All items are measured on a 7-point Likert scale (1=Strongly disagree, 7=Strongly agree). Additionally and similarly to Sundar et al. (2014), Bruun et al. also demostrated that these three factors shape the overall coolness judgment, which was measured through the item “This device is cool”.

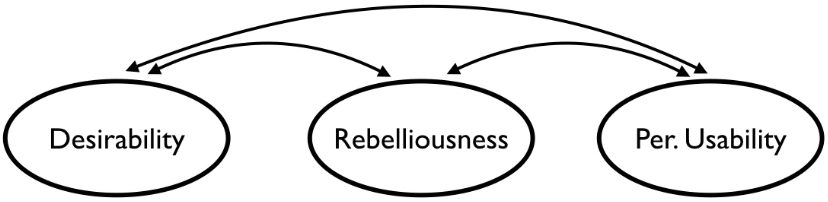

Figure 2: Three factor structure of inner cool (Bruun et al., 2016).

Desirability	Rebelliousness	Perceived Usability
This device can make me better This device is meant for people like me This device can make me happy This device can make me look good This device totally connects with me This device can make me look in control of things	This device moves against the current This device is unconventional This device is different This device is outside the ordinary This device is rebellious	This device is easy to operate This device is easy to use This device is easy to learn This device is simple to use This device is effortless to use

Table 2: Factors and items from the "COOL questionnaire" (Bruun et al., 2016).

3 Questionnaires for Measuring Established UX Factors

As mentioned in the introduction of the paper, Bargas-Avila and Hornbæk (2011) point towards the existence of a "dimensionality explosion" in relation to measuring UX. A critical question here is whether coolness further fuels this explosion and if its factors, to some extent, overlap with existing factors. In order to answer this question a set of established UX factors are presented in the following paragraphs.

Many widely considered UX factors concern the aesthetic appeal of interaction designs (Bargas-Avila and Hornbæk, 2011). Lavie and Tractinsky (2004) proposed a questionnaire for assessing the level of websites aesthetics. Their questionnaire has since then been applied to evaluate UX of various products, such as mobile phones (Sonderegger et al., 2012). The questionnaire is based on the two factors of classic and expressive aesthetics. Items of the classic aesthetics factor consist of a set of adjectives such as "Pleasant", "Clean" and "Symmetric". Lavie and Tractinsky (2004) state that this factor deals with traditional notions of aesthetics. The expressive aesthetics factor represents qualities that go beyond the classical design principles and includes items such as "Creative", "Fascinating" and "Sophisticated". Table 3 shows all the questionnaire items, which are rated on 7-point Likert scale (1=Strongly disagree, 7=Strongly agree).

Classic aesthetics	Expressive aesthetics
This device has:	*This device has*:
Aesthetic design	Creative design
Pleasant design	Fascinating design
Clear design	Use of special effects
Clean design	Original design
Symmetric design	Sophisticated design

Table 3: Factors and items from the aesthetics questionnaire (Lavie and Tractinsky, 2004).

Attrakdiff (Hassenzahl et al., 2003) is also a widely recognized questionnaire for measuring UX (Bargas-Avila and Hornbæk, 2011). Like the aesthetics questionnaire it was created with a focus on websites, but it has also been successfully applied to assess the UX of various types of products, e.g. culturally adaptive applications (Reinecke and Bernstein, 2011). The shortened version of this questionnaire (Attrakdiff2) is based on a two factor structure concerning hedonic and pragmatic qualities and two evaluative constructs (Van Schaik et al., 2012). The hedonic quality factor deals with the overall appeal of a product and includes items related to aesthetics (e.g. "I judge the product to be stylish") as well as items about excitement (e.g. "I judge the product to be captivating"). The pragmatic quality factor revolves around utilitarian and usability aspects with underlying items such as "I judge the product to be confusing/structured" or "I judge the product to be impractical/practical". All items (Table 4) are assessed on a 7-point scale (e.g. 1=dull, 7=captivating).

Hedonic quality	Pragmatic quality	Evaluative constructs
I judge the device to be:	*I judge the device to be:*	*I judge the device overall to be:*
Dull-Captivating	Confusing-Structured	Bad-Good
Tacky-Stylish	Impractical-Practical	Ugly-Beautiful
Cheap-Premium	Unpredictable-Predictable	
Unimaginative-Creative	Complicated-Simple	

Table 4: Factors and items from Attrakdiff2 questionnaire (Van Schaik et al., 2012).

Finally, attractiveness is also considered as an established UX factor and since it was identified as a core factor at the Sundar study we chose to include a reliable questionnaire that measures it. Quinn and Tran (2010) developed a five 7-point scale to measure attractiveness of a product and they used it to assess the attractiveness of mobile phones. The underlying items (Table 5) deal with aesthetic notions similar to those suggested in Lavie and Tractinsky (2004), but they also relate to the hedonic factor of Van Schaik et al. (2012). In order to differentiate between the attractiveness factor of Sundar et al. (2014) and attractiveness factor of Quinn and Tran (2010) we will refer to the first as attractiveness(cool) and the latter as attractiveness(UX).

Attractiveness (UX)
I judge the device to be: Attractive-Unattractive Beautiful-Ugly Eye catching-Plain Interesting-Boring I like the way this phone looks

Table 5: Items from the attractiveness questionnaire (Quinn and Tran, 2010).

4 Research Hypotheses

As pointed out by Bargas-Avila and Hornbæk (2011) when a new UX factor is proposed, the tools that measure it should be tested for their reliability and validity, and the factor should also be compared against other established UX factors. In this paper we report two studies that follow this suggestion. First we compared the two existing cool questionnaires alone (study 1), and then we compared them against the established UX factors presented in the previous section (study 2). In the following we discuss our hypotheses in relation to how cool factors and UX factors converge.

4.1 Study 1: Converging Factors of Coolness

At a first glance the factors of attractiveness (Sundar et al., 2014), desirability and usability (Bruun et al., 2016) seem different, i.e. they represent different aspects of coolness. Based on their items, attractiveness deals with aesthetic appeal (e.g. “This device is stylish”), while desirability relates to personal desires (e.g. “This device can make me happy”). Perceived usability is different from these as it concerns perceived learnability, utility and operability of the device (e.g. “This device is effortless to use”).

However, there is some overlap between the questionnaires. In particular, we find the originality (Sundar et al., 2014) and rebelliousness (Bruun et al., 2016) factors to be similar as they both deal with unconventional and novel aspects of a product, e.g. “This device is unconventional” vs. “This device is out of the ordinary”. The subculture factor from Sundar et al. relates more to the people using a product than the product itself. Yet, the topic of being different and unique is essential (e.g. “This device makes people who use it different from other people”).

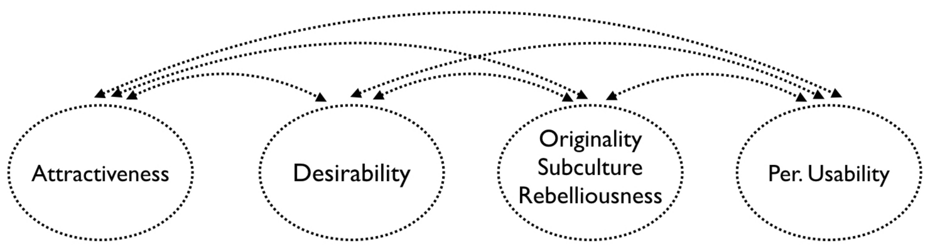

Figure 3: Hypothesized converging factors of cool when combining Sundar et al. (2014) and Bruun et al. (2016).

Figure 1 and Figure 2 present the two existing models for evaluating coolness of products, each with three factors. Based on the seemingly comparable factors of originality/subculture and rebelliousness, we hypothesize that a four factor structure would emerge when combining the items from the two questionnaires (see Figure 3). Thus, we hypothesize that the combination of items from Sundar et al. (2014) and Bruun et al. (2016) would lead to the following:

H1. Coolness can be measured through the four factors of attractiveness, desirability, perceived usability and originality/ subculture/rebelliousness.

4.2 Study 2: Converging Factors of Coolness and Established UX Factors

We will start this section with our hypotheses on overlapping factors among established

UX factors and we will continue with their relation to the cool factors. From Lavie and Tractinsky's classic and expressive factors there are items such as "Pleasant design", or "Fascinating design", which seem to overlap with items from Van Schaik et al.'s hedonic factor, e.g. "I judge the product to be captivating". These in turn are similar to the items presented in Quinn and Tran's attractiveness(UX) factor, e.g. "I judge the product to be interesting". Thus, the factors of classic aesthetics, expressive aesthetics, hedonic and attractiveness(UX) all deal with observable aesthetic characteristics. This is in line with Hassenzahl and Monk (2010) and Diefenbach et al. (2014), who argue that hedonic quality is similar to expressive aesthetics in specific contexts. The pragmatic factor in Van Schaik et al. (2012) seems to stand apart with items related to the perceived usability of a product, e.g. "I judge this product to be complicated".

Compared to the coolness questionnaires, we do see similar items to those posed in the established UX questionnaires. The attractiveness(cool) factor suggested in Sundar et al. (2014) deals with outer appearance and its items seem comparable to those from classic/expressive aesthetics/hedonic quality/attractiveness(UX) presented above. Also, the pragmatic quality factor of AttrakDiff2 relates to instrumental aspects of a product, which is similar to the perceived usability factor presented in Bruun et al.'s (2016) coolness questionnaire. As an example, consider the item "I judge the product to be simple" (Van Schaik et al., 2012) versus "This device is simple to use" (Bruun et al., 2016). The hypothesized structure of coolness in Figure 3 also suggests factors of originality/ subculture/rebelliousness and desirability. Respectively, these factors deal with unconventional notions of a product and personal desire, and, thus seem to be independent to other established UX factors.

On the basis of this discussion, we hypothesize that a four factor structure will emerge when comparing the suggested factors of coolness (Bruun et al., 2016; Sundar et al., 2014) and the established UX factors (Lavie and Tractinsky, 2004; Quinn and Tran, 2010; Van Schaik et al., 2012). This hypothesized four factor structure is shown in Figure 4 along with two-way arrows indicating suggested correlations between factors.

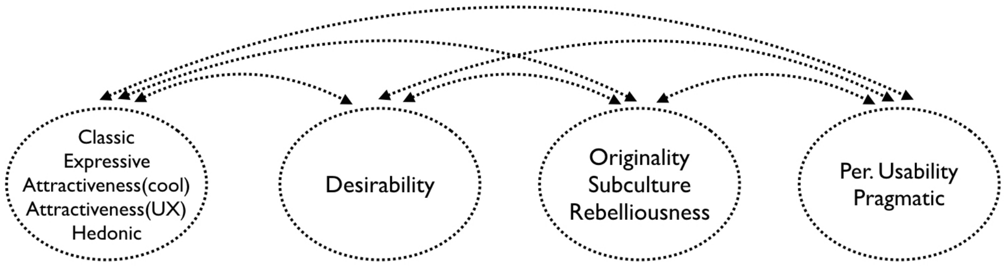

Figure 4: Hypothesized converging factors when combining coolness questionnaires (Bruun et al., 2016; Sundar et al., 2014) and established UX questionnaires (Lavie and Tractinsky, 2004; Quinn and Tran, 2010; Van Schaik et al., 2012). Attractiveness(UX) refer to the factor from (Quinn and Tran, 2010) and attractiveness(cool) refers to (Sundar et al., 2014).

Thus, our second hypothesis (H2) relates coolness to established UX factors:

H2a. The coolness factor of attractiveness(cool) converges on established UX factors of classic/expressive aesthetics/hedonic quality/ attractiveness(UX).

H2b. The coolness factor of perceived usability converges on the established UX factor of pragmatic quality.

H2c. The coolness factors of originality/subculture/rebelliousness and desirability do not converge on any of the established UX factors.

5 Method

In order to test our research hypotheses, we applied the statistical techniques of exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) using SPSS v. 23 and AMOS v. 22 respectively. EFA is based on an iterative process where items are removed from an initial pool of items, based on how much they contribute to measuring a particular factor. We conducted two EFA studies where we used the Bartlett's Test of Sphericity to test homogeneity of variances, the Kaiser-Meyer-Olkin Measure (KMO) to test sampling adequacy, and Principal Axes Factoring as extraction method using an oblique rotation (as recommended in literature, e.g. Bulmer (1979) and

Field (2009)). The number of the extracted factors was determined through a Scree test and through parallel analysis (using Monte Carlo PCA, Watkins). In the two EFA studies we removed items by applying two criteria: low communalities (<.5) and low factor loadings (<.65).

CFA is, as the name implies, of confirmatory nature and it is used to validate the factor structure that emerged though an EFA. In CFA there are also item loadings on factors, but also covariances between factors, denoting how variances between any two pairs of factors are correlated. The goodness of a factor model is determined by a range of fit-indices, which collectively indicate whether, or not the factor structure is appropriate and reliable (Schreiber et al., 2006). In the following sections we present the used indices for each CFA.

We conducted two CFAs to test our hypotheses based on Structural Equation Modelling (SEM) with Maximum Likelihood Estimation. When conducting SEM it is necessary to conduct a pre analysis to examine whether SEM assumptions are met in the data sample. These assumptions are related to missing data, normality, linearity and multicollinearity (Schreiber et al., 2006). We had no missing data and the CFA datasets had univariate normality with skewness values between -1 and .43 and kurtosis values between -1.4 and .6. These are within acceptable thresholds to assume that data is normally distributed (Tabachnick and Fidell, 2013). Due to the strong factor loadings (>.6) identified during all EFAs we also assume linearity between latent and manifest variables. The level of multicollinearity was also acceptable according to Kutner et al. (2004) with Variance Inflation Factor levels between 1.5 and 4.5. Therefore, in both CFA studies all the necessary assumptions were met.

In our studies we included a large number of participants to be able to do statistical analysis. We recruited these participants from Amazon Mechanical Turk

(MTurk). MTurk participants have been used successfully in other studies within HCI and have been shown to provide valuable results (e.g. Boujarwah et al., 2012; Heer and Bostock, 2010; Heimerl et al., 2012). We limited our selection to people living in the US to avoid language barriers and to follow the recommendations of Ross et al. (2010) and Huff and Tingley (2015). Ross et al. (2010) conducted a profiling study of MTurk workers and collected data for their gender, age, income, and level of education. Their findings show that the sample of US MTurk workers is balanced in relation to income and gender, while there are slightly more workers of younger age. Their education level is similar to that of the whole US population (OECD 2016). Huff and Tingley (2015) compared a large sample of US MTurk workers to a nationally representative sample. They focused on age, gender, race, ideology, occupation and the areas participants live. The MTurk sample was identical to the representative one in relation to ideology, occupation and the areas participants live, and slightly imbalanced in relation to age, gender and race. As an additional quality measure we only recruited MTurk workers with 95% approval ratings, as recommended by Ross et al. (2010).

The need for a large pool of participants that we satisfied through MTurk also informed our experimental setup. Since it is not possible to interact physically with MTurk workers and show them physical artifacts, we were inspired by other studies in HCI where participants made ratings based on images. For example, Lindgaard et al. (2006) used images of websites that were shown to the participants through a PC for 50 to 500 milliseconds. Their aim was to study how fast people shape a judgement for a website's visual appeal. Tractinsky (1997) used a projector and collectively showed images of different ATM layouts to his participants while asking them to rate their perceived usability and beauty. Hoegg et al. (2010) used images of cookware and electric mixers that varied aesthetically to test the belief that "what is beautiful is good".

In a similar manner to these previous studies, we chose as evaluation objects mobile devices and we created a website, which on the left side showed an image of a mobile device and on the right side listed the questionnaire items. Using images also allowed us to experimentally control for external parameters. The included 13 mobile devices were of the same color, were presented without any indication of their brand and had their screens turned off to exclude any effect from the operating system (Figure 5).

Figure 5. The thirteen mobile devices used in the studies.

2239 MTurk workers participated in our studies, used our website and filled in the two established cool questionnaires (Sundar et al., 2014; Bruun et al., 2016) and the three established UX questionnaires (Lavie and Tractinsky, 2004; Quinn and Tran, 2010; Van Schaik et al., 2012). All questionnaire items were presented randomly and were rated on a 7-point Likert scale from “strongly disagree” to “strongly agree. Each participant was asked to assess only one mobile device and participated in one EFA or CFA study. Participants were paid an incentive ranging between 0.25$ and 0.35$, which was in line with MTurk’s guidelines on how to fairly pay them. From the 2239 participants, we removed responses where participants had a considerably lower completion time than the average, which is in line with Kittur et al. (2008). We also removed all participants that reported prior experience with the mobile device they were asked to evaluate, as prior experience may significantly affect user experience factors, and in particular perceived usability (Sauro 2011). This left us with a sample of 1790 participants that was balanced in relation to gender (892 females), had a large variety of age groups (18 to 72 years old, M=33.6, SD=10.8) and included a variety of races (self-identified as Caucasian, African, Hispanic, Asian, Arab, etc.). 1251 of them participated in the EFAs and 539 in the CFAs, with an average of 156.4 participants per device in the EFAs and 107.8 participants per device in the CFA's.

6 Study 1: Convergence of Coolness Factors (H1)

In this section we present how we addressed our first research hypothesis (H1) by examining the convergence of factors from the two cool questionnaires.

6.1 EFA – Exploring the Coolness Factor Model

To study H1 we asked 822 participants to rate one mobile device each and we included three devices (274 participants per device). All participants rated the device based on the 56 question items, i.e. all items from all questionnaires (EFA1, Table 6).

Study	**n**	**Devices**	**i**
EFA1	822	Samsung Galaxy S6, Blackberry Priv, and Vertu Signature Touch	56

Table 6. EFA1. n=number of participants, i=number of items used as input.

Initially we examined the reliability of each individual factor. With the exception of pragmatic quality, all factors had exceptional Cronbach α scores, indicating that, for the case of mobile devices, all individual factors can be reliably measured through their respective items (Table 7).

Questionnaire	Factor	i	Cronbach α
Capturing cool (Sundar et al., 2014)	Subculture	5	.913
	Attractiveness	5	.911
	Originality	5	.921
Cool questionnaire (Bruun et al., 2016)	Desirability	6	.902
	Rebelliousness	5	.886
	Perceived Usability	5	.918
Aesthetics (Lavie and Tractinsky, 2004)	Classic Aesthetics	5	.869
	Expressive Aesthetics	5	.872
Attrakdiff2 (Van Schaik et al., 2012)	Hedonic Quality	4	.892
	Pragmatic Quality	4	.705
Attractiveness (Quinn and Tran, 2010)	Attractiveness	5	.937

Table 7. Reliability analysis of factors from all questionnaires, i=number of items.

To define the factor model of the two combined coolness questionnaires, we relied on exploratory factor analysis. Combined, the coolness questionnaires (Bruun et al., 2016; Sundar et al., 2014) consist of 31 of the total 56 items (with the remaining 25 representing items from the other established UX questionnaires). We removed items by applying the cut-off criteria (low communalities (<.5) and low factor loadings (<.65)), and in the end 18 items remained. The final four factor structure was identified through Scree Tests and Parallel Analysis (Monte Carlo PCA, Watkins). This model explained 74.84% of the variance, KMO was .930 and Bartlett's Test of Sphericity was significant.

Factor A contains items that emerged from attractiveness (Sundar et al., 2014) and Factor B mainly from rebelliousness (Bruun et al., 2016). Items from originality and subculture (Sundar et al., 2014) mostly converged on rebelliousness, but were removed using the cut-off criteria. Factor C is about usability (Bruun et al., 2016), while Factor D deals with desirability (Bruun et al., 2016). Details are presented in Table 8.

	Factor:	A	B	C	D
	Eigenvalue:	*7.361*	*1.780*	*3.518*	*0.813*
	Cronbach α:	*.911*	*.918*	*.912*	*.812*
Attractiveness	This device is stylish[2]	**.919**	.016	-.051	-.062
	This device is hip[2]	**.861**	-.030	-.027	.077
	This device is sexy[2]	**.780**	.011	.027	.040
	This device is hot[2]	**.779**	.023	.026	.133
	This device is on the cutting edge[2]	**.730**	.054	.153	.026
Perceived Usability	This device is simple to use[1]	-.065	**.916**	-.023	.010
	This device is easy to use[1]	.046	**.890**	.004	-.023
	This device is easy to learn[1]	-.055	**.882**	.030	-.014
	This device is easy to operate[1]	.001	**.882**	-.017	.034
	This device is effortless to use[1]	.084	**.759**	.004	.015
Rebelliousness	This device is different[1]	.062	-.002	**.887**	-.062
	This device is outside the ordinary[1,2]	.076	.022	**.861**	-.019
	This device is unconventional[1]	-.106	-.014	**.845**	.022
	This device is unique[2]	.181	.032	**.831**	-.077
	This device moves against the current[1]	-.093	-.019	**.824**	.137
Desirability	This device can make me better[1]	-.059	-.005	.043	**.917**
	This device can make me happy[1]	.099	.089	-.088	**.778**
	This device can make me look in control of	.234	.004	.102	**.618**
	Sum of Squared Loadings (Total variance explained): 74.84%				

Table 8. Pattern Matrix with item loadings per factor in EFA1 containing only the cool questionnaires. A= Attractiveness, B=Perceived Usability, C=Rebelliousness, D=Desirability. [1] originates from Bruun et al. (2016), [2] originates from Sundar et al. (2014).

6.2 CFA1 – Confirming the Coolness Factor Model

In order to confirm the four-factor structure of coolness as emerged from EFA1, we conducted a CFA study. This included the 18 items that emerged from EFA1 and 206 participants rated one device each (Table 9).

Study	n	Devices	i
CFA1	206	Apple iPhone 6s Plus, and Huawei Ascend Y530	18

Table 9. CFA1. n=number of participants, i=number of items used as input.

In the first run all indices suggested acceptable values, i.e. it was not necessary to go through modification indices to increase model fit. Table 10 shows the respective item loadings and model-fit indices obtained in CFA1 where all loadings are significant. We also validated the model by examining the matrix of standardized residuals. A model with a good fit will have residuals centered around zero and we found none larger than ±2, hereby indicating a good model fit (Schreiber et al., 2006).

	CFA1 – Coolness Factor Model	
Attractiveness	This device is stylish[2]	.74
	This device is hip[2]	.83
	This device is sexy[2]	.76
	This device is hot[2]	.75
	This device is on the cutting edge[2]	.72
Perceived Usability	This device is simple to use[1]	.91
	This device is easy to use[1]	.86
	This device is easy to learn[1]	.87
	This device is easy to operate[1]	.88
	This device is effortless to use[1]	.72
Rebelliousness	This device is different[1]	.83
	This device is outside the ordinary[1,2]	.84
	This device is unconventional[1]	.63
	This device is unique[2]	.79
	This device moves against the current[1]	.63
Desirability	This device can make me better[1]	.78
	This device can make me happy[1]	.86
	This device can make me look in control of things[1]	.72
Model Fit Indices		
Ratio of $\chi 2$ to df (CMIN/df, acceptance threshold $\leq$ 3)		1.5
Normed Fit Index (NFI, acceptance threshold $\geq$ .95)		.92
Incremental Fit Index (IFI, acceptance threshold $\geq$ .95)		.97
Tucker-Lewis Index (TLI, acceptance threshold $\geq$ .95)		.96
Comparative Fit Index (CFI, acceptance threshold $\geq$ .95)		.97
Goodness-of-Fit Index (GFI, acceptance threshold $\leq$ .95)		.9
Adjusted Goodness-of-Fit Index (AGFI, acceptance threshold $\leq$ .95)		.87
Root Mean Square Error of Approx. (RMSEA, accept. threshold $\leq$ .06)		.05
p of close fit (PCLOSE, acceptance threshold > .05)		.43

Table 10. Item loadings per factor and model-fit indices for the CFA1 study. All are within acceptable thresholds, indicating good model-fit. [1] originates from Bruun et al. (2016), [2] originates from Sundar et al. (2014)

Based on the model-fit indices we found that our data supports a four-factor model representing coolness. Table 11 shows the correlation matrix between the four factors. The diagonal elements in bold represent the square root of average variance extracted

(AVE) as and the Cronbach α values (in parentheses). Since the square roots of AVEs are bigger than all factor correlations we can conclude that the discriminant validity is more than adequate. The same is the case with internal consistency (construct reliability) with high Cronbach α values.

	Attractiveness	Per. Usability	Rebelliousness	Desirability
Attractiveness	**.761 (.874)**			
Per. Usability	.35	**.850 (.923)**		
Rebelliousness	.55	.2	**.750 (.861)**	
Desirability	.74	.42	.41	**.789 (.829)**

Table 11. Factor correlation matrix for CFA1. Values in bold indicate the square root of average variance extracted (AVE) and Cronbach α (in parentheses).

6.3 Study 1 Results

In this section we present our results in relation to the research hypothesis H1:

H1. Coolness can be measured through the four factors of attractiveness, desirability, perceived usability and originality/subculture/rebelliousness.

We confirm this hypothesis. Findings from the EFA1 and CFA1 show the existence of 18 items distributed over four factors for measuring perceived coolness: attractiveness, perceived usability, rebelliousness and desirability. Attractiveness stems exclusively from Sundar et al. (2014) while perceived usability and desirability stems from Bruun et al. (2016). The rebelliousness factor includes items from both questionnaires. Thus, we confirm H1 for the case of assessing perceived coolness of mobile devices.

7 Study 2: Convergence of Coolness on UX Factors

In the previous section we confirmed the factor structure related to coolness where the total of six factors from Sundar et al. (2014) and Bruun et al. (2016) converged on four factors. In the following we examine the emerging factor structure when combining the factors identified across the coolness studies as well as the other established UX factors related to aesthetics (Lavie and Tractinsky, 2004), attractiveness(UX) (Quinn and Tran, 2010), and hedonic and pragmatic quality (Van Schaik et al., 2012). By examining the convergence of these factors we seek to test the second set of hypotheses (H2a, H2b and H2c). In the following we present our findings from two EFA studies exploring possible factor structures describing the relation between coolness and established UX factors.

7.1 EFA – Exploring the Cool-UX Factor Model

We started our analysis using the same dataset as before (EFA1, Table 6), but this time we included all 56 items (31 from the cool questionnaires and 25 from the other established UX questionnaires). Through an exploratory factor analysis, we produced a total of five models with four and five factor structures. Items were removed by applying the cut-off criteria and only if they did not belong to any factors in any model. In all five models the KMO was >.925, fulfilling the criteria for sampling adequacy and Bartlett's Test of Sphericity was significant (<.001). In the end, the initial 56 items were reduced to 33.

Given the remaining relatively large number of items, we chose to conduct an additional EFA study (EFA2). In EFA2 we had 429 additional participants rate one device each and we included a total of 5 devices (~86 participants per device, Table 12).

Study	n	Devices	i
EFA2	429	HTC One M8, OnePlus One, Tag Heuer Meridiist, Nokia 222, Philips Fluid	33

Table 12. EFA2. n=number of participants, i=number of items used as input.

The 429 participants in EFA2 rated the remaining 33 items and a five-factor structure was identified through Scree Tests and Parallel Analysis (Monte Carlo PCA, Watkins). By applying the cut-off criteria, the number of items was reduced from 33 to 22 and our final model had a KMO of .954. Cumulatively this five-factor model explained 79.39 % of the variance. In Table 13 we present the emerged five factor structure after EFA2, along with the loadings of each item on factors.

	Factor:	A	B	C	D	E
	Eigenvalue:	*7.862*	*5.184*	*2.284*	*1.346*	*0.790*
	Cronbach α:	*.946*	*.933*	*.921*	*.790*	*.905*
Hedonic	I find this device: plain/eye catching[3]	**.909**	-.003	.127	-.005	-.079
	I judge this device to be: cheap/premium[4]	**.879**	-.082	-.092	-.073	-.006
	I judge this device to be: dull/captivating[4]	**.865**	.010	-.009	.026	.123
	I find this device: boring/interesting[3]	**.851**	.015	.035	-.045	.092
	I judge this device to be: unimaginative/creative[4]	**.767**	.037	.236	-.009	.006
Perceived Usability	This device is simple to use[1]	-.004	**.924**	-.003	.002	-.033
	This device is easy to use[1]	-.005	**.924**	-.006	.010	-.018
	This device is easy to operate[1]	.058	**.919**	-.057	-.035	-.039
	This device is easy to learn[1]	-.106	**.842**	.084	-.103	.005
	This device is effortless to use[1]	.024	**.799**	-.025	.034	.118
Rebelliousness	This device moves against the current[1]	-.039	.106	**.910**	.059	-.097
	This device is outside the ordinary[1,2]	.057	-.009	**.866**	-.013	.051
	This product stands apart from similar products[2]	.015	-.033	**.833**	-.068	.116
	This device is different[1]	.093	-.076	**.815**	-.013	.046
	This device is unconventional[1]	.071	-.069	**.802**	.015	-.014
Classic Aesthetics	This device has clear design[5]	-.065	.043	.055	**-.922**	.001
	This device has clean design[5]	.116	.032	-.090	**-.840**	.016
Desirability	This device can make me better[1]	-.101	.035	.015	.091	**.938**
	This device can make me look in control of things[1]	-.023	-.047	.120	-.112	**.810**
	This device can make me look good[1]	.067	-.111	.086	-.127	**.779**
	This device can make me happy[1]	.138	.132	-.110	-.005	**.755**
	This device totally connects with me[1]	.289	.115	-.114	.003	**.678**
	Sum of Squared Loadings (Total variance explained): 79.39%					

Table 13. Pattern matrix with item loadings per factor in EFA2. A= Attractiveness, B=Perceived Usability, C=Rebelliousness, D=Classic Aesthetics and E= Desirability. [1] originates from Bruun et al. (2016), [2] originates from Sundar et al. (2014), [3] originates from Quinn and Tran (2010), [4] originates from Van Schaik et al. (2012), [5] originates from Lavie and Tractinsky (2004).

Throughout the EFA studies we observed a trend on how items converged on the five factors. Items from the two attractiveness factors (Sundar et al., 2014; Quinn and Tran, 2010), the expressive aesthetics (Lavie and Tractinsky, 2004) and the hedonic quality (Van Schaik et al., 2012) converged around Factor A (hedonic). Furthermore, items from pragmatic quality (Van Schaik et al., 2012) converged on Factor B (perceived usability) with low factor loadings, which led to their removal. Thus, Factor B consists of items from the perceived usability factor identified in Bruun et al. (2016). Items from subculture and originality (Sundar et al., 2014) and rebelliousness (Bruun et al., 2016) converged on Factor C (rebelliousness). Items from the classic aesthetics factor (Lavie and Tractinsky, 2004) solely define Factor D (classic aesthetics) with no convergence on other factors. Finally, a few specific questions from expressive aesthetics, classic aesthetics and attractiveness converged on Factor E (desirability) with low factor loadings. Thus, Factor E is defined by items from desirability (Bruun et al., 2016).

7.2 CFA – Confirming the Cool-UX Factor Model

To confirm the proposed five-factor model from EFA2, we had 333 participants rate one mobile device, and we included three different devices in total (CFA2, Table 14), i.e. each device was assessed by 111 participants on average. None of the participants had taken part in the previous studies.

Study	n	Devices	i
CFA2	333	Nexus 6P, North Face M8, and Blackberry Classic	22

Table 14. CFA2. n=number of participants, i=number of question items used as input.

To obtain an acceptable model fit we went through four iterations where we removed one item at a time based on the largest modification indices. After that point all indices suggested a good model fit, i.e. it was not necessary to go through further iterations. Through these iterations we removed four items from the 22 identified in EFA2. Thus, the final CFA model consists of 18 items. Table 15 shows the respective item loadings and model-fit indices obtained in CFA2 where all loadings are significant. We also validated the model by examining the matrix of standardized residuals. A model with a good fit will have residuals centered around zero and we found none larger than ±2, hereby indicating a good model fit (Schreiber et al., 2006).

	CFA2 – Cool-UX Factor Model	
Hedonic	I find this device: plain/eye catching[3]	.89
	I judge this device to be: dull/captivating[4]	.88
	I find this device: boring/interesting[3]	.90
	I judge this device to be: unimaginative/creative[4]	.84
Perceived Usability	This device is simple to use[1]	.87
	This device is easy to use[1]	.88
	This device is easy to operate[1]	.91
	This device is easy to learn[1]	.83
Rebelliousness	This device moves against the current[1]	.7
	This device is outside the ordinary[1,2]	.86
	This product stands apart from similar products[2]	.79
	This device is different[1]	.87
Classic aesthetics	This device has clear design[5]	.73
	This device has clean design[5]	.76
Desirability	This device can make me better[1]	.78
	This device can make me look in control of things[1]	.74
	This device can make me look good[1]	.86
	This device can make me happy[1]	.78
Model Fit Indices		
Ratio of $\chi 2$ to df (CMIN/df, acceptance threshold ≤ 3)		1.9
Normed Fit Index (NFI, acceptance threshold $\geq .95$)		.95
Incremental Fit Index (IFI, acceptance threshold $\geq .95$)		.98
Tucker-Lewis Index (TLI, acceptance threshold $\geq .95$)		.97
Comparative Fit Index (CFI, acceptance threshold $\geq .95$)		.97
Goodness-of-Fit Index (GFI, acceptance threshold $\leq .95$)		.93
Adjusted Goodness-of-Fit Index (AGFI, acceptance threshold $\leq .95$)		.9
Root Mean Square Error of Approx. (RMSEA, accept. threshold $\leq .06$)		.05
p of close fit (PCLOSE, acceptance threshold $> .05$)		.42

Table 15. Item loadings per factor and model-fit indices for the CFA2 study. All are within acceptable thresholds, indicating good model-fit. [1] originates from Bruun et al. (2016), [2] originates from Sundar et al. (2014), [3] originates from Quinn and Tran (2010), [4] originates from Van Schaik et al. (2012), [5] originates from Lavie and Tractinsky (2004).

Table 16 presents the correlation matrix between the five factors, which shows that none of the factors have a 1-1 correlation. The diagonal elements in bold represent the square root of average variance extracted (AVE) as well as the Cronbach α values (in parentheses). Since the elements exceed all factor correlations except one, discriminant validity is adequate. That said, the hedonic and desirability factors do have a correlation of .79, which indicates that these factors are closely related and one can be used to predict the other. Nevertheless, all factors are consistently separated throughout our EFA and CFA studies, i.e. they are measuring different UX factors. Furthermore, in relation to internal consistency (construct reliability) Cronbach α values are high, which shows that the items can reliably measure the five factors.

	Hedonic	Per.Usability	Rebelliousness	Classic Aesthetics	Desirability
Hedonic	**.878 (.929)**				
Per. Usability	.29	**.873 (.927)**			
Rebelliousness	.66	.14	**.806 (.878)**		
Classic Aesthetics	.47	.66	.14	**.745 (.712)**	
Desirability	.79	.41	.48	.55	**.791 (0.868)**

Table 16. Factor correlation matrix for CFA2. Diagonal values in bold indicate: the square root of average variance extracted (AVE) and Cronbach α (in parentheses).

7.3 Study 2 Results

Our second hypothesis related coolness to other established UX factors and is divided in three parts (H2a, H2b and H2c). We will address each hypothesis individually.

H2a. The coolness factor of attractiveness(cool) converges on established UX factors of classic/expressive aesthetics/hedonic quality/attractiveness(UX).

This hypothesis is not supported. As in previous studies (Hassenzahl and Monk, 2010; Diefenbach et al., 2014) our results showed that expressive aesthetics and hedonic quality converge. We also identified that these also converged with the two attractiveness factors (Quinn and Tran, 2010; Sundar et al., 2014). At the same time, classic aesthetics (Lavie and Tractinski, 2004) formed an independent factor. Thus, even though we expected all five factors to be merged into one, they merged into two: A) a hedonic factor which consisted of two question items from Quinn and Tran's (2010) attractiveness factor, and two items from Van Schaik et al.'s (2012) hedonic quality factor, and B) a classic aesthetics factor that emerged solely from Lavie and Tractinsky (2004). For these reasons we falsify hypothesis H2a.

H2b. The coolness factor of perceived usability converges on the established UX factor of pragmatic quality.

Table 15 shows that the emerged perceived usability factor consists exclusively of question items stemming from the perceived usability factor as suggested by Bruun et al. (2016). This happened because even though all items from Van Schaik et al.'s (2012) pragmatic quality factor consistently followed the perceived usability items, they all had loadings below the cut-off level. Additionally, since we did not observe any of the pragmatic items converging on any other factor, we verify hypothesis H2b.

H2c. The coolness factors of originality/subculture/rebelliousness and desirability does not converge on any of the established UX factors.

Finally, in relation to H2c both factors of rebelliousness and desirability that emerged during the EFA studies were retained in the five-factor model from CFA2. Items from originality and subculture did converge on the rebelliousness factor, but with lower loadings. Thus, hypothesis H2c is also verified.

8 Discussion

Our purpose with this paper was twofold. Firstly, we wanted to make a systematic comparison of the underlying factors between the recently proposed questionnaires for measuring perceived coolness. Thus, in our first study we examined the extent of convergence between the suggested coolness factors and we demonstrated that there are differences as well as overlaps. Secondly, we also wanted to position the coolness factors in relation to established UX ones and thus contribute to the dimensionality explosion challenge (Bargas-Avila and Hornbæk, 2011). Through our research effort in our second study we managed to combine all factors into a single model, thus we strengthened the position of coolness within UX research. In the following subsections we discuss implications for research and practice.

8.1 Implications for UX research

When studying the two cool questionnaires (Sundar et al., 2014: Bruun et al., 2016) we produced a model that narrowed down the initial six factors into four (Table 8: attractiveness, perceived usability, rebelliousness and desirability). When we followed the same process by including established UX questionnaires, a new model emerged that contained not eleven, but five factors (Table 15: hedonic quality, classic aesthetics, desirability, perceived usability, and rebelliousness). What we believe is interesting for our research community is to understand the relation among these factors.

In this part of the discussion we contribute to understanding the relation among these factors by linking back to theory. According to literature, when people observe a product (or a person) they immediately make a judgement on its overall coolness (Pountain and Robbins, 2000). When there is no actual usage with a product, as was the case in our study, this judgement is initially based on the externally observable aesthetic attributes (outer coolness), which people use to infer a judgement of personality characteristics (inner coolness). Then, both inner coolness and outer coolness shape the overall judgment of coolness (Pountain and Robbins, 2000; Raptis et al., 2015; Bruun et al., 2016). Figure 6 shows the theoretical relation among perceived inner coolness, perceived outer coolness and overall coolness judgment.

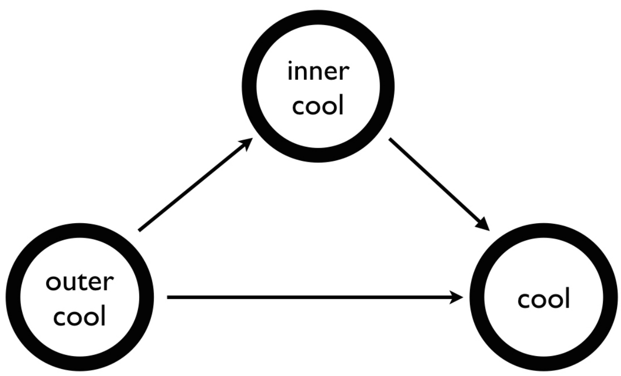

Figure 6. Theoretical relationships among perceived inner cool, perceived outer cool and overall coolness judgement.

We took this theoretical model one step further by including the emerged five factors of hedonic quality, classic aesthetics, desirability, usability and rebelliousness. We hypothesize a UX inference model in a similar manner as Hassenzahl and Monk (2010) and Van Shaik et al. (2012), who showed what we perceive as beautiful is also perceived as good, which in turn is perceived as usable, i.e. *"What is beautiful is good and what is good is usable"*. In order to validate our proposed inference model we used inferential statistics (Partial Least Squares, Vinzi et al., 2010). Based on our analysis, we argue that hedonic quality and classic aesthetics contribute to outer coolness, since they both relate to the aesthetic attributes of a product, while perceived usability, rebelliousness and desirability contribute to inner cool. In the Appendix the final validated inference model is presented along with standardized regression coefficients, T-statistics, the percentage of explained variance, as well as, details on the process followed. In the following figure, we present a simplified version of this model that depicts how the emerged five factors are clustered around inner cool and outer cool, and how they both contribute to the overall coolness judgment.

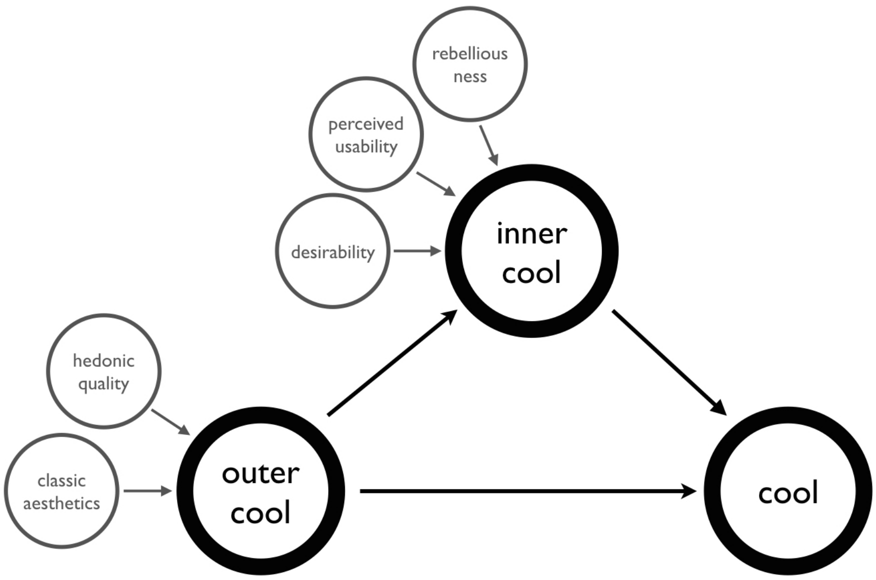

Figure 7. Simplified inference model showing that hedonic quality and classic aesthetics cluster around outer cool, while desirability, rebelliousness and perceived usability cluster around inner cool. Both perceived outer and inner cool shape the overall coolness judgment. All paths are significant (details in the Appendix).

The proposed inference model of users' experiences with mobile devices through outer, inner and overall cool (Figure 7) is a valid and useful tool for researchers. Firstly, it demonstrates the existence of an inference rule: *"The perception of product aesthetics influences perceived product personality characteristics, and both shape the overall coolness judgement"*. Secondly, it shows that people do infer inner cool from outer cool when they believe that it is a relevant rule for the situation (for example, during first impression with mobile devices, i.e. without actual usage). Thirdly, it demonstrates that overall coolness cannot be inferred only by factors related to externally observable attributes. Inner coolness is also needed. Fourthly, our inference model strongly positions coolness within user experience research, and it can be used to explain and/or predict the relationship among the five emerged factors.

Another important implication of our research work, is that we showed that the established user experience factors we use in our research community converge. This leads to some factors being more prevalent than others while others converge, e.g. pragmatic quality and perceived usability. This reduces the number of factors that are relevant to consider. This finding taps directly into the UX community discussion of the dimensionality explosion. Bargas-Avila and Hornbæk (2011) for instance mention: *"the main problem of this dimensionality explosion is that the relation to established constructs is rarely made clear"*. Our study provides a way of dealing with this through the proposed measurement model for the case of assessing perceived UX of mobile devices.

Of course, further research is needed for other products than mobile devices, for different contexts of use (for example after long term usage), and for different cultural groups, in order to test the applicability of inferring inner-cool from outer-cool rule and its performance (e.g. is the direction of the effect always the same?). Furthermore, we find of particular interest to compare our inference model with existing ones, e.g. (Van Schaik et al., 2012), as such comparisons may provide answers to important research questions like: Are these models applicable in all contexts? For all products? Furthermore, an important research activity that we believe our community needs to pursue is to continue the study of convergence of the rest of UX factors. Do other established UX factors that we did not include in our studies (such as pleasure) converge to the five we managed to identify in this paper? Are there other influential UX factors which are currently unknown?

Finally, we believe there is a need for more research on the relation between hedonic quality and classic aesthetics. Since both of them emerged as unique and distinct factors and both measure aesthetic attributes, it is crucial to understand what they actually measure. We consider this challenge as important since it was also identified in previous research work (Bruun et al., 2016). Based on our findings we propose that the two items of classic aesthetics that remained in our study (clean/clear design) are related to the cognitive process of recognition, while the four hedonic items relate to an intentional (or even unintentional) evaluation process of a product's aesthetic appeal that occurs afterwards. As an example, we believe our participants recognized the product they experienced in our study as a mobile device, and then they evaluated its appeal. More studies with this emphasis are needed in order to verify our assumption and we consider them as important as they may shed light on the cognitive mechanisms people use to aesthetically evaluate our produced designs.

8.2 Implications for UX Practice

Our findings also have implications for practice. Firstly, we re-confirmed that all five established questionnaires we included in our studies can reliably measure their pertinent factors for the case of perceived UX of mobile devices (Table 7). The interesting issue though for practitioners is how to use these questionnaires in practice.

If practitioners want to measure a specific perceived UX factor, e.g. the perceived usability of a produced design, then any of these questionnaires that measures perceived usability can be used, as it will provide reliable results. A challenge though exists if the purpose is to have a holistic evaluation of a product's perceived attributes. In such cases practitioners should administer combinations of all these questionnaires in efforts to assess various aspects of UX. However, measuring all UX factors would mean that participants will have to answer a relatively large set of questions (56 in case of the questionnaires included in this study alone). Our findings help practitioners deal with dimensionality explosion and the large number of question items through the proposed five-factor model (Figure 7) and its items (Table 15). We demonstrated that the two established UX factors that converge on outer coolness (hedonic quality and classic aesthetics) can be measured using 6 items. Inner coolness can be measured through 11 items from the Cool Questionnaire (Bruun et al., 2016) and 1 item from the Capturing Cool questionnaire (Sundar et al., 2014). Finally, the overall coolness judgment can be measured by 3 items (details on how they emerged can be found in the Appendix). Thus, instead of answering 56 items which are related to several perceived UX factors that practitioners do know that they overlap, participants can answer the more manageable 21 items, which belong to factors that are independent. The final questionnaire that measures both perceived inner and outer coolness and the overall coolness judgment as well as their resulting items is shown in Figure 8.

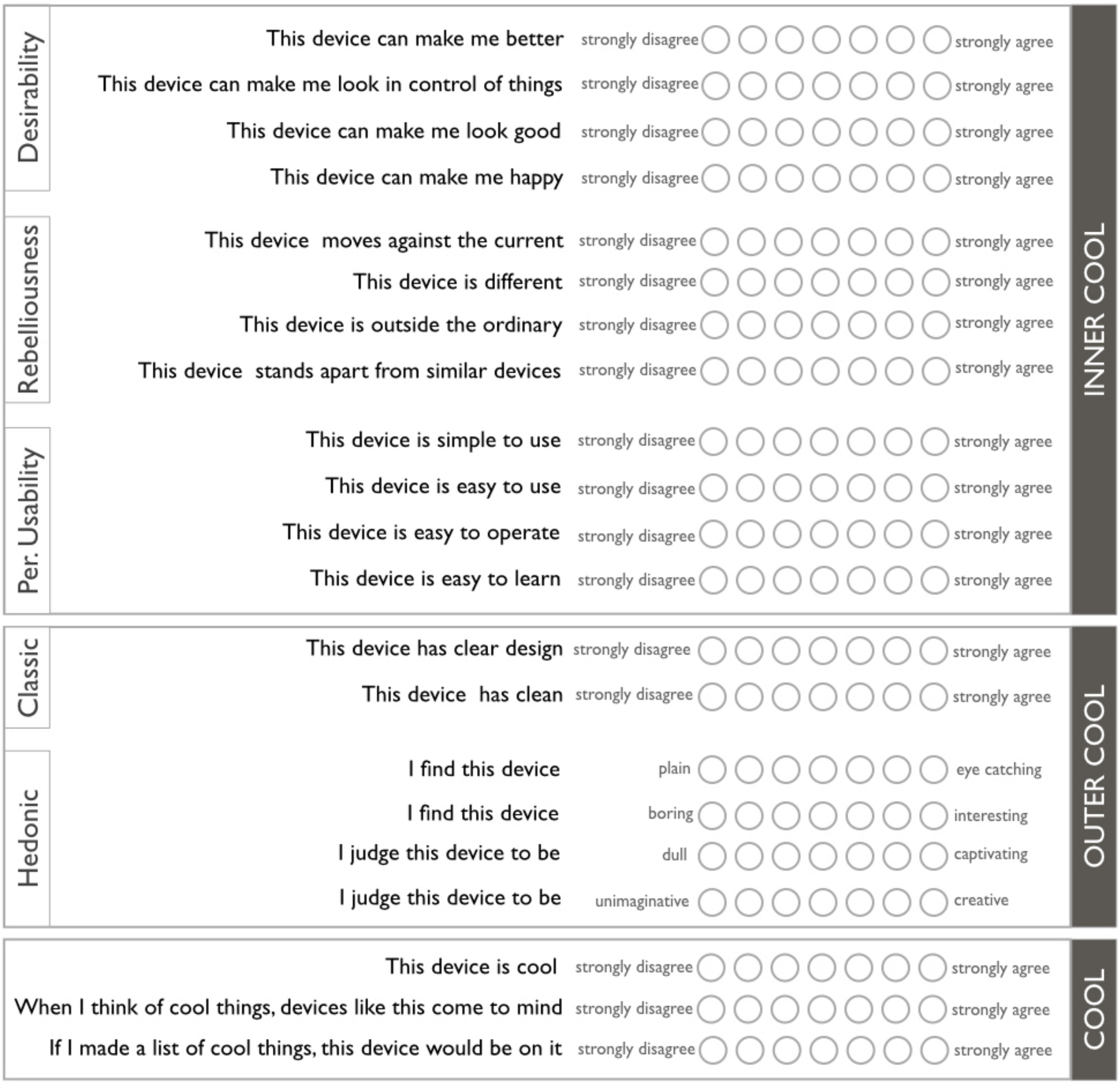

Figure 8. Final perceived inner cool and outer cool questionnaire. Measured inner cool and out cool factors and their items as well as the three evaluative items for the overall coolness judgment. The word "device" can be replaced with a suitable product.

Furthermore, we argue that our questionnaire/model will be more useful to practitioners if combined with qualitative methods. For example, if a product scores low on rebelliousness and this has a negative impact on its overall coolness, then practitioners can through, e.g. in-depth interviews, identify specific design elements in relation to rebelliousness that need to be changed, produce re-designs and then re-evaluate them. Thus, coolness becomes an essential design criterion, which not only can be measured, but also understood in relation to other established UX factors.

Finally, we define three challenges for our combined questionnaire/model that practitioners need to be aware of, which should be researched more in order to increase its applicability. The first is related to its performance. At this moment, we do not know what it means for a product to score, e.g. 5 on usability or 3.5 on rebelliousness, i.e. we do not know if such scores are good or bad for a particular product. In order to understand the behavior of the model, research approaches that were used for other questionnaires in the past, should be applied. For example, for the SUS scale Bangor et al. (2008) conducted a meta-analysis of previous studies, and concluded on the meaning of an SUS score in relation to a product's usability. Secondly, since coolness is deeply rooted to the cultural communities people belong to (O'Donnell and Wardlow, 2000), we need to test the model's behavior in different communities. For example, we may have different results in Asian, or European cultures, but also within different subcultures. Finally, we need to study the model's behavior in relation to time. For example, it is known that users who interact with products for long periods of time, change their perception of usability (Sonderegger et al., 2012). To what extent do the rest of the factors have similar behavior? Such knowledge can be extremely useful for practitioners as it will allow them to understand how users' experiences develop over time, hereby leading to decisions, e.g. on when to re-design products.

9 Conclusions

In this paper we explored how a large set of cool and user experience factors converge. Our paper contributes to the dimensionality explosion discussion (Bargas-Avila and Hornbæk, 2011) in two ways. First, we showed that the two existing questionnaires for measuring perceived coolness converge on four factors (Table 8). Secondly, we established coolness within user experience research by comparing it against established

UX factors. Our resulting model shows that eleven distinct cool and UX factors (Table 7) converge on five (Figure 7 and Table 15).

Our research identified a number of implications for researchers and practitioners. In relation to user experience research, first we positioned coolness within UX by demonstrating how it relates to established UX factors. Additionally, we moved a step further by proposing an inference model, which is based on the emerged five factors and has a strong theoretical foundation that distinguishes between outer cool (the perceived aesthetic characteristics of a product) and inner cool (the perceived personality characteristics of a product). The model can be used to explain and/or predict people's judgmental mechanisms in relation to perceived coolness of mobile devices.

In relation to practice, first we demonstrated that each of the deployed questionnaires in our two studies are reliable. Second, we compared the two cool questionnaires and produced a valid tool for measuring perceived inner coolness (Table 10). Finally, we produced a questionnaire (Figure 8), that measures both perceived inner and perceived outer coolness through 5 distinct factors with 18 items, and 3 evaluative items for the overall cool judgement. Both the model as well as the final questionnaire can be used to holistically assess users' experiences with mobile devices.

One of the limitations of our study is that participants were asked to evaluate images and not the actual mobile phones. As others have also used this approach, it would be interesting to investigate if the use of images as test object leads to differences in results when compared to the use of physical artifacts. Furthermore, we believe that both our model and its resulting questionnaire is applicable to other types of digital artifacts, as it was the case with most established UX questionnaires (such as Attrakdiff2), by replacing the word "device" with the product under evaluation.

Nevertheless, further research is needed to test them in different conditions and for different products, to verify their generalizability and applicability.

10 Acknowledgements

We would like to thank Dr. XXXXX XXXXXX for her invaluable help as well as all the people that participated in our studies.

11 References

Bangor, A., Kortum, P., and Miller, J., 2008. An empirical evaluation of the system usability scale. *International Journal of Human-Computer Interaction*, 24 (6), 574–594.

Bargas-Avila, J.A., and Hornbæk, K., 2011. Old Wine in New Bottles or Novel Challenges? A Critical Analysis of Empirical Studies of User Experience. In *Proc. CHI 2011*, 2689-2698.

Bruun, A., Raptis, D., Kjeldskov, J., and Skov, M. B., 2016. Measuring the Coolness of Interactive Products: The COOL Questionnaire. *Behaviour & Information Technology*, 35 (3), 233-249.

Boujarwah, F., Abowd, G., and Arriaga, R., 2012. Socially Computed Scripts to Support Social Problem Solving Skills. In *Proc. CHI 2012*, 1987-1996.

Bulmer M.G., 1979. *Principles of Statistics*. Courier Dover Publications.

Chin, W.W., 1998. The partial least squares approach to structural equation modelling. In G.A., Marcoulides (ed.), Modern Methods for Business Research. Lawrence Erlbaum Associates, 295–336.

Chin, W.W., Marcolin, B.L., and Newsted, P.N., 2003. A partial least squares latent variable modeling approach for measuring interaction effects: Results from a monte carlo simulation study and an electronic-mail emotion/adoption study. *Information Systems Research,* 14, 2, 189–217.

Conan, N., 2008. Behind the viral video: what's fake, what's real. National Public Radio.

Culén, A.L., and Gasparini, A.A., 2012.Situated Techno-Cools: factors that contribute to making technology cool in a given context of use. *PsychNology*, 10 (2), 117-139.

Dar-Nimrod, I., Hansen, I.G., Proulx, T., Lehman, D.R., Chapman, B.P., Duberstein, P.R., 2012. Coolness: an empirical investigation. *Journal of Individual Differences* 33 (3), 175–185.

Diefenbach, S., Kolb, N., and Hassenzahl, M., 2014. The 'Hedonic' in Human-Computer Interaction – History, Contributions, and Future Research Directions. In *Proc. DIS 2014,* ACM Press, 305-314.

Falk, R.F., and Miller, N.B., 1992. *A primer for soft modelling*. The University of Akron Press, Ohio.

Field, A., 2009. *Discovering Statistics using SPSS*. SAGE Publications.

Gioia, T., 2009. *The birth and death of the Cool*. Speck Press.

Hassenzahl, M., Burmester, M., and Koller, F., 2003. Attrakdiff: Ein Frageboten zur Messung wahrgenommener hedonischer und pragmatischer Qualität. *Mensch & Computer 2003. Interaction in Bewegung*, 187-196.

Hassenzahl, M. and Monk, A., 2010. The inference of perceived usability from beauty. *Human-Computer Interaction 25*(3), 235–260.

Heer, J., and Bostock, M., 2010. Crowdsourcing Graphical Perception: Using Mechanical Turk to Assess Visualization Design. In *Proc. CHI 2010*, 203-212.

Heimerl, K., Gawalt, B., Chen, K., Parikh, T. and Hartmann, B., 2012. Community Sourcing: engaging local crowds to perform expert work via physical kiosks. *In Proc CHI 2012*.

Henseler, J., Ringle, C., and Rayner, R., 2009. The use of partial least squares modelling in international marketing. In T. Cavusgil, R. Sinkovicks, and P. Ghauri (eds), New Challenges in International Marketing, Emerald, 277-319.

Hoegg, J.A., Alba, J.W., and Dahl, D.W. 2010. The good, the bad, and the ugly: Influence of aesthetics on product feature judgments. *Journal of Consumer Psychology*, 20 (4), 419-430.

Holtzblatt, K., 2011. What makes things cool? Intentional design for innovation. *Interactions,* 18 (6), 40-47.

Horton, M., Read, J.C., Fitton, D., Toth, N., and Little, L., 2012. Too Cool at School-Understanding Cool Teenagers. *PsychNology*, 10 (2), 73-91.

Huff, C., and Tingley, D. 2015. "Who are these people?" Evaluating the demographic characteristics and political preferences of MTurk survey respondents. *Research and Politics*, 2 (3), 1-10.

Janlert, L.E., and Stolterman, E., 1997. The character of things. *Design Studies*, 18 (3), 297–314.

Jordan, P.W., 1997. Products as personalities. In (ed.) M.A. Hanson, *Contemporary Ergonomics*, 449-453.

Jöreskog, K.G., and Wold, H., 1982. The ML and PLS techniques for modelling with latent variables: Historical and comparative aspects. In H. Wold and K.G. Jöreskog (eds), *Systems under indirect observation: Causality, structure, prediction*, 1, 263-270.

Kerner, N., Pressman, G., and Essex, A., 2007. *Chasing Cool: Standing out in Today's Cluttered Marketplace*. Atria, New York.

Kittur, A., Chi, E.H., and Suh, B. 2008. Crowdsourcing user studies with Mechanical Turk. In *Proc. CHI 2008*, ACM Press, 453-456.

Kutner, M. H., Nachtsheim, C. J., and Neter, J. 2004. *Applied Linear Regression Models 4th ed*. McGraw-Hill, Irwin.

Lavie, T., and Tractinsky, N., 2004. Assessing dimensions of perceived visual aesthetics of web sites. *Int. J. Human-Computer Studies*, 60, 269-298.

Law, E. L.-C., Roto, V., Hassenzahl, M., Vermeeren, A. and Kort, J., 2009. Understanding, Scoping and Defining User eXperience: A Survey Approach. In *Proc. CHI*, ACM Press, Boston, 719-728.

Levy, S., 2006. *The Perfect Thing: How the iPod Shuffles Commerce, Culture, and Coolness*. Simon & Schuster, New York.

Lindgaard, G., Fernandes, G., Dudek, C., and Brown, J. 2006. Attention web designers: You have 50 milliseconds to make a good first impression!. *Behaviour & Information Technology*, 25 (2), 115-126.

Lowry P.B., and Gaskin, J., 2014. Partial least squares (PLS) structural equation modeling (SEM) for building and testing behavioral causal theory: When to choose it and how to use it. *IEEE Transactions on Professional Communication*, 57 (2), 123-146.

MacAdams, L., 2001. *Birth of the Cool. Beat, Bebop, and the American Avant-Garde*. The Free Press.

McCrickard, D.S., Barksdale, J., and Doswell, F., 2012. Understanding Cool: An analytic exploration of contributing factors for teens. *PsychNology*, 10 (2), 93-102.

Nancarrow, C., Nancarrow, P., and Page, J., 2001. An analysis of the concept of cool and its marketing implications. *Journal of Consumer Behaviour*, 1 (4), 311-322.

OECD. 2016. Population with tertiary education. Available online: DOI:10.1787/0b8f90e9-en

O'Donnell, K.A., and Wardlow, D.L., 2000. A Theory on the origins of Coolness. *Advances in Consumer Research*, 27, 13-18.

Pountain, D., and Robbins, D., 2000. *Cool Rules. Anatomy of an Attitude*. Reaction Books.

Quinn, J.M., and Tran, T.Q., 2010. Attractive phones don't have to work better: independent effects of attractiveness, effectiveness, and efficiency on perceived usability. In *Proc. CHI 2010*, 353-362.

Raptis, D., Kjeldskov, J., and Skov, M.B., 2013. Understanding "Cool" in Human-Computer Interaction Research and Design. In *Proc. OzCHI 2013*, 53-62.

Read, C.R., Horton, M., and Fitton, D., 2012. Being Cool-Getting Persona-l. *CHI Workshop on Cool aX Continents, Cultures and Communities 2012*.

Reinecke, K. and Bernstein, A., 2011. Improving Performance, Perceived Usability, and Aesthetics with Culturally Adaptive User Interfaces. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 18, 2, Article No. 8.

Ringle, C.M., Wende, S., and Will, A., 2006. *SmartPLS 2.0*. University of Hamburg, Hamburg.

Ross, J., Irani, L., Silberman, M.S., Zaldivar, A., and Tomlinson, B. 2010. Who are the crowdworkers?: shifting demographics in mechanical turk. In *Proc. CHI EA 2010*, ACM Press, 2863-2872.

Sauro, J., 2011. Does Prior Experience Affect Perceptions Of Usability? Available online: http://www.measuringu.com/blog/prior-exposure.php

Schreiber, J.B., Nora, A., Stage, F.K., Barlow, E.A. and King, J., 2006. Reporting Structural Equation Modeling and Confirmatory Factor Analysis Results: A Review. *Journal of Educational Research,* 99, 6, 323-338.

Sobel, M.E., 1982. Asymptotic Confidence Intervals for Indirect Effects in Structural Equation Models. *Sociological Methodology*, 13, 290-312.

Sonderegger, A., Zbinden, G., Uebelbacher, A., and Sauer, J., 2012. The influence of product aesthetics and usability over the course of time: a longitudinal field experiment. *Ergonomics*, 55 (7), 713-730.

Sundar, S.S., Tamul, D.J., and Wu, M., 2014. Capturing “cool”: Measures for assessing coolness of technological products. *Int. J. Human-Computer Studies*, 72, 169-180.

Tabachnick, B. G., & Fidell, L. S., 2013. *Using Multivariate Statistics 6th ed*. Pearson Education Inc., New Jersey.

Tenenhaus, M., Vinzi, E.V., Chatelin, Y.M., and Lauro, C., 2005. PLS Path Modeling. *Computational Statistics & Data Analysis*. 48, 1, 159-205.

Tractinsky, N. 1997. Aesthetics and apparent usability: empirically assessing cultural and methodological issues. In *Proc. CHI 1997*, ACM Press, pp. 115–122.

Van Schaik, P., Hassenzahl, M., and Ling, J., 2012. User-Experience from an Inference Perspective. *TOCHI,* 19, 2, Article No. 11.

Vermeeren, A.P.O.S., Law, E.L.C., Roto, V., Obrist, M., Hoonhout, J., Väänänen-Vainio-Mattila, K., 2010. User Experience Evaluation Methods: Current State and Development Needs. In *Proc. NordiCHI 2010*, 521-530.

Vinzi, V.E, Chin, W.W., Herseler, J. and Wang, H., 2010. *Handbook of Partial Least Squares: Concepts, Methods and Applications in Marketing and Related Fields*. Springer.

Watkins, M.W. *Monte Carlo PCA for Parallel Analysis*. Online: http://edpsychassociates.com/Watkins3.html

Wilson, B., 2010. *Using PLS to Investigate Interaction Effects Between Higher Order Brand Constructs*. In V. Esposito Vinzi et al. (eds), Handbook of Partial Least Squares: Concepts, Methods and Applications in Marketing and Related Fields. Springer, Berlin.

Appendix: Developing the PLS inner cool - outer cool inference model

In order to test the inferential relationship between outer cool, inner cool and the overall coolness judgment, we applied a Partial Least Squares (PLS) analysis (Vinzi et al., 2010). This is an ideal technique for models with high complexity and when a theoretical understanding of a domain needs to be tested (Jöreskog and Wold, 1982; Falk and Miller, 1992).

In order to have enough data to conduct PLS we merged the data from EFA1, EFA2 and CFA2 (1584 participants) and we focused only on final 18 items that emerged from the CFA2 (Table 15). These had to be compared with an overall coolness judgment. Bruun et al. (2015) the measure overall coolness judgment through the item "This device is cool", while Sundar et al. (2014) do the same using the same item and 8 additional ones. All nine of them were measured throughout all EFAs and CFAs. In order to narrow down their number we conducted a reliability analysis on the overall coolness judgment, which resulted to 3 items: "When I think of cool things, devices like this come to mind", "This device is cool", and "If I made a list of cool things, this device would be on it" (Cronbach α=.904).

The dataset was analyzed using the SmartPLS v2.0 software (Ringle et al., 2006). Since we had two second-level formative factors (perceived inner and outer cool) we used the hierarchical components approach, which is the most popular when estimating higher order factors with PLS (Chin et al., 2003; Tenenhaus et al., 2005; Wilson 2010). PLS produces the standardized regression coefficients (path estimates) between factors and we assessed the significance of path coefficients through bootstrap analysis (with N=5000, as proposed by Henseler et al., 2009).

In the first step, we applied the standardized regression coefficients between desirability, perceived usability and rebelliousness to reflect the inner cool factor, followed by the coefficients between hedonic and classic aesthetics to reflect the outer cool factor. The second step in the process was to analyze coefficients between the inner cool and outer cool factors on the overall coolness judgment for a mobile device. The final inference model can be found the following figure.

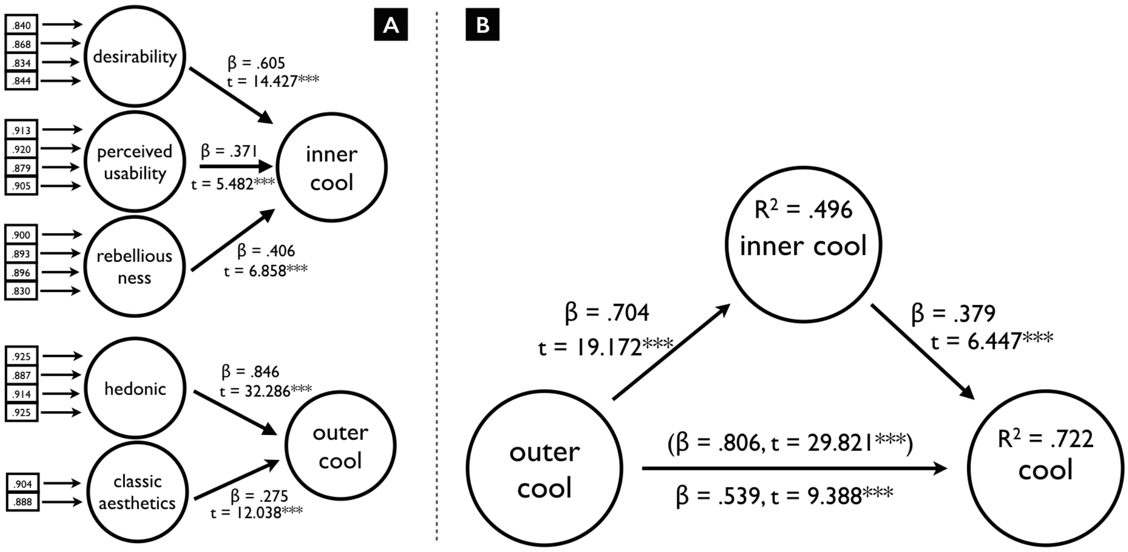

Figure 9. A) PLS measurement and structural model for the first-order formative constructs of outer cool and inner cool with factor loadings per item. B) PLS structural model for the second-order formative constructs of inner and outer cool and the overall judgment about the coolness of a mobile device. Values in parentheses indicate effects without inner cool. β stands for standardized regression coefficients, t for T-statistic, and R^2 for percentage of explained variance. *** p<.001.

Two parameters are usually applied for assessing the goodness of such models: the strength of each path coefficient and the percentage of explained variance (R^2). All path coefficients in our model were statistically significant ($p<.001$) and we had one substantial and one moderate R^2 value (Chin 1988). As a last step in the process we tested the significance of the mediation effect of inner cool, i.e. whether inner cool could be excluded from the model, through a Sobel test (Sobel 1982) as recommended by Lowry and Gaskin (2014). The Sobel test value (6.206) was statistically significant ($p<.001$), which means that inner cool partially mediates outer cool in determining the overall coolness judgment of a product, and thus it cannot be ignored.

Figure and Table List

Table 2: Factors and items from the "COOL questionnaire" (Bruun et al., 2016).

Table 3: Factors and items from the aesthetics questionnaire (Lavie and Tractinsky, 2004).

Table 4: Factors and items from Attrakdiff2 questionnaire (Van Schaik et al., 2012).

Table 5: Items from the attractiveness questionnaire (Quinn and Tran, 2010).

Table 6. EFA1. n=number of participants, i=number of items used as input.

Table 7. Reliability analysis of factors from all questionnaires, i=number of items.

Table 8. Pattern Matrix with item loadings per factor in EFA1 containing only the cool questionnaires. A= Attractiveness, B=Perceived Usability, C=Rebelliousness, D=Desirability. [1] originates from Bruun et al. (2016), [2] originates from Sundar et al. (2014).

Table 9. CFA1. n=number of participants, i=number of items used as input.

Table 10. Item loadings per factor and model-fit indices for the CFA1 study. All are within acceptable thresholds, indicating good model-fit.

Table 11. Factor correlation matrix for CFA1. Values in bold indicate the square root of average variance extracted (AVE) and Cronbach α (in parentheses).

Table 12. EFA2. n=number of participants, i=number of items used as input.

Table 13. Pattern matrix with item loadings per factor in EFA2. A= Attractiveness, B=Perceived Usability, C=Rebelliousness, D=Classic Aesthetics and E= Desirability. [1] originates from Bruun et al. (2016), [2] originates from Sundar et al. (2014), [3] originates from Quinn and Tran (2010), [4] originates from Van Schaik et al. (2012), [5] originates from Lavie and Tractinsky (2004).

Table 14. CFA2. n=number of participants, i=number of question items used as input.

Table 15. Item loadings per factor and model-fit indices for the CFA2 study. All are within acceptable thresholds, indicating good model-fit. [1] originates from Bruun et al. (2016), [2] originates from Sundar et al. (2014), [3] originates from Quinn and Tran (2010),

[4] originates from Van Schaik et al. (2012), [5] originates from Lavie and Tractinsky (2004).

Table 16. Factor correlation matrix for CFA2. Diagonal values in bold indicate: the square root of average variance extracted (AVE) and Cronbach α (in parentheses).