Taylor & Francis
Taylor & Francis Group

Behaviour & Information Technology

ISSN: 0144-929X (Print) 1362-3001 (Online) Journal homepage: http://www.tandfonline.com/loi/tbit20

Measuring the coolness of interactive products: the COOL questionnaire

Anders Bruun, Dimitrios Raptis, Jesper Kjeldskov & Mikael B. Skov

BEHAVIOUR & INFORMATION TECHNOLOGY, 2016
VOL. 35, NO. 3, 233–249
http://dx.doi.org/10.1080/0144929X.2015.1125527

Taylor & Francis
Taylor & Francis Group

Measuring the coolness of interactive products: the COOL questionnaire

Anders Bruun, Dimitrios Raptis, Jesper Kjeldskov and Mikael B. Skov

Department of Computer Science, Aalborg University, Selma Lagerlöfs Vej 300, DK-9220 Aalborg Oest, Denmark

ABSTRACT

Coolness has recently started to be explored as a design goal for interactive products from practitioners as well as researchers within human–computer interaction (HCI), but there is still a need to further operationalise the concept and explore how we can measure it. Our contribution in this paper is the COOL questionnaire. We based the creation of the questionnaire on literature suggesting that perceived coolness is decomposed to outer cool (the style of a product) and inner cool (the personality characteristics assigned to it). In this paper, we focused on inner cool, and we identified 11 inner cool characteristics. These were used to create an initial pool of question items and 2236 participants were asked to assess 16 mobile devices. By performing exploratory and confirmatory factor analyses, we identified three factors that can measure the perceived inner coolness of interactive products: desirability, rebelliousness and usability. These factors and their underlying 16 question items comprise the COOL questionnaire. The whole process of creating the questionnaire is presented in detail in this paper and we conclude by discussing our work against related work on coolness and HCI.

ARTICLE HISTORY
Received 25 January 2015
Accepted 19 November 2015

KEYWORDS
Perceived coolness; inner cool; outer cool; COOL questionnaire; attractiveness; hedonic quality; pragmatic quality; aesthetics

1. Introduction

Coolness is frequently used to describe approbation of various products. Cars, shoes, home appliances and interactive products are often characterised as being cool, or perceived as being cool, as, for example, several Apple products. Since coolness is used to characterise many products, designing for coolness has become increasingly important over the last years for practitioners and researchers. Coolness is now being integrated into interactive design processes and became a part of interaction design and human–computer interaction (HCI). As pointed out by Sundar, Tamul, and Wu (2014), coolness in interactive products has become an essential psychological criterion for designers and developers when creating new systems, applications, interfaces or devices.

In relation to coolness there are at least three challenges that need to be addressed. Firstly, coolness as a goal of a design process is somewhat new and various notions on coolness exist. Recently, few studies started to investigate coolness as a concept within interaction design, for example, Holtzblatt, Rondeau, and Holtzblatt (2010), Holtzblatt (2011), Fitton et al. "Climbing the Cool Wall" (2012), Fitton et al. "Constructing the Cool Wall" (2012), Read et al. (2011), Read, Horton, and Fitton (2012) and McCrickard, Barksdale, and Doswell (2012). These papers present studies investigating coolness for a variety of interactive products and within different contexts of use. These studies constitute the first steps in the process of understanding coolness, but at the same time coolness is not as well defined as other product qualities, such as usability (ISO 9241-11, 1998) or pleasure (Jordan 1997). Thus, we see a challenge in defining coolness in relation to interactive products.

Secondly, in order to apply coolness as a design goal, interaction designers need to be able to assess the coolness of design ideas. Thus, there is a need for specific tools and techniques that will allow us to measure the coolness of prototypes/products. In the past, this was the case for other product qualities too. For example, regarding usability and user experience (UX), the HCI community has developed theories, guidelines, methods and principles and provided design practitioners with an array of tools and constructs. Currently in the 'cool' literature there are few papers that provide such tools; for example, there are studies that outline techniques for identifying and analysing coolness through card sorting (Fitton et al., "Climbing the Cool Wall," 2012; Fitton et al., "Constructing the Cool Wall," 2012), or one questionnaire that measures coolness (Sundar, Tamul, and Wu 2014). A challenge here is to develop not only reliable and valid instruments to measure coolness, but also instruments that can be quickly and easily used by practitioners during product development.

Thirdly, since the scope is broadened by the fact that coolness is being explored as a design goal, there is a

CONTACT Anders Bruun bruun@cs.aau.dk

challenge in positioning coolness in relation to other established qualities such as usability and hedonic qualities. These are known from other established questionnaires, yet we know little of the extent to which constructs overlap. Such a challenge is important for researchers as it allows for a more detailed understanding of coolness, as well as for practitioners as they need to know what concepts they address in their designs when using different tools.

In this paper, we attempt to address these challenges and our main contribution is the design and construction of the COOL questionnaire, a questionnaire suitable for assessing the coolness of interactive products. Informed by previous literature on coolness, we derive a series of key characteristics that describe the essence of coolness, addressing the first challenge. Based on this, we present the process of developing the questionnaire using the statistical techniques of exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). The COOL questionnaire is the result of this process, which addresses the second challenge. The questionnaire is presented in detail in Section 4 along with the detailed explanations on the conception of its items (Sections 5 and 6). Finally, the third challenge is addressed as we compare our COOL questionnaire with other established UX questionnaires (Section 6). We conclude our paper by discussing our work in relation to that of others, its relation to UX and interaction design, its generalisability as well as avenues for further research.

2. Related work

Within HCI and interaction design, a number of research papers investigated the notion of cool, its meaning and its relation to interactive product design. This research effort has contributed to the initial understanding of the term cool both conceptually and methodologically and also presented different ways of working with it in design. At the same time due to its novelty, there are also studies that advise caution in introducing coolness to design processes (Cowan, Avramides, and Beale 2013).

In 2011, Holtzblatt (2011) discussed the concept of cool and stressed that cool products bring joy to our lives and contribute to our personal feelings of accomplishment, connection with others, identity and delightful experiences. Methodologically Holtzblatt presents 'The Wheel of Joy' and 'The Triangle of Design' as tools for defining the aspects of life and experience that designers should focus on when designing for cool. Additionally, in an early study on cool in HCI, Read, Horton, and Fitton (2012) developed a conceptual framework of 'Being Cool', by 'Doing Cool Things' and by 'Having Cool Stuff' facilitating the design of cool artefacts and interactive products for teenagers, displaying characteristics such as being rebellious, antisocial, retro, authentic, rich and innovative (Horton et al. 2012; Read, Horton, and Fitton 2012). These characteristics were derived through a detailed literature review and the same was the case with Culén and Gasparini (2012) where they argued that perceived coolness is affected by factors such as fun, mastery, adding value, useful, successful, self-presentation and innovative. As a next step, many studies took these characteristics, explored their applicability in different contexts and studied their relations. For example, McCrickard, Barksdale, and Doswell (2012) investigated challenges and opportunities in designing cool user interfaces for young people and concluded that the characteristics identified by Read, Horton, and Fitton (2012) and Horton et al. (2012) are suitable for understanding coolness, but they also elicit different reactions in people from different demographics (age, gender, expertise, etc.). Furthermore, De Guzman (2012) and Fitton et al. "Climbing the Cool Wall" (2012) and Fitton et al. "Constructing the Cool Wall" (2012) presented techniques for identifying and analysing coolness through 'Cool Card Sorting', or mapping out people's coolness preferences using the 'Cool Wall'.

As identified in many of the previous papers, one next step in incorporating coolness to interaction design processes is to define ways to measure it. For example, McCrickard, Barksdale, and Doswell (2012) highlighted the need for a 'cool engineering' approach, which would enable design activities, such as understanding how target users think, evaluating designs and in the end, producing cool interactive products. Towards this end, three studies recently appeared where experiments were conducted, with the aim either to produce tools that measure coolness, or to understand its relation to other constructs.

Sundar, Tamul, and Wu (2014) conducted a literature review as a starting point to identify potential coolness factors and to derive a questionnaire. From an exploratory study with 315 participants, they extracted key factors that contributed to coolness. These factors were then transformed to questions and two more studies were performed in order to reduce the number of question items. In this study by Sundar, Tamul, and Wu (2014), images of a variety of software and hardware artefacts (USB drive, Wii, Prezi, Warcraft, etc.) were presented to a large number of participants (1150 in total) and the statistical techniques of EFA and CFA were applied. The end result of the Sundar, Tamul, and Wu (2014) paper is a questionnaire, which concludes that the coolness of a product can be conceptualised as a matter of originality, attractiveness and subcultural appeal.

Similarly, Farnsworth, Holtzblatt, Pai, et al. (2014) and Farnsworth, Holtzblatt, Held, et al. (2014) describe the process of creating a questionnaire for measuring coolness, without publishing the questionnaire itself. They created their questionnaire by conducting several studies that included 900 consumers and over 2000 business professionals from around the world, over a 3-year period, and they argue that cool can be measured by 7 constructs: accomplishment, connection, identity, sensation, direct into action, the hassle factor and the learning delta.

Warren and Campbell (2014) investigated the relation between coolness and the concept 3 of autonomy, as well as the factors that may affect it. Coolness in their study was measured using two scales ('how cool or uncool do you consider the design' and 'how cool or uncool would your friends consider the design'). Autonomy was measured using an established questionnaire. In their six studies, the participants rated everyday objects (such as bottle designs) and brands (using logos). The authors concluded that brands as well as objects that diverge from the norm (have increased autonomy) in a way that seems appropriate are perceived as cool (Warren and Campbell 2014).

All the above research studies have contributed considerably to our initial understanding of cool, how to design for it and how to measure it. However, these papers have also strongly highlighted the need for further research. Although early conceptualisations have been put forward, cool still remains a very difficult concept to operationalise within interaction design as it is difficult to grasp and measure. As a result, it is very difficult to design for. In response to this challenge our efforts will, similar to those of Sundar, Tamul, and Wu (2014), focus on breaking down the concept of cool into smaller entities and then using these systematically as building blocks in order to produce a questionnaire that can measure the perceived coolness of an interactive product. In order to decompose coolness, we conducted an extensive literature review focusing on the domains of interaction design, marketing and the music/movie industry (Raptis, Kjeldskov, and Skov 2013). The aim of the review was to understand what coolness is and to identify its basic characteristics. Through this process we have identified three core principles of coolness, which shaped our experimental design: (1) Coolness consists of *inner* and *outer cool*, (2) Coolness is recognised immediately and (3) Coolness is grounded in people's communities (Raptis, Kjeldskov, and Skov 2013).

2.1. Coolness consists of inner and outer cool

According to the literature, coolness can be divided into inner cool and outer cool (MacAdams 2001; Nancarrow, Nancarrow, and Page 2001). Inner cool is about someone's or something's personality or character. Outer cool is about how something presents itself through a certain style in physical appearance.

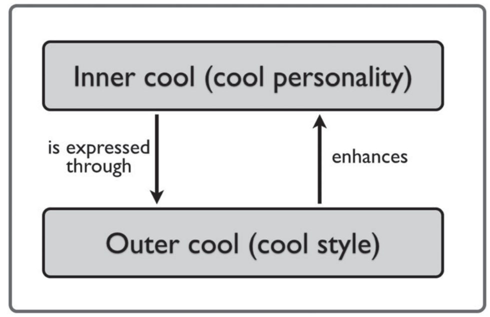

Figure 1. Inner cool and outer cool.

Inner cool for people relates to the way somebody's personality is perceived. Inner cool for objects refers to the perceived personality traits that people assign to them in order to help them understand and relate to them (Janlert and Stolterman 1997; Jordan 1997). Outer cool for people is typically a matter of somebody's clothes, accessories, language and pose (Gioia 2009). For objects it is a matter of similar aesthetic qualities in their design, for example, physical form, materials, lines and colours. Consequently, if someone believes that in the movie 'Rebel Without a Cause' James Dean is cool, this does not only refer to the way he looks (outer cool), but also to his perceived personality in that movie (inner cool). Additionally, when an object is characterised as cool, for example, an iPhone, this does not only refer to its physical appearance (outer cool), but also to its perceived character as a design object (inner cool).

Inner and outer cool are interrelated, where outer cool is typically expressed through cool style and inner cool conversely enhances the coolness perception of a person's or an object's perceived personality and character (Figure 1).

From our literature review on cool (Raptis, Kjeldskov, and Skov 2013), a number of characteristics have emerged for inner as well as outer cool. Although we consider outer cool as very important, we chose to focus on inner cool. The inner cool characteristics we have identified in the literature are presented in Table 1 and were used as the starting point in the process of creating our COOL questionnaire.

2.2. Coolness is recognised immediately

The literature generally agrees that coolness is something we perceive and recognise immediately based on what is observed in the moment (e.g. Frank 1997; Pountain and Robbins 2000) and thus, coolness is not something we need to think about a lot. This applies to both people and objects and it is based on an immediate judgement of both their inner and outer coolness. The interesting

Table 1. Identified characteristics of inner coolness.

Inner cool characteristic	Sources
Being rebellious	Pountain and Robbins (2000)
Being antisocial	Pountain and Robbins (2000) and Nancarrow, Nancarrow, and Page (2001)
Embracing authenticity	Southgate (2003)
Embracing innovation	Pountain and Robbins (2000)
Seeking exclusivity	Charles (2002) and Nancarrow, Nancarrow, and Page (2001)
Seeking pleasure	Pountain and Robbins (2000) and MacAdams (2001)
Seeking personal development	Pountain and Robbins (2000) and MacAdams (2001)
Being/appearing in control	Pountain and Robbins (2000) and MacAdams (2001)
Making hard thing appear easy	Pountain and Robbins (2000) and Majors (2001)
Being detached or emotionally neutral	Majors and Billson (1993), Stearns (1994), Pountain and Robbins (2000) and MacAdams (2001)
Being strongly tied to a group	O'Donnell and Wardlow (2000), MacAdams (2001) and Gioia (2009)

issue here is that people do not identify and rate all the qualities of a person or a product in order to make this judgment, but they perceive and/or *infer* instead only the qualities they consider as *relevant* to the specific context (Kruglanski and Gigerenzer 2011; Van Schaik, Hassenzahl, and Ling 2012).

2.3. Coolness is grounded in people's communities

While one's perception of coolness in people and objects is subjective, and therefore often different from person to person, it is not simply a matter of individual preference. Instead the 'rules' about what is cool and what is not cool are deeply grounded in the communities we belong to, for example, cultural, social, political and sexual groups (e.g. Horton et al. 2012; MacAdams 2001; O'Donnell and Wardlow 2000; Thompson 1973). It is through participation in these different communities that our individual perception of coolness is shaped. Young adults who like to skateboard belong to a very different community in society than middle-aged bank employees. Consequently they have very different ideas about what is cool. One group might like to listen to Eminem, while the other might like Johnny Cash. But these communities have not formed because of shared taste in music. The taste in music is formed through participation in the community, as meaning is shared (Gioia 2009; O'Donnell and Wardlow 2000) and is constantly negotiated (Rodkin et al. 2006). This common perception about cool not only ties a group together, but also separates a group from other groups as well as from mainstream society (Horton et al. 2012; Saxton 2005). As a consequence, we cannot universally conclude about the coolness of persons or products, without also contemplating the communities they are parts of and are considered in (Gerber and Geiman 2012).

3. Measuring cool

The primary contribution of this paper is the COOL questionnaire, shown in Figure 2. In order to produce the listed questionnaire items, we went through a process of careful application of statistical methods (described in detail in Sections 4 and 5). This process took us from the 11 characteristics identified in the literature (Table 1) to three factors suitable for measuring the perceived inner coolness of an interactive product. The COOL questionnaire represents a validated proposal for measuring perceived inner coolness through 17 items on a 7-point Likert-like scale. It can specifically be used to measure an interactive product's perceived inner coolness and can be directly applied to an empirical study of coolness in the form presented in Figure 2.

The COOL questionnaire measures desirability through six items related to how a specific interactive product appeals to personal desire, for example, 'This device can make me happy' or 'This device can make me look good'. Rebelliousness is measured through five statements regarding the conventionality of the device, for example, 'This device is outside the ordinary' or 'This device is unconventional'. Usability is measured through five items related to ease of use, for example, 'This device is easy to operate' or 'This device is effortless to use'.

Furthermore, the COOL questionnaire contains the control question 'This device is cool'. This question was included in the questionnaire as a direct measure of coolness in a similar matter that the ugly/beautiful, bad/good items are included to Attrakdiff2 questionnaire (Van Schaik, Hassenzahl, and Ling 2012). The control question will allow practitioners to understand which of the three inner cool factors mostly affect the cool perception in specific cases. For example, usability might be considered as the dominant factor for shaping perceived inner coolness when a bank employee interacts with a new IT system that makes his work easier, or rebelliousness when we are buying a new watch. Finally, the questionnaire was purposefully designed in order to be easily used for a variety of products by replacing the word 'device' with, for example, 'website'.

4. Developing the COOL questionnaire

Creating questionnaires require meticulous statistical exploration and confirmation in order to ensure

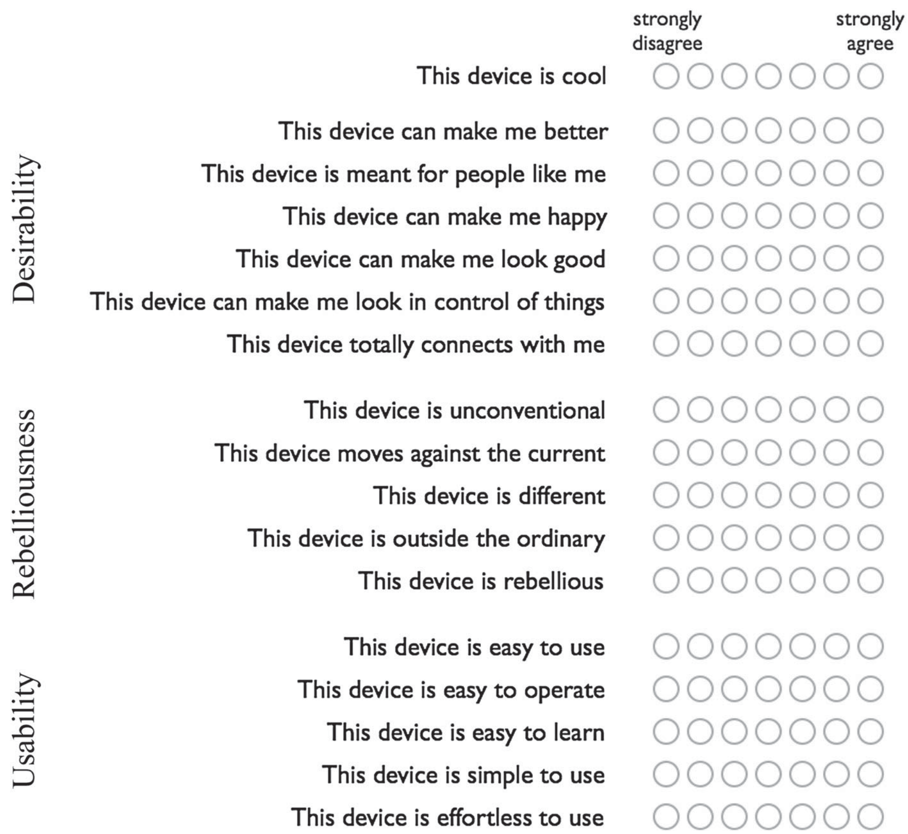

Figure 2. The COOL questionnaire.

robustness, for example, to be certain that the proposed question items measure particular factors reliably. In order to create the COOL questionnaire, we applied EFA and CFA by following the example of Lavie and Tractinsky (2004) when they created their classic/expressive aesthetics questionnaire. Similar process was also followed in the creation of other UX questionnaires, such as Attrakdiff (Hassenzahl, Burmester, and Koller 2003).

EFA is based on an iterative process where questionnaire items (individual questions) are removed from an initial pool of items, based on the degree by which they contribute to measuring a particular factor. The extent of item contribution is measured through factor loadings, where a high loading essentially means that the item correlates highly with the given factor. If an item has a low loading (in our case below a cut-off level of 0.6) on all factors, then that item should be removed. This is an iterative process where according to Lavie and Tractinsky (2004), after removing a number of items, another EFA should be conducted on a new dataset until each of the remaining items has a sufficiently high level of correlation with a single factor. In relation to this, it is important to note that reliability is also about having items that correlate highly with *only one* factor and not with multiple, as one particular item should not measure multiple constructs. This is partly what lies behind the concept of 'multicollinearity' which should be low (for details, Lavie and Tractinsky 2004).

CFA is, as the name implies, of confirmatory nature. Where the EFA is about exploring (removing and sometimes adding items depending on the emerging factors), the CFA is applied to make a final validation of these factors and the items that measure them (Schreiber et al. 2006). In CFA the notion of loadings is used in the same way as in EFA described above. In CFA there are not only item loadings on factors, but also covariance 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. More specifically, we can say that when getting acceptable levels on the fit indices, we can expect to observe similar patterns of loadings and covariance on a different data sample (assuming this has the same demographic characteristics and based on evaluation of similar products).

For validity it is crucial that the CFA is not conducted on the same datasets as the EFA; otherwise, you end up confirming a factor model on the same data on which it was derived, and thereby run the risk of inflated model-fit measures. In CFA several measures are applied to determine model fit. For details on this, we refer to Schreiber et al. (2006).

4.1. Process

In order to measure perceived inner coolness of an interactive product, we produced the COOL questionnaire starting from the 11 inner cool characteristics derived from the literature (Raptis, Kjeldskov, and Skov 2013). The process of doing this consisted of three phases: (1) creating candidate questions, (2) EFA studies and (3) CFA studies. Figure 3 outlines this process along with the number of question items removed after each study in the three phases. Specific question items removed after each study can be seen in the Appendix.

In the first phase of the process, we created candidate items for each of the 11 inner cool characteristics. Two of the authors created 15 items per characteristic (165 items in total). This was done by taking inspiration from other questionnaires in the literature and from our own understanding of each characteristic. For example, for *authenticity*, some of the produced candidate items were as follows: 'This device is one of a kind', 'This device is original' and 'This device mimics other devices'. Afterwards, six senior researchers evaluated the appropriateness of the 165 initial candidate items. Finally, the items per characteristic were reduced from 15 to 13 (143 in total).

In the second phase, we conducted iteratively four EFA studies on different datasets in order to discover the underlying factor structure of inner coolness as well as the items that contributed most to each factor. In the fourth EFA study, we also included external questionnaires in order to test for convergent validity, that is, to assess whether or not the COOL questionnaire measured the same constructs proposed by other UX questionnaires. This allowed us to see how our identified factors relate to established UX qualities. For this reason, we included the Attrakdiff2 questionnaire (Van Schaik, Hassenzahl, and Ling 2012), an aesthetics questionnaire (Lavie and Tractinsky 2004) and an attractiveness questionnaire (Quinn and Tran 2010).

Finally, in the third phase, we conducted three CFA studies to confirm the factor structure suggested in phase two.

4.2. Products

The products of coolness assessment for the development of our questionnaire were a number of mobile devices. In all EFA and CFA, we asked participants to assess a mobile device by answering a number of items through a web page, which on the left side showed the device (Figure 4) and on the right side listed the items. We decided to present the mobile devices through images based on the facts that (a) coolness is perceived immediately, (b) making a judgement about a product that it is boxed and we cannot not interact with is something that happens every day in product choice situations and (c) the same approach is used to assess other established perceived qualities, such as visual appeal (Lindgaard et al. 2006) or perceived usability (Flavián, Guinalíu, and Gurrea 2006; Sauro 2010). All items were presented randomly in order not to feel repetitive to the participants (we did the same when we included the external questionnaires) and were all rated on a 7-point Likert scale from 'strongly disagree' to 'strongly agree'.

In total, our participants made assessments of 16 different mobile devices, with an average of 98.6 participants per device in the EFA studies, and 152.3 participants per device in CFA studies. Since we wanted to measure the perceived inner coolness of the devices and we did not want our participants to be influenced by external parameters, we had to experimentally control them. To achieve this, we presented the participants with mobile devices of the same colour, without any visible indication of brand, and with their screens turned off in order to remove possible effects of the interface (Figure 4).

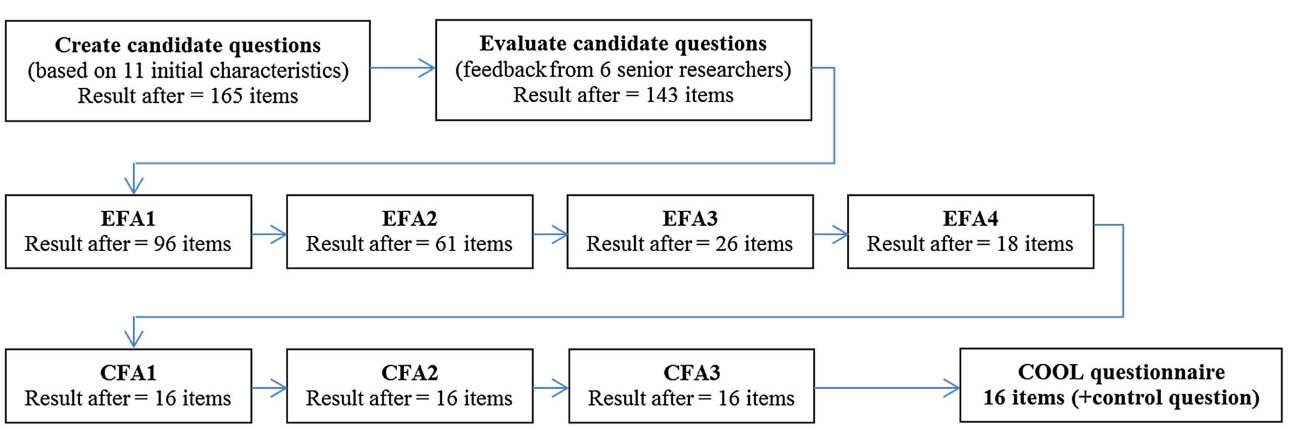

Figure 3. Process of creating the COOL questionnaire.

Figure 4. Experimental set-up.

4.3. Participants

In order to perform the necessary statistical analyses that would lead to a questionnaire, we had to include a large number of participants who belonged to a variety of communities. We opted for Amazon Mechanical Turk (MTurk) in order to achieve a large sample size. Several studies have shown MTurk participants to be reliable for studies similar to ours, including studies within HCI (Boujarwah, Abowd, and Arriaga 2012; Gottlieb et al. 2012; Gupta et al. 2012; Heer and Bostock 2010; Heimerl et al. 2012) as well as in other fields, such as political science (Berinsky, Huber, and Lenz 2012). Thus using MTurk allowed us to reach the large sample sizes needed for our statistical analyses without compromising the quality of the data. To reduce language barriers, we wanted to reach native English-speaking participants and recruited participants only from the USA. Although this is a limited geographical area, there was a large variety of communities represented in our dataset. We had participants with varying ages (18–67 years), ethnicities (self-identified as Caucasian, African-American, Hispanic, Asian, Arab, etc.) and a variety of geographical locations within the USA. By including a large number of participants from various communities, we made sure that our sample was heterogeneous, as suggested by Kline (1993). Thus our produced questionnaire is not community specific; that is, it reliably measures how different people perceive coolness. The same approach was used in the development of other UX questionnaires in the past (see e.g. Lavie and Tractinsky 2004).

The participants assessed only one mobile device each and participated only in one study. All participants who reported prior experience with the specific mobile device were removed from the datasets. The reason for removing those participants is that research has shown that prior experience with an IT product affects evaluation results (Langdon, Lewis, and Clarkson 2007; Sauro 2011). Additionally, we wanted to measure perceived inner coolness at first sight, not after actual usage. MTurk also provided us with the participants' average completion times and participants with considerable lower completion times than the average were also removed, as this was an indication of not filling in the questionnaire properly. Depending on the number of questions they had to answer, participants were paid an incentive ranging between $0.30 and $1.10. These incentives were provided based on MTurk's guidelines on how to fairly pay participants.

In the end, after the ones we had to remove, we had a total of 2236 participants (916 females) across the EFA and CFA studies, with a mean age of 28.36 (SD = 7.58) years. A total of 1322 of them participated in the EFA studies and 914 in the CFA studies. Details about the number of participants per study can be found in Tables 1 and 4.

4.4. Key EFA and CFA decisions

A number of decisions were made in order to conduct EFAs and CFAs. For EFAs the most important decisions were in relation to the adequacy of our sample, the method and the criteria for factor extraction, and the factor rotation method. Sampling adequacy was tested through the Kaiser–Meyer–Olkin (KMO) measure. By following recommendations from the literature (e.g. Bulmer 1979; Field 2009), we have used Principal Axes Factoring as the extraction method, and since we did not know beforehand if the extracted factors would be orthogonal or not, we opted for oblique rotation method. The number of extracted factors was determined through a Scree test and through parallel analysis (using Monte Carlo Principal Component Analysis [PCA], Watkins). In all four EFA studies, we removed items by applying two criteria: low communalities (<.5) and low factor loadings (<.6). Finally, in order to conduct EFA, homogeneity of variances should be secured. In order to test this assumption, we used Bartlett's Test of Sphericity, which had to be significant.

The CFA was 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 (Schreiber et al. 2006). This is then followed by a post-analysis in which model verification is conducted. In the pre-analysis, we examined our CFA samples in terms of missing data, normality, linearity and multicollinearity as suggested by Schreiber et al. (2006). In this regard we had no missing data. In terms of normality, SEM leans on the assumption of normally distributed data. In our case, all CFA datasets had univariate normality. Normality is assessed by measures of Skewness and Kurtosis. Skewness is an expression of the extent to which the distribution leans towards left or right relative to the normal. Kurtosis determines the level of peaknedness of the distribution. A perfect normal distribution has Skewness and Kurtosis values of 0. In our SEM datasets, Skewness values were between −.6 and .6 and Kurtosis values between −1.3 and 1.3. According to Bulmer (1979), these are within acceptable boundaries, that is, individual variables are assumed to be normally distributed.

Linearity between latent and manifest variables is also an assumption made in SEM analysis. Based on the strong factor loadings (>.67) identified during all of our EFA studies, we assume linearity in our CFA datasets. Furthermore, SEM is based on the assumption of absence of multicollinearity; that is, latent factors do not correlate to an extent to which they can be said to measure the same construct. Based on the relatively low correlations between latent variables (<.5), we assume absence of multicollinearity in our CFA datasets.

Finally, we used SPSS as the statistical software for EFAs and the AMOS package v. 22 for CFAs. In particular for CFAs, these tools provide a broad range of indices, which we applied to determine the degree of model fit, for example, the ratio of χ^2 to df, Normed Fit Index (NFI), Comparative Fit Index (CFI), Goodness-of-Fit Index (GFI) and several others. Table 6 provides an overview of all applied fit indices and we refer to Schreiber et al. (2006) for more details on these.

5. Results

In this section, we present findings from our EFAs and CFAs.

5.1. EFA results

Thousand three hundred and twenty-two participants were used in all four EFA studies and they evaluated a total of 13 mobile devices through the experimental set-up we have described before. In all EFA studies, Bartlett's Test of Sphericity was significant (<.001) for all the models we have produced. Table 2 presents an overview of the four EFA studies.

A total of 310 people participated in EFA1 and rated 143 items. The purpose for conducting EFA1 was twofold: (a) we wanted to reduce the total number of items and (b) we wanted to check the appropriateness of each item. A total of seven models were produced in EFA1 with

Table 2. The four EFA studies.

Study	*n*	Devices	*i*
EFA1	310	LG Optimus LS II and a concept Blackberry	143
EFA2	195	Caterpillar B15, Apple iPhone 5 and Huawei Ascend P6	96
EFA3	537	HTC One Mini, Sony Experia Z Ultra, Nokia 8210, Motorola RAZR and a concept phone	61
EFA4	280	Apple iPhone 5c, Nokia Lumia 625 and Huawei U8300	26 + 23[a]
Total	1.322	A total of 13 mobile devices, mean = 98.6 participants per device, SD = 33.5	

Note: *n* = number of participants, *i* = number of items used as input.
[a]Items from external questionnaires.

three-, four- and five-factor structures. Items were removed only if they did not belong to any of the factors in any of the seven models. In all seven models, the KMO was >.940, fulfilling the criteria for sampling adequacy. Another interesting observation of EFA1 was that all the negative items were clustered together in one factor. Since there are studies that pinpoint the issues of including negative phrased questions in questionnaires (e.g. Sauro and Lewis 2011), we decided to change them to positive. Additionally, a number of items were rephrased, as they were characterised as confusing by the participants. After EFA1 the total number of items was reduced from 143 to 96 and were used as input for EFA2.

In all, 195 participants rated 96 items in EFA2. In EFA2, our data suggested a three- or possibly four-factor structure. A total of five models were produced and the lowest KMO was .957. The process of removing items was the same as in EFA1 and we kept all the items that belonged to at least one of the factors in at least one of the models. After EFA2 the number of items was reduced from 96 to 61. In EFA3, 537 participants rated 61 items and finally a 3-factor structure for perceived inner coolness was confirmed through Scree tests and parallel analysis (Monte Carlo PCA, Watkins). Our best model had a KMO of .954 and the number of items was reduced from 61 to 26.

In EFA4, we included three established UX questionnaires in order to check the convergent validity of the COOL questionnaire: Attrakdiff2 (Van Schaik, Hassenzahl, and Ling 2012) that measures *hedonic* (four items) and *pragmatic quality* (four items), an aesthetics questionnaire (Lavie and Tractinsky 2004) that measures *classic* (five items) and *expressive aesthetics* (five items) and an attractiveness questionnaire (Quinn and Tran 2010) that measures the *attractiveness* (five items) of mobile devices. In total, 280 people rated 26 + 23 items. By applying the cut-off criteria, we ended up with five factors: three from our own dataset and two from the external questionnaires. Based on the items going into each of our three factors (Table 3), we chose to name them *desirability, rebelliousness* and *usability*. Items from the external questionnaires either clustered around two new factors or were discarded due to loadings being below the cut-off level of 0.6. In total, 18 question items from the established questionnaires were removed. The remaining items merged around two factors, which we named as *new attractiveness* and *classic aesthetics*. One of these factors was named *new attractiveness* as it was formed on the basis of items from two different questionnaires (Table 3).

Cumulatively this five-factor model explained 77.91% of the variance and Bartlett's Test of Sphericity was significant (<.001), while KMO was .921. Thus, we found homogeneity in variances and the sample size is deemed adequate. We ended up having 18 items for perceived inner coolness. Thus, in EFA4 we removed a further 8 out of 26 items. Results regarding EFA4 are presented in Table 3.

Table 3. Item loadings per factor in EFA4.

	Factor:	A	B	C	D	E
	Eigenvalue:	*9.112*	*4.476*	*2.458*	*1.033*	*0.841*
	Cronbach's α:	*.952*	*.928*	*.914*	*.864*	*.737*
Desirability	This device can make me better	0.861	0.160	0.012	−0.136	0.016
	This device is meant for people like me	0.835	−0.091	−0.003	0.040	0.051
	This device can make me happy	0.831	−0.036	0.020	0.080	−0.008
	This device can make me look good	0.787	0.019	−0.001	0.093	0.016
	This device can make me look in control of things	0.771	0.072	0.065	0.029	0.062
	This device totally connects with me	0.770	−0.060	0.007	0.187	0.004
	This device can please me[a]	0.731	−0.026	0.048	0.076	0.109
Rebelliousness	This device is unconventional	−0.040	0.874	−0.016	0.002	−0.024
	This device moves against the current	−0.042	0.836	0.007	0.035	−0.002
	This device is different	−0.063	0.825	0.018	0.055	−0.022
	This device is outside the ordinary	−0.056	0.806	−0.056	0.049	0.172
	If it was a person, this device would be a rebel[a]	0.202	0.789	−0.057	−0.049	−0.042
	This device is rebellious	0.121	0.736	0.145	0.052	−0.160
Usability	This device is easy to use	−0.05	0.003	0.928	0.053	−0.002
	This device is easy to operate	0.025	0.052	0.850	−0.033	−0.009
	This device is easy to learn	−0.109	0.049	0.820	0.068	0.051
	This device is simple to use	0.102	−0.095	0.770	−0.067	−0.020
	This device is effortless to use	0.103	−0.019	0.675	−0.016	0.110
New attractiveness	I find this device: boring/interesting[b]	0.138	−0.025	0.043	0.898	−0.081
	I find this device: plain/eye-catching[b]	−0.036	0.124	0.011	0.825	0.030
	I judge the device to be: dull/captivating[c]	0.168	0.059	−0.042	0.604	0.188
Classic aesthetics	This device has clean design[d]	0.048	−0.017	0.068	−0.025	0.747
	This device has clear design[d]	0.092	−0.014	0.077	0.071	0.617
	Sum of squared loadings (total variance explained): 77.91%					

Note: A = desirability; B = rebelliousness; C = usability; D = new attractiveness; E = classic aesthetics.
[a]Removed in CFA1.
[b]Originates from attractiveness questionnaire.
[c]Originates from Attrakdiff2.
[d]Originates from aesthetics questionnaire.

Table 4. Factor correlation matrix for EFA4.

	Desirability	Rebelliousness	Usability	New attractiveness
Desirability	1.000			
Rebelliousness	0.223	1.000		
Usability	0.412	0.033	1.000	
New attractiveness	0.605	0.422	0.195	1.000
Classic aesthetics	0.511	−0.077	0.487	0.345

Furthermore, Table 4 presents the correlations between the five factors. The factors of new attractiveness and classic aesthetics from external questionnaires correlate with the three perceived inner coolness factors. However, the findings outlined in Table 3 also show that new attractiveness and classic aesthetics have relatively low factor loadings on the three inner coolness factors, while most items from established UX questionnaires were below the cut-off level. This indicates that established UX questionnaires measure a different construct from inner coolness. That said, the correlations in Table 4 indicate a relation between the three inner cool factors and the two factors of new attractiveness and classic aesthetics. We discuss this relation in Section 6.1.

In terms of the three emerged factors of perceived inner coolness, desirability accounted for 39.62% of the variance and consisted of seven question items. We found this factor to be influenced by items initially proposed to belong to the inner cool characteristics (Table 1) of *seeking personal development*, *strongly tied to a group*, *seeking pleasure* and *being/appearing in control*. Rebelliousness was made up of six items and accounted for 19.46% of the variance. All six came from those initially proposed to belong to the inner characteristic of *being rebellious*. Usability accounted for 10.646% of the variance and consisted of five items from the characteristics of *being/appearing in control* and *making hard things appear easy*.

Note here that we went from the initial 11 inner cool characteristics to the 3 factors mentioned above. Thus, through the EFA studies we discarded numerous (125) of the original items, and even entire characteristics such as *seeking exclusivity, embracing innovation, embracing authenticity, being antisocial* and *being detached* as their related items did not provide sufficient contributions to any of the emerging factors, for the case of mobile devices. A detailed list of the items that were removed on each EFA study can be found in the appendix.

5.2. CFA results

In order to prove stability for a single sample dataset, Schreiber et al. (2006) argues that there is a general consensus on applying sample sizes of at least 10 participants per estimated parameter. Another approach is to prove stability by conducting multiple CFAs on different datasets. In our study we did both as detailed below.

A total of 914 participants were used over three CFA studies and they evaluated a total of 6 mobile devices (Table 5). None of the participants had taken part in the EFA studies.

The first CFA study was based on the three-factor model of perceived inner cool as emerged from EFA4 and we assessed the model fit of the factors desirability, rebelliousness and usability. This model consisted of 35 distinct parameters to be estimated; hence we needed at least 350 participants to ensure stability. We based the dataset on 355 participants, where each evaluated a single device. In the first run, all model fit indices suggested a close fit with the exception of PCLOSE. This was below the 0.05 threshold. PLCOSE represents a *p*-value indicating whether the null hypothesis of root-mean-square error of approximation (RMSEA) = 0.05 (indicating a close fitting model) is significantly different from the observed value. Thus, a PLCOSE value larger than 0.05 indicates no significant difference from this null hypothesis; that is, in that case the model fit would be close. To increase model fit, we went through two iterations of reviewing modification indices in order to determine which items to remove (i.e. removing items causing high modification indices reduces the noise on model fit). In each iteration, we removed one item and recalculated model fit indices. In the end, we removed the following two items: 'This device can please

Table 5. The three CFA studies.

Study	*n*	Factor model	Devices	*i*
CFA1	355	Three factors: Desirability, Rebelliousness, Usability	Alcatel One Touch 282, ASUS Padfone 2 and Samsung Galaxy S4 Mini	18
CFA2	559	Five factors: desirability, rebelliousness, usability, new attractiveness, classic aesthetics	Apple iPhone 5c, Nokia Lumia 625 and Huawei U8300	16 + 5[a]
CFA3	914	Five factors (merged datasets): desirability, rebelliousness, usability, new attractiveness, classic aesthetics	All of the above	16 + 5[a]
A total of 6 devices, mean = 152.3 participants per device, SD = 37.2				

Note: *n* = number of participants, *i* = number of question items used as input.
[a]Items from external questionnaires.

me' and 'If it was a person this device would be a rebel'. Thus, after 2 iterations and removing 2 items, we reached the final model consisting of the 3 factors of desirability, rebelliousness and usability as well as 16 underlying items. Table 6 shows the respective item loadings and model-fit indices obtained in the first CFA where indices are within acceptable thresholds.

In CFA2, we evaluated a 5-factor model including the 3 factors related to perceived inner cool (with the 16 underlying items) and the 2 factors of new attractiveness (3 items) and classic aesthetics (2 items). This model consisted of 52 distinct parameters to be estimated, which indicates that at least 520 participants were needed. CFA2 was based on a different dataset from CFA1 and we had 559 participants. In total they evaluated three devices. In the first run, we found all model-fit indices to be within acceptable ranges (Table 6).

In the third and final CFA study, we merged the datasets from CFA1 and CFA2 and re-evaluated the five-factor model confirmed in CFA2. Thus, CFA3 is based on 914 participants evaluating a total of 6 different devices. Model-fit indices were all within acceptable ranges after the first run (Table 6). All item loadings on factors are significant. We further validated the five-factor model by examining the matrix of standardised residuals. Residuals denote the deviations between observed values and the estimated parameters; that is, residuals express the level of observed differences between all pairs of variables (items and factors) and their estimated loadings. A model with a good fit will have the residuals centred around zero. We examined the matrix of standardised residuals provided in AMOS for excessively high values and found none larger than ±2, which indicates a good model fit (Schreiber et al. 2006).

Based on the model-fit indices, sample sizes and CFA replications, we found a three-factor model of the perception of perceived inner cool and a five-factor model also including new attractiveness and classic aesthetics to be supported by our data. Thus, our model can be considered reliable in the prediction of data.

Table 7 presents the correlation matrix between the five factors obtained in CFA3. These correlations are comparable to those found in EFA4 (see Table 4). It again shows the factors of new attractiveness and classic aesthetics to be correlating with the three factors of perceived inner coolness. None of our inner coolness factors (desirability, usability and rebelliousness) have a 1–1 correlation to measures of new attractiveness or classic aesthetics. However, we found the correlation between desirability and new attractiveness to be 0.794. This indicates that those concepts are closely related, yet measuring separate dimensions.

Table 6. Item loadings per factor and model-fit indices for the three CFA studies.

		CFA1	CFA2	CFA3
Desirability	This device can make me better	0.79	0.79	0.85
	This device is meant for people like me	0.86	0.84	0.87
	This device can make me happy	0.87	0.86	0.86
	This device can make me look good	0.86	0.86	0.86
	This device can make me look in control of things	0.77	0.80	0.82
	This device totally connects with me	0.87	0.87	0.90
Rebelliousness	This device is unconventional	0.77	0.72	0.74
	This device moves against the current	0.62	0.77	0.84
	This device is different	0.79	0.85	0.76
	This device is outside the ordinary	0.74	0.84	0.79
	This device is rebellious	0.73	0.73	0.86
Usability	This device is easy to use	0.90	0.90	0.90
	This device is easy to operate	0.89	0.88	0.88
	This device is easy to learn	0.88	0.85	0.84
	This device is simple to use	0.90	0.88	0.85
	This device is effortless to use	0.71	0.75	0.75
New attractiveness	I find this device: boring/interesting	–	0.94	0.88
	I find this device: plain/eye-catching	–	0.86	0.94
	I judge the device to be: dull/captivating	–	0.90	0.86
Classic aesthetics	This device has clean design	–	0.84	0.82
	This device has clear design	–	0.80	0.79
Model-fit indices				
Ratio of χ^2 to df (CMIN/df, acceptance threshold ≤3)		2.2	2.4	2.8
NFI (acceptance threshold ≥.95)		.94	.95	.97
Incremental Fit Index (acceptance threshold ≥.95)		.97	.97	.98
Tucker-Lewis Index (acceptance threshold ≥.95)		.96	.97	.97
CFI (acceptance threshold ≥.95)		.97	.97	.98
GFI (acceptance threshold ≤.95)		.93	.93	.95
Adjusted Goodness-of-Fit Index (acceptance threshold ≤.95)		.90	.91	.94
RMSEA (acceptance threshold ≤.06)		.06	.05	.04
p of close fit (PCLOSE, acceptance threshold >.05)		.06	.40	.97

Note: All are within acceptable thresholds, indicating good model fit.

Table 7. Factor correlation matrix for CFA3.

	Desirability	Rebelliousness	Usability	New attractiveness
Desirability	1.000			
Rebelliousness	0.199	1.000		
Usability	0.341	−0.051	1.000	
New attractiveness	0.794	0.417	0.144	1.000
Classic aesthetics	0.579	−0.131	0.534	0.438

Table 8. Results for factors' CR, Cronbach's α and AVE for the three CFA studies.

	CFA1 ($n = 355$)			CFA2 ($n = 559$)			CFA3 ($n = 914$)		
	CR	Cronbach's α	AVE	CR	Cronbach's α	AVE	CR	Cronbach's α	AVE
Desirability	.934	.933	.702	.934	.934	.704	.937	.934	.711
Rebelliousness	.790	.780	.443	.912	.912	.677	.897	.889	.636
Usability	.933	.929	.738	.926	.924	.714	.927	.929	.719
New attractiveness	–	–	–	.910	.940	.772	.922	.924	.799
Classic aesthetics	–	–	–	.823	.820	.699	.789	.801	.651

To further test for consistency and validity of our factors, we measured composite reliability (CR), Cronbach's α and average variance extracted (AVE). For all five factors (usability, desirability, rebelliousness, new attractiveness and classic aesthetics), all values, except for rebelliousness in CFA1, were within acceptable ranges: AVE>.5 (suggested by Fornell and Larcker 1981), CR>.7 (suggested by Hair et al. 1998) and Cronbach's $\alpha > .8$ (suggested by Furr and Bacharach 2013). We then assessed discriminant validity using the Fornell and Larcker (1981) technique. For all pairs of factors, both AVEs were larger than the shared variance (square of their correlation). More details can be found in Table 8.

Based on the above, we can conclude that the three-factor model for perceived inner coolness is stable and the proposed factor differs from established perceived qualities for the case of mobile devices.

All the 3 factors contribute to the perception of inner coolness of a product and they have emerged through a meticulous process that started with 143 items and ended up with 16. According to our findings, all the three factors contribute positively to the perception of inner coolness, but as we know from the literature, some factors might be considered more relevant than others depending on the context (Kruglanski and Gigerenzer 2011; Van Schaik, Hassenzahl, and Ling 2012). For example, when somebody buys new digital calipers (Kelly 2013, 57), usability might play a crucial role on perceiving the tool as cool, because it allows the owner to complete tasks easily and with more control. At the same time, somebody else might intentionally buy the OnePlus 2 phone and not another more common Android phone as he or she might believe it is rebellious as it is very difficult to acquire (it is sold only online and through a waiting list).

6. Discussion

Our aim in this paper was to identify the factors that contribute to perceived inner coolness and produce a tool that can reliably measure them. Based on this, we designed and developed the COOL questionnaire using the statistical techniques of EFA and CFA. Our data suggest that the perception of inner coolness in an interactive product is determined by the following:

Desirability. The degree to which individuals believe that a product supports personal development and pleasure and the degree the individuals relate to it.

Rebelliousness. The degree to which individuals believe that a product stands out from the rest.

Usability. The degree to which individuals believe that a product is usable and makes a task more easy or appear more easy.

6.1. Coolness and UX research

This paper contributes to UX research in two ways. Firstly, we showed that coolness can be considered as a new dimension of UX, at least in the case of mobile devices. According to our findings, perceived inner coolness acts as an umbrella construct, which can be decomposed to desirability, usability and rebelliousness. In our case of assessing mobile devices, we found that our usability factor performed better than the factor of pragmatic quality suggested in existing questionnaires. Items measuring pragmatic quality contributed to low factor loadings on the usability factor. At the same time, our data suggest that desirability and rebelliousness are different constructs from the existing ones, which is indicated by values of discriminant validity and internal consistency. These were within acceptable ranges.

Secondly, we contributed to the ongoing discussion regarding the relationships between established UX qualities in different contexts of use (Bargas-Avila and Hornbæk 2011). Our data support the claim that many of the existing UX constructs overlap in specific contexts of use. For the case of mobile devices, items from the constructs of hedonic quality and expressive aesthetics were clustered around attractiveness and classic aesthetics, although with relatively low factor loadings. Thus, our data suggest that, for mobile devices, the combination of our three inner cool factors along with new attractiveness (three items) and classic aesthetics (two items) is more complete in measuring perceived UX of mobile devices. Of course, more studies are needed for different interactive products and while including more UX qualities in order to see if these findings are more general. We argue that such research direction is important as it could lead to a specific set of qualities and questionnaire items that practitioners could use in their evaluations.

Finally, our CFA studies showed that perceived inner coolness can be measured by desirability, usability and rebelliousness, qualities that are different from attractiveness and classic aesthetics. If we link back to our theoretical model, which proposes that coolness consists of inner and outer cool, the above may suggest that attractiveness and classic aesthetics are in fact measuring outer coolness. Since our experiment was not purposefully designed to research this issue, more studies are needed to have conclusive results. We consider this research direction as really important as it can lead to a model that will describe how our first impression with an interactive product (outer cool) is shaping our perception on its personality characteristics (inner cool) and how they both contribute to the overall impression regarding its coolness.

6.2. *The COOL questionnaire in comparison to other cool factors*

There are a few studies that have investigated how to systematically measure coolness of interactive products. Our study is in line with the initial results of Read, Horton, and Fitton (2012) and Horton et al. (2012). In those studies, it was found that coolness is related to constructs such as authenticity and innovativeness. This is very similar to the construct of rebelliousness that was included in our questionnaire. Our findings also overlap with the constructs identified by Culén and Gasparini (2012) who found that coolness is affected by mastery, usefulness and self-presentation. These constructs are closely related to our constructs of desirability and usability. Furthermore, our factors of desirability and usability overlap with constructs such as accomplishment, connection, identity and direct into action as identified in Farnsworth, Holtzblatt, Held, et al. (2014). Thus, there are considerable overlaps between constructs in our study and related work, yet there are other notions with which cool can be considered (see e.g. Read, Horton, and Fitton 2012). These relations among constructs need to be explored in future studies in order to have a better understanding on what coolness is.

To our knowledge, only Sundar, Tamul, and Wu (2014) published a questionnaire for measuring coolness. They decomposed this into the following three components: subculture, attractiveness and originality. In the following, we will compare our work to the Sundar, Tamul, and Wu (2014) study. Our findings show that inner coolness can be measured through usability, desirability and rebelliousness, while Sundar, Tamul, and Wu (2014) show that coolness can be measured through subculture, attractiveness and originality. Firstly, the underlying items of the originality factor in Sundar, Tamul, and Wu (2014) are similar to those included in our rebelliousness factor. Secondly, in relation to the subculture factor, we included most of these items in our initial pool of items, as can be seen in the appendix. However, our EFA studies indicated that these items did not load sufficiently high onto any factors, which is why we removed them. Thirdly, items that Sundar, Tamul, and Wu (2014) included in their attractiveness factor are comparable to several items that we either discarded during our EFA studies (i.e. 'This device is sexy', 'This device has style', 'This device is classy' and 'This device is innovative'), or belonged to the attractiveness factor of the Quinn and Tran questionnaire (2010). Finally, through our EFA studies we also found five items to be loading on the factor of usability, which was confirmed through the CFA studies. In comparison, Sundar, Tamul, and Wu (2014) initially had a factor denoted as utility, which bears close resemblance to the factor of usability in our study. However, Sundar, Tamul, and Wu (2014) eventually chose to remove utility.

So, why do we see these differences between studies? The first explanation is related to how the initial pool of items used in EFAs and CFAs were created. We created our initial pool of questions based on the 11 cool characteristics identified in a literature review (Raptis, Kjeldskov, and Skov 2013). Sundar, Tamul, and Wu (2014) applied an empirical approach and created their initial pool of items by asking participants to self-report their perceptions on 14 digital artefacts (ranging from websites to USB dongles and even World of Warcraft).

Second, our distinction between inner and outer cool explains why attractiveness is not part of our questionnaire. In this paper, we focused on inner cool and since attractiveness, according to our theoretical framework, belongs to outer cool, it was measured through other established questionnaires, for example, the aesthetics questionnaire by Lavie and Tractinsky (2004).

The third explanation for the differences between the studies is related to the experimental methods that were followed to create the questionnaires. Sundar, Tamul, and Wu (2014) assessed 18 different devices and software applications of various types. This is commendable in terms of testing the questionnaires' generalisability across products. In Sundar, Tamul, and Wu (2014), they had 16–36 participants assess each product. Lavie and Tractinsky (2004) used around 75 participants per website assessed to produce their questionnaire. In our case, we had an average of 125 participants per mobile device assessed. Looking across similar studies, there seems to be a trade-off in terms of the variety of products assessed and the number of participants assessing each. With enough participants, however, the high level of variability in products provides results that are more generalisable across product types. The high variability in products could also have led to the removal of the utility factor in Sundar, Tamul, and Wu (2014). The utility factor in Sundar, Tamul, and Wu (2014) is comparable to the usability factor derived in our study. Utility may not be as crucial in games, for example, World of Warcraft, as on websites or mobile phones.

In sum, there are several differences between our study and the one presented in Sundar, Tamul, and Wu (2014). Our findings show that inner coolness can be measured through usability, desirability and rebelliousness and suggest that outer coolness can be measured by new attractiveness and classic aesthetics. This distinction between inner cool and outer cool can also explain why some artefacts, even though they are perceived as ugly, are considered as cool, for example, the old VW Beetle (Imseng 2011). Given the differences between studies, it would be relevant to compare the two questionnaires in comparable experimental conditions in order to determine their relative strengths, weaknesses and similarities for assessing product coolness.

7. Conclusions

We have addressed the concept of cool in interaction design by researching literature on the concept both inside and outside the HCI field. Based on the insights gained from this, we have constructed a questionnaire for measuring the inner coolness of interactive products – we call it the COOL questionnaire and it is the key contribution of this paper. The questionnaire is shown in Figure 2. The questionnaire was developed from an offset in 11 inner cool characteristics derived from the literature, and was validated with an emphasis on mobile phones. Through a process of iteratively applying the statistical methods of EFA and CFA, we arrived at three factors that measure the perceived coolness of an interactive product: *desirability*, *rebelliousness* and *usability*. The questionnaire consists of 16 specific items that measure these 3 factors of perceived inner coolness, plus the control question 'this device is cool'.

Our questionnaire is meant as a practical tool that can be used to measure the perceived inner coolness of an interactive product during a design process. We believe that it can be a useful tool for interaction design practitioners when designing for cool in industrial projects. Furthermore, we think it will be also useful for interaction design researchers as a practical tool to understand the coolness dimension of UX.

8. Future work

The study of coolness in HCI and interaction design has just recently started and current studies seem to open a multitude of avenues for further research. This is also stressed by Holtzblatt, Rondeau, and Holtzblatt (2010) and Holtzblatt (2011) who list several challenges for conducting research on cool, for example, UX design process integration. Our work extends previous work on coolness within interaction design, for example, Fitton et al. "Climbing the Cool Wall" (2012), Fitton et al. "Constructing the Cool Wall" (2012), Read et al. (2011) and Read, Horton, and Fitton (2012), as we provide a validated questionnaire that measures perceived inner coolness of mobile devices. Holtzblatt (2011) raised the question of how designers can identify aspects of coolness for specific product categories and how this can be incorporated into the product requirements.

We believe that in order to use the COOL questionnaire in practice, there are at least five challenges that need to be addressed through further studies. Firstly, in both CFA studies we had an average of 152 participants from various communities assessing each mobile device. Therefore, we argue that our model is generalisable in terms of the perception of inner coolness of mobile devices. Yet, this points to a potential limitation of generalisability *beyond* mobile devices, which is relevant to discuss. We believe our questionnaire is suitable for other types of interactive products, as this was also the case with Lavie and Tractinsky's (2004) and the Attrakdiff2 (Hassenzahl, Burmester, and Koller 2003; Van Schaik, Hassenzahl, and Ling 2012) questionnaires. These were also created by focusing on one type of interactive product and then were successfully used to assess a variety of products. Of course, more studies are needed in order to have conclusive results.

Secondly, we believe that it would be important for practitioners to have more data on how our questionnaire performs. As an example of practical use, the iPhone 5c had the following mean scores: desirability = 3.8, rebelliousness = 2.9 and usability = 5.1. In comparison the Huawei U8300 scored lower in terms of desirability (2.9) and usability (4.7) but higher on rebelliousness (4.6). This difference may be explained by the fact that the Huawei U8300 moves against mainstream trends by reintroducing the tactile keyboard known from phones in the early 2000s; that is, it can be considered more rebellious than the iPhone 5c. At this moment we do not have enough data to understand what it means for a mobile device, for example, to score 5 on usability and 3.5 on rebelliousness in relation to inner coolness. Is 5 a good score? Is 3.5 a bad one? Additionally, what does it mean for a mobile device to be, for example, extremely rebellious? – a challenge that was pointed also by Horton et al. (2012). Do extremes have a negative impact on coolness? Such important issues were tackled in the past for other questionnaires too, where datasets from many studies were examined and analysed (e.g. the system usability scale, Bangor, Kortum, and Miller 2008). We argue that the same should be conducted for the COOL questionnaire too.

Thirdly, we also find it important to study the relation between coolness and actual use of interactive devices and systems in practice. This points towards the limitation of only assessing static images, which was the case in our study. Thus, our findings are limited to participants' initial perceptions of devices without any tactile sense of their shape and weight. However, as discussed earlier, coolness is something we recognise immediately from our first impression of objects and people, which makes the COOL questionnaire valid for initial assessments. However, as mentioned in Holtzblatt (2011) and Culén and Gasparini (2012), it is also relevant to study the extent to which perceptions of coolness change over time.

Fourthly, it is relevant to further study what characterises products that are considered as *universally cool* (Schiller 2012) or *classic* (Nancarrow, Nancarrow, and Page 2001). If perceived coolness does indeed change over time, or stays stable for some products, then interaction designers would obviously benefit from understanding this process better. Our questionnaire can play a role on that, as it will allow collecting different datasets that could be compared and analysed in order to understand how time influences perception of coolness.

Finally, our findings also suggest that the COOL questionnaire measures inner cool, while attractiveness and classic aesthetics contribute to outer cool. The contribution of attractiveness and classic aesthetics on outer cool is something that needs to be explored further and validated with more research data.

Acknowledgement

We would like to thank Dr Veronika Hinkle for her invaluable help.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The study was partly financed by the Danish Research Council [grant number 09-065143].

References

Bangor, A., P. Kortum, and J. Miller. 2008. "An Empirical Evaluation of the System Usability Scale." *International Journal of Human-Computer Interaction* 24 (6): 574–594.

Bargas-Avila, J. A., and K. Hornbæk. 2011. "Old Wine in New Bottles or Novel Challenges? A Critical Analysis of Empirical Studies of User Experience." Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Vancouver, May 7–12, 2689–2698.

Berinsky, A. J., G. A. Huber, and G. S. Lenz. 2012. "Evaluating Online Labor Markets for Experimental Research: Amazon.com's Mechanical Turk." *Political Analysis* 20: 351–368.

Boujarwah, F., G. Abowd, and R. Arriaga. 2012. "Socially Computed Scripts to Support Social Problem Solving Skills." Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Austin, May 7–12, 1987–1996.

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

Charles, G. 2002. "Cool for Cash." *Marketing Week* 25(33): 24.

Cowan, B. R., K. Avramides, and R. Beale. 2013. "Should Cool Be a Design Goal?" British Computer Society Human Computer Interaction Conference, London, September 9–13, 1–5.

Culén, A. L., and A. A. Gasparini. 2012. "Situated Techno-Cools: Factors that Contribute to Making Technology Cool in a Given Context of Use." *PsychNology* 10 (2): 117–139.

De Guzman, E. S. 2012. "'Cool' Card Sort: Adapting Desirability Studies to Assess 'Cool'-ness." ACM CHI Workshop on Cool aX Continents, Cultures and Communities 2012, Austin, May 7–12.

Farnsworth, C., K. Holtzblatt, T. Held, and S. Pai. 2014. "Cool in Business: Developing a Data-Based Instrument Measuring 'Cool'." Proceedings of HCII 2014, HCI International, Heraklion, June 22–24, 232–243.

Farnsworth, C., K. Holtzblatt, S. Pai, K. Wagg, T. Held, E. Wylen, S. L. Kennedy, and P. Kutty. 2014. "Measuring Products 'Coolness' – Developing a Measurement Instrument." Proceedings and Extended Abstracts CHI'14, Toronto, April 26–May 1, 893–896.

Field, A. 2009. *Discovering Statistics Using SPSS*. London: SAGE Publications.

Fitton, D., M. Horton, J. C. Read, L. Little, and N. Toth. 2012. "Climbing the Cool Wall: Exploring Teenage Preferences of Cool." Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Austin, Austin, May 7–12, 2093–2098.

Fitton, D., J. C. Read, M. Horton, L. Little, N. Toth, and Y. Guo. 2012. "Constructing the Cool Wall: A Tool to Explore Teen Meanings of Cool." *PsychNology* 10 (2): 141–162.

Flavián, C., M. Guinalíu, and R. Gurrea. 2006. "The Role Played by Perceived Usability, Satisfaction and Consumer Trust on Website Loyalty." *Information & Management* 43 (1): 1–14.

Fornell, C., and D. Larcker. 1981. "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error." *Journal of Marketing Research* 18: 39–50.

Frank, T. 1997. *The Conquest of Cool. Business Culture, Counterculture, and the Rise of Hip Consumerism*. Chicago: The University of Chicago Press.

Furr, R. M., and V. R. Bacharach. 2013. *Psychometrics: An Introduction*. London: SAGE Publications.

Gerber, J. P., and C. Geiman. 2012. "Measuring the Existence of Cool Using an Extended Social Relations Model." *PsychNology* 10 (2): 103–115.

Gioia, T. 2009. *The Birth and Death of the Cool*. Golden, CO: Speck Press.

Gottlieb, L., J. Choi, P. Kelm, T. Sikora, and G. Friedland. 2012. "Pushing the Limits of Mechanical Turk: Qualifying the Crowd for Video Geo-location." Proceedings of the ACM Multimedia 2012 Workshop on Crowdsourcing for Multimedia (CrowdMM), Nara, October 29–November 2, 23–28.

Gupta, A., W. Thies, E. Cutrell, and R. Balakrishnan. 2012. "mClerk: Enabling Mobile Crowdsourcing in Developing Regions." Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Austin, May 7–12, 1843–1852.

Hair, J. F., R. E. Anderson, R. L. Tatham, and W. C. Black. 1998. *Multivariate Data Analysis*. Upper Saddle River: Prentice Hall International.

Hassenzahl, M., M. Burmester, and F. Koller. 2003. "Attrakdiff: Ein Frageboten zur Messung wahrgenommener hedonischer und pragmatischer Qualität." In *Mensch & Computer 2003. Interaction in Bewegung*, edited by B. G. Teubner, 187–196. Stuttgart: Springer.

Heer, J., and M. Bostock. 2010. "Crowdsourcing Graphical Perception: Using Mechanical Turk to Assess Visualization Design." Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Atlanta, April 10–15, 203–212.

Heimerl, K., B. Gawalt, K. Chen, T. Parikh, and B. Hartmann. 2012. "Community Sourcing: Engaging Local Crowds to Perform Expert Work Via Physical Kiosks." Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Austin, May 7–12.

Holtzblatt, K. 2011. "What Makes Things Cool? Intentional Design for Innovation." *Interactions* 18 (6): 40–47.

Holtzblatt, K., D. B. Rondeau, and L. Holtzblatt. 2010. "Understanding 'Cool'." Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Atlanta, April 10–15, 3159–3162.

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

Imseng, D. 2011. *Think Small. The Story of the World's Greatest Ad*. Phoenix: Full Stop Press.

ISO 9241-11. 1998. *Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs) – Part 11: Guidance on Usability*. Geneva: International Standardization Organization (ISO).

Janlert, L. E., and E. Stolterman. 1997. "The Character of Things." *Design Studies* 18 (3): 297–314.

Jordan, P. W. 1997. "Products as Personalities." In *Contemporary Ergonomics*, edited by M. A. Hanson, 449–453. London: Taylor & Francis.

Kelly, K. 2013. *Cool Tools*. Published by KK*.

Kline, P. 1993. *An Easy Guide to Factor Analysis*. Oxford: Routledge.

Kruglanski, A. W., and G. Gigerenzer. 2011. "Intuitive and Deliberate Judgments Are Based on Common Principles." *Psychological Review* 118 (1): 97–109.

Langdon, P. M., T. Lewis, and P. J. Clarkson. 2007. "The Effects of Prior Experience on the Use of Consumer Products." *Universal Access in the Information Society* 6 (2): 179–191.

Lavie, T., and N. Tractinsky. 2004. "Assessing Dimensions of Perceived Visual Aesthetics of Web Sites." *International Journal of Human-Computer Studies* 60: 269–298.

Lindgaard, G., G. Dudek, G. Fernandes, and J. Brown. 2006. "Attention Web Designers: You Have 50 Milliseconds to Make a Good First Impression!" *Behaviour & Information Technology* 25 (2): 115–126.

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

Majors, R. 2001. "Cool Pose: Black Masculinity and Sports." In *The Masculinities Reader*, edited by S. M. Whitehead and F. Barrett, 209–217. Cambridge: Polity Press.

Majors, R., and J. M. Billson. 1993. *Cool Pose: The Dilemmas of Black Manhood in America*. New York: Touchstone.

McCrickard, D. S., J. Barksdale, and F. Doswell. 2012. "Understanding Cool: An Analytic Exploration of Contributing Factors for Teens." *PsychNology* 10 (2): 93–102.

Nancarrow, C., P. Nancarrow, and J. Page. 2001. "An Analysis of the Concept of Cool and Its Marketing Implications." *Journal of Consumer Behaviour* 1 (4): 311–322.

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

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

Quinn, J. M., and T. Q. Tran. 2010. "Attractive Phones Don't Have to Work Better: Independent Effects of Attractiveness, Effectiveness, and Efficiency on Perceived Usability." Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Atlanta, April 10–15, 353–362.

Raptis, D., J. Kjeldskov, and M. B. Skov. 2013. "Understanding 'Cool' in Human-Computer Interaction Research and Design." Australian Conference on Human-Computer Interaction, Adelaide, Australian Conference on Human-Computer Interaction, Adelaide, November 25–29, 53–62.

Read, J., D. Fitton, B.Cowan,, R. Beale, Y. Guo and M. Horton. 2011. "Understanding and Designing Cool Technologies for

Teenagers." CHI '11 Extended Abstracts on Human Factors in Computing Systems, Vancouver, May 7–12, 1567–1572.
Read, C. R., M. Horton, and D. Fitton. 2012. "Being Cool-getting Persona-l." CHI Workshop on Cool aX Continents, Cultures and Communities 2012, Austin, May 7–12.
Rodkin, P. C., T. W. Farmer, R. Pearl, and R. Van Acker. 2006. "They're Cool: Social Status and Peer Group Supports for Aggressive Boys and Girls." *Social Development* 15 (2): 175–204.
Sauro, J. 2010. "5 Second Usability Tests." http://www.measuringu.com/five-second-tests.php.
Sauro, J. 2011. "Does Prior Experience Affect Perceptions of Usability?" http://www.measuringu.com/blog/prior-exposure.php.
Sauro, J., and J. R. Lewis. 2011. "When Designing Usability Questionnaires, Does It Hurt to Be Positive?" Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Vancouver, May 7–12, 2215–2224.
Saxton, G. 2005. "Collections of Cool." *Young Consumers* 6 (2): 18–27.
Schiller, J. 2012. "Who Is the Coolest of Them All." CHI Workshop on Cool aX Continents, Cultures and Communities 2012, Austin, May 7–12.
Schreiber, J. B., A. Nora, F. K. Stage, E. A. Barlow, and J. King. 2006. "Reporting Structural Equation Modeling and Confirmatory Factor Analysis Results: A Review." *Journal of Educational Research* 99 (6): 323–338.
Southgate, N. 2003. "Coolhunting, Account Planning and the Ancient Cool of Aristotle." *Marketing Intelligence & Planning* 21 (7): 453–461.
Stearns, P. N. 1994. *American Cool: Constructing a Twentieth-Century Emotional Style*. New York: University Press.
Sundar, S. S., D. J. Tamul, and M. Wu. 2014. "Capturing 'Cool': Measures for Assessing Coolness of Technological Products." *International Journal of Human-Computer Studies* 72: 169–180.
Thompson, R. F. 1973. "An Aesthetic of Cool." *African Arts* 7 (1): 41–91.
Van Schaik, P., M. Hassenzahl, and J. Ling. 2012. "User-Experience from an Inference Perspective." *ACM Transactions on Computer-Human Interaction* 19(2), Article No. 11. doi:10.1145/2240156.2240159.
Warren, C., and M. C. Campbell. 2014. "What Makes Things Cool? How Autonomy Influences Perceived Coolness." *Journal of Consumer Research* 41 (2): 543–563.
Watkins, M. W. *Monte Carlo PCA for Parallel Analysis*. http://edpsychassociates.com/Watkins3.html.

Appendix

Items used in EFAs and CFAs without the ones that were removed in EFA1. The reason for not presenting the items from EFA1 is that after EFA1 many items were rephrased. On each item it is depicted the EFA or CFA study that was removed, or added. For example, the first item 'I would love to have this device' was removed in EFA4. The items in bold belong to the final version of the questionnaire.

I would love to have this device (-EFA4)	This device is normal (-EFA2)	**This device is outside the ordinary**
If I buy this device, my friends will follow (-EFA3)	This device is a step forward (-EFA2)	This device is pleasurable (-EFA3)
If it was a person this device would be a rebel (-CFA1)	This device is against the rules (-EFA3)	This device is popular (-EFA2)
It will be so nice to show this device to my friends(-EFA3)	This device brings new things (-EFA2)	This device is progressive (-EFA2)
My friends will love this device (-EFA2)	This device is all fun (-EFA2)	**This device is rebellious**
This device is ahead of the competition (-EFA3)	This device can empower me (-EFA3)	This device is relaxed (-EFA2)
This device can boost my performance (-EFA2)	This device is authentic (-EFA3)	This device is revolutionary (-EFA3)
This device can create new opportunities for me (-EFA2)	This device is calm (-EFA2)	This device is risky (-EFA3)
This device is at the center of attention (-EFA2)	This device is classy (-EFA3)	This device is satisfactory (-EFA2)
This device can give me authority (-EFA2)	This device is common (-EFA3)	This device is sexy (-EFA2)
This device can have a positive impact on my image (-EFA3)	This device is conservative (-EFA2)	**This device is simple to use**
This device can help me improve (-EFA3)	This device is desirable (-EFA4)	This device is special (-EFA2)
This device can help me master new skills (-EFA2)	This device is detached (-EFA2)	This device is state of the art (-EFA4)
This device can make difficult activities look easy (-EFA2)	**This device is different**	This device is stereotypical (-EFA3)
This device is suitable for people the age of my parents (-EFA2)	This device has quality (-EFA3)	**This device can make me better**
This device is straightforward (+EFA3 -EFA4)	**This device is easy to learn**	This device is trivial (-EFA2)
This device can make me look good	**This device is easy to use**	**This device is unconventional**
This device can make me look in control of things	This device is easy-going (-EFA2)	This device is underground (-EFA2)
This device can make me look more powerful (-EFA3)	This device is eccentric (-EFA2)	This device is unique (-EFA3)
This device can make me more productive (-EFA3)	This device is effective (-EFA3)	This device can please me (-CFA1)
This device can make me proud (-EFA3)	This device is efficient (-EFA3)	**This device moves against the current**
This device can make me unique (-EFA4)	This device is exciting (-EFA4)	This device resists authority (-EFA2)
This device is worth recommending (-EFA4)	This device is exclusive (-EFA3)	This device sets new standards (-EFA2)
This device can save me a lot of time (-EFA3)	This device is fresh (-EFA2)	This device starts a new era (-EFA3)
This device disobeys known norms (-EFA3)	This device is innovative (-EFA3)	This device stimulates me (-EFA3)
This device disrupts social order (-EFA2)	This device is magical (-EFA2)	**This device totally connects with me**
This device will influence future devices (-EFA2)	This device is mainstream (-EFA3)	This device does not care (-EFA2)
This device easily convinces me (-EFA2)	This device is meant for few (-EFA3)	This device has style (-EFA4)
This device fits well with my personal values (-EFA3)	This device has character (-EFA2)	**This device is easy to operate** (+EFA3)
This device is meant for people like me	This device is obedient (-EFA2)	**This device is effortless to use** (+EFA3)
This device has my name on it (-EFA3)	This device is obvious (-EFA2)	**This device can make me happy**
This device is easily recognizable (-EFA2)	This device is one of a kind (-EFA3)	This device is useful (+EFA2 -EFA3)
Very few individuals have this device (-EFA3)	This device is opposing society(-EFA2)	This device is normal (+EFA2 -EFA3)
This device imitates other devices (-EFA2)	This device is original (-EFA2)	