Let's Design a Product Coolness Scale! a Hands-on Teaching Activity to Understand Scale Development

Carine LALLEMAND [1, 2] and Regina BERNHAUPT [1]
[1] Eindhoven University of Technology, the Netherlands
[2] University of Luxembourg, Luxembourg

Standardized scales are popular evaluation instruments in Human-Computer Interaction (HCI) and design, both for research and practice. If scale development processes and psychometrics are well integrated into psychology or social sciences curriculums, students in design or HCI curriculums are less familiar with these notions. Knowing how these scales are built and understanding their underlying psychometric properties is essential for design researchers and practitioners alike. In this contribution, we present a 2-hours interactive teaching activity consisting of the design of a product coolness scale. Guided by step-by-step instructions, groups of students engage in several tasks to learn about scale development processes and requirements, and to better understand the psychometric properties of standardized scales. We share insights from our teaching practice.

CCS CONCEPTS • **Human-centered computing ~Human computer interaction (HCI) ~HCI design and evaluation methods**

Additional Keywords and Phrases: Standardized scale; Psychometrics properties; Teaching activity; Instructional technique; Human-Computer Interaction.

ACM Reference Format:
Carine Lallemand and Regina Bernhaupt. 2022. Let's Design a Product Coolness Scale! a Hands-on Teaching Activity to Understand Scale Development. In *4th Annual Symposium on HCI Education (EduCHI'22), April 30–May 1, 2022, New Orleans, LA, USA.* ACM, New York, NY, USA, 10 pages.

1 INTRODUCTION

Standardized scales are popular evaluation instruments in Human-Computer Interaction (HCI) and design, both for research and practice. While scale development processes and psychometrics are well integrated into psychology or social sciences teaching curriculums, students in design or HCI curriculums are less familiar with these notions. Yet design and HCI students will highly likely use standardized scales during their studies and later in their career. Knowing how these scales are built and understanding their underlying psychometric properties is essential for design researchers and practitioners alike. Developing a critical attitude towards these instruments, and being able to make informed choices about their use requires familiarity with scale development processes [11]. In the literature and research handbooks, scale development stages are rather consensual, following these main stages: (a) definition of the construct, (b) generation of a pool of items and choice of response format, (c) evaluation of the items by a panel of experts, (d) revision and selection of the items to retain, (e) scale administration and evaluation of its psychometric properties. Some authors sometimes label these stages differently or group substages into different clusters (see Kyriazos and Stalikas [15] for an overview of the scale development process described by multiple sources).

A few scholars in psychology have designed activities to engage students in scale development / psychometric courses. For instance Mandernach and Hackathorn [20] invited students to analyze psychometric information in popular media to engage discussions around statistical evidence and onboard them on topics related to psychological measurement.

EduCHI'22, April 30–May 1, 2022, New Orleans, LA, USA.

Similarly, students learned about the concept of validity by analyzing the validity of the tests presented in popular magazines (e.g., Cosmopolitan or Men's Fitness): they first completed the measure before answering a set of questions around validity and providing recommendations in which the test developers could demonstrate validity. As reported by Mandernach and Hackathorn [20], taking the test is one of students' favorite course activity "and comparing the validity evidence (or lack thereof) provides a deeper appreciation for the elements necessary for a test to be relevant and useful". Others authors have suggested activities specifically addressing reliability [9], or measurement theory [29].

Building on our experience of teaching methodological classes to a variety of audiences, from HCI/design students to students of other disciplines or practitioners, we developed hands-on teaching activities related to user evaluation methods Amongst them, this teachable moment aims at introducing HCI/design students to principles of scale construction, without requiring prior knowledge. We believe that this activity can be interesting for the community of HCI or design educators, to be applied to their curriculum or to inspire future initiatives.

2 INSTITUTIONAL CONTEXT AND LEARNING OBJECTIVES

The teaching activity we present has been designed in the context of a course on User Evaluation Methods, which is part of the Bachelor curriculum within the department of Industrial Design at the Eindhoven University of Technology. On average, 50 students are enrolled in the course each semester. A requirement to participate in the course is to have completed an introductory course about user-centered design processes and methods. Thus students have been briefly introduced to some standardized evaluation scales relevant for system and product design, such as the System Usability Scale [1] or the AttrakDiff scale [10]. During the very first sessions of our courses, they have also experimented with several scales by filling them out themselves [16]. However, the introductory course did not entail explanations about how these scales are built nor about their underlying psychometrics properties. Over the course of 10 weeks, each session of the course starts with a hands-on group activity supporting the learning experience.

2.1 Learning objectives

This teachable moment contribution is inspired by Kolb's cycle of experiential learning [14], starting with a concrete experience, reflecting on the experience and conceptualizing what they have learned, to move back to an active experimentation by – later in the course - trying out what they have learned and testing the implications of this abstracted knowledge in new situations. Introducing practical applications before presenting theoretical constructs is an effective approach to the teaching of scale validation techniques [22]. The learning objectives are the following:

- Getting familiar with the basic principles of scale construction and psychometrics
- Understanding notions of scale dimensionality, validity, fidelity, and sensitivity
- Knowing the differences between validated and ad hoc evaluation scales and pros/cons of each scale type

Our teaching activity "Let's design a product coolness scale" has been successfully tested during a remote teaching semester, but can of course be used on campus following the same format.

3 DESCRIPTION OF THE ACTIVITY

The "Let's design a product coolness scale" teaching moment is a 1 hour group challenge to design an evaluation scale by following the principles of summated rating scales construction. Guided by step-by-step instructions, groups of students engage in several tasks to learn about scale development processes and requirements, and to better understand

the psychometric properties of standardized scales (validity, fidelity, sensitivity). The duration of the challenge is set to 60 min working on the scale development challenge as a group, followed by a pitch of 5 min in front of the whole class and a debriefing session with lecturers for about 30 min (Figure 1). Working with a class of 50 students approximately, we grouped students in teams of 5-6. The format can be adjusted to a smaller class size by creating groups of 3 students. A point of attention that can considerably extend the duration of the activity is the pitching stage, which is proportional to the number of teams. If you have a large number of groups, a possible solution is to invite only a few teams to pitch.

In the following section, we describe the challenge following the instructions given to students. These instructions were presented in the form of a shared presentation document (Google Slides) divided into as many sections as groups, with one step / task per slide. We illustrate each stage with the results of some student groups.

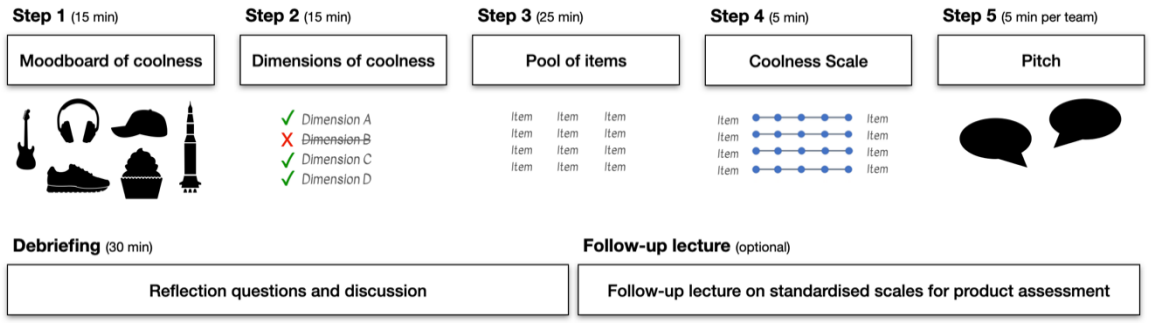

Figure 1: Planning and stages of the Product Coolness Scale teachable moment

3.1 Description of the Challenge as Presented to Students

As we design new products, the compliment that we often crave for is "This is Cool!". Coolness has become a major goal for designers. Think about it: if we aim to craft "coolness" into our products, we should understand what coolness entails and be able to measure it. Claims of the obviousness of what cool is ("I know when I see it") do not help reify the concept; nor are they helpful to designers and researchers [24].

We challenge you to design a standardized scale to measure the coolness of products. By creating this coolness scale, we expect you to learn about scale development processes and requirements, and to better understand the psychometrics properties of standardized scales (validity, fidelity, sensitivity). Understanding the tools we use, and ultimately how not to misuse them, is relevant for design researchers and practitioners alike. As a team, complete your mission by following the step-by-step activities presented in the template. An indicative timing is given for each step, so that you can estimate when to move on to the next stage.

Step 1 – Discover and reflect on what makes things cool (15 min)

You can benchmark a few products that you/people think are "cool" and ask yourself what are the sub-dimensions of coolness. You can base this on your experience, or try to distinguish dimensions of coolness in the explanation people give (for instance product customer reviews).

1.1. Copy-paste here a couple of pictures of cool products (you can search for "cool products" online)

1.2. Annotate them with what makes them cool. Think both about their characteristics and the emotions / experiences triggered by coolness.

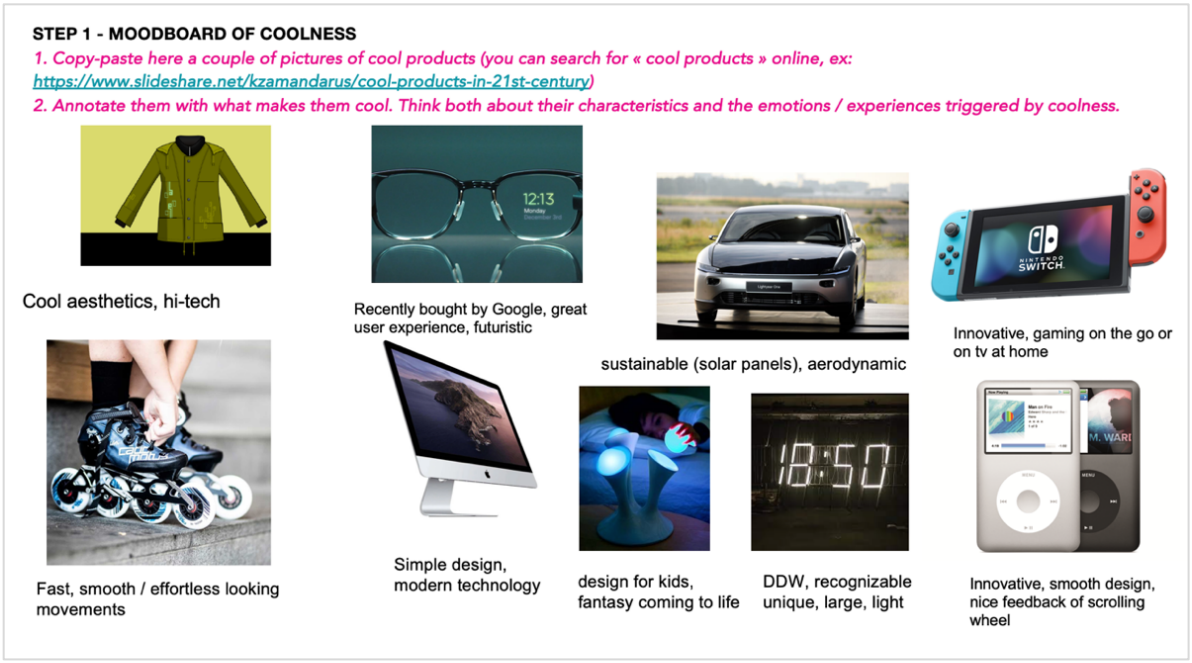

Figure 2: Moodboard of coolness created by a group of students as part of the activity

Step 2 – Define the dimensions of coolness (15 min)

Now that you have found inspiration and scoped the concept of coolness, attempt to abstract the concept to dimensions.

2.1. List all possible dimensions of coolness (free brainstorm).

2.2. Then, discuss them as a team, and select the ones that you think are relevant to measure the construct of "coolness".

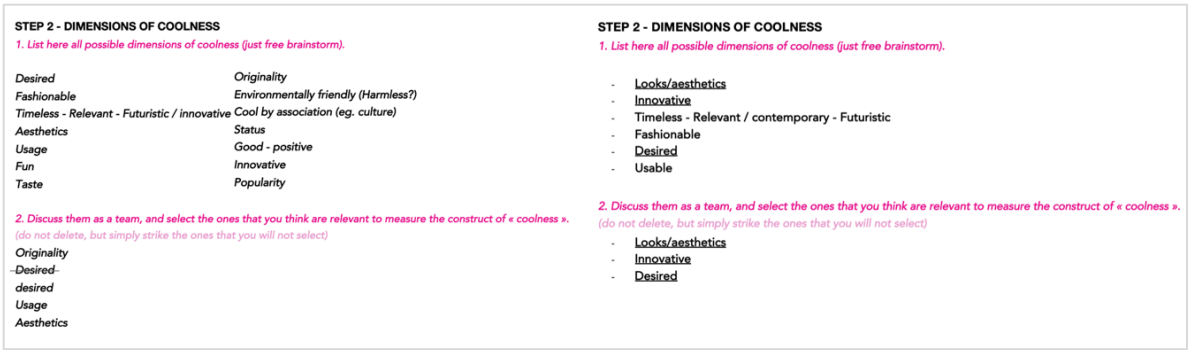

Figure 3: Examples of dimensions of coolness created by two group of students

Step 3 – Creating the items of your scale (25 min)

3.1. For each dimension of coolness you identified, create a set of items to measure the dimension. It is OK and even desirable to have too many items at the start.

3.2. Decide on the format of your items: how should people give their answers? (think of traditional Likert scales for instance, statements with an agreement scale, contrasting words, or other formats)

3.3. It is up to you to decide how many items your scale should have at this stage. If you have too many, reduce your number of items to keep the most relevant ones only.

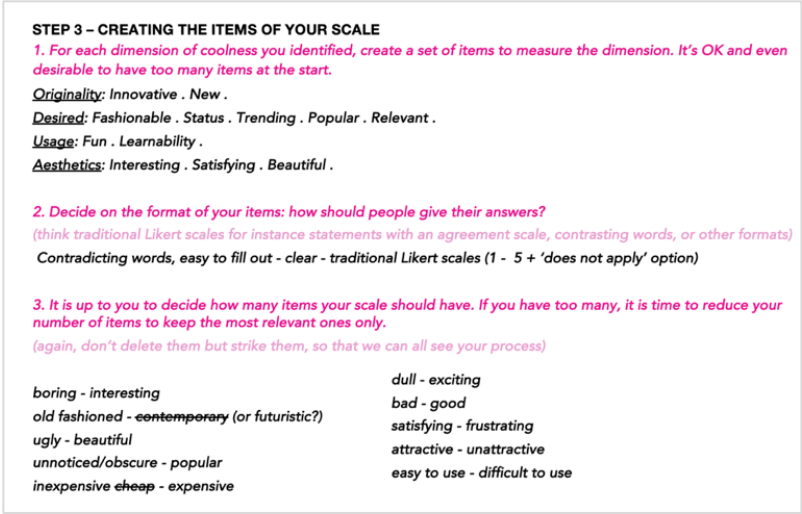

Figure 4: Students create the pool of items composing their scale and decide on the format of the scale

Step 4 – Assemble your scale (5 min)

Insert your final scale here, as would be presented to respondents. Do not indicate the names of the dimensions the items belong to and do not share them yet with other teams.

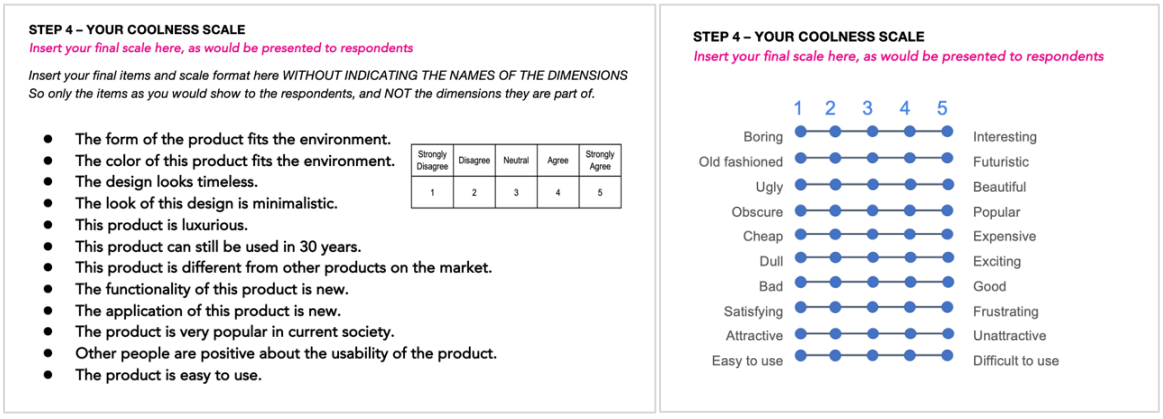

Figure 5: resulting coolness scales created by two groups of students

Step 5 – Pitch and test the face validity (5 min per team)

We invited students to pitch their process by starting at the end: showcasing the final scale they produced. This allowed us to introduce the notion of *face validity* as the extent to which a tool appears to measure what it is supposed to measure. Students could experience it hands-on by asking students from other groups to try to name the dimensions covered by each item. The students then explained their process by showing the slideshow.

3.2 Reflection Questions

Following the pitches, we asked students to reflect on the following aspects and to share their perspectives with the class.

Step 1 – Besides the moodboard you created, can you think of other sources to use to define the dimensions of coolness?

Step 2 – How much overlaps vs. differences were there in what each of you considered as a cool product?

Step 3.2 – Why did you choose this specific response scale format? What are the pros/cons as compared to the ones chosen by the other teams?

Step 3.3 – What do you think is the appropriate number of items to include in such a scale? Do you know any techniques that can help you in deciding which items to keep or to filter out?

Step 4 – Did you create instructions to accompany your scale? Why or why not? Which role do you think the instructions play in the administration of such a scale?

3.3 Debriefing and Showcase of Existing Coolness Scales and their Development

After the reflection questions, we conducted a debriefing on the case by presenting how coolness has been conceptualized in the literature (Table 1). We discussed how to review the literature to define and scope a concept and explained the notion of construct validity. We also showcased two existing coolness scales [2,24]. The COOL questionnaire by Sundar, Tamul and Wu [24] is a 19 items instruments composed of 3 subdimensions (subculture, attractiveness and originality) and presented under the form of 9-points Likert scales. The COOL questionnaire by Bruun, Raptis, Kjeldskov and Skov [2] entails 16 items in 3 subdimensions (desirability, rebelliousness and usability) accompanied by a single item generic measure of coolness reading "this device is cool".

Table 1: Dimensions of coolness identified in the literature

Authors	Definition
Holtzblatt, 2011 [12]	Cool products bring **joy** to our lives and contribute to our personal **feelings of accomplishment, connection with others, identity and delightful experiences**
Fitton et al. 2012 [7]	'Being Cool', by 'Doing Cool Things' and by 'Having Cool Stuff'. For teenagers, cool is **rebellious**, **antisocial**, **retro**, **authentic**, **rich** and **innovative**
Culén & Gasparini, 2012 [4]	Perceived coolness is affected by factors such as **fun**, **mastery**, **adding value**, **useful**, **successful**, **self-presentation** and **innovative**.
Raptis, Kjeldskov, & Skov, 2013 [21]	(1) Coolness consists of **inner and outer cool,** (2) Coolness is **recognized immediately** and (3) Coolness is **grounded in people's communities**
Sundar, Tamul, & Wu, 2014 [24]	Coolness of a product can be conceptualized as a matter of **originality**, **attractiveness** and **subcultural appeal**.
Farnsworth et al., 2014 [5]	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 [25]	Brands or objects that **diverge from the norm** (have increased autonomy) in a way that seems appropriate are perceived as cool

By analyzing these scales in detail, we explained to students that there is not a single way to measure a construct: distinct scales might include different dimensions (especially if they rely on a different theoretical model), items, scale format and instructions. The COOL questionnaire scale by [2] also gives the opportunity to talk about the pros and cons of single-item measures as it includes an additional item "this device is cool" as a generic and potential single-item measure of coolness.

The two existing product coolness scale in the literature both rely on Likert scales. However, since the student groups used different response scale formats (Likert scale, semantic differential and a form of visual analogue scale), it was easy to introduce a discussion around the landscape of existing formats as well as typical debates around response scales (e.g., presence of a middle option). We raised awareness on the fact that the format of a response scale can influence the response just as much as the format of the question itself. In the follow-up lecture, we introduce additional scales useful in product design, which represent a variety of formats [10,18,19,23,26].

4 SUCCESSES AND CHALLENGES

4.1 Insights from Students Reflections on the Course

At the end of the activity, we asked students to give their first impression on this exercise in the chat. Many expressed that the challenge was "hard", "difficult", or "harder than I thought it would be but also really fun". The second part was considered harder than the mood board exercise. Part of the difficulties was that there was "a lot of different opinions to manage" and "everyone has different ideas of cool". Similarly, step 3 was described as tedious and challenging. Students appreciated the flow and step-by-step structure, which they humoristically described as "cool". Most students stated that they had now a better understanding of evaluation scales, which they encountered or used already without questioning how these are build. They also admitted to have misused standardized scales in the past because they knew nothing about psychometrics qualities: they often removed or transformed items themselves to better fit the context of their evaluation, did not duplicate the instructions of the scale precisely, or quickly translated items in the language of the target audience without any consideration for validity. In the course evaluation form, students have stated that they could immediately put the methods and concepts discovered during the course into practice in their design studio projects.

4.2 Insights from the Instructors Perspectives

Our key objective was to design an engaging activity to provide students with a basis foundation for understanding the development of standardized evaluation scales. Although we did not conduct a comparative study between this active way of teaching students about scale development and the more classical ways we have adopted in the past years as instructors in various HCI or design courses, we are satisfied with the activity outcomes and perceived engagement and understanding by the students. Previous literature on teaching (in psychology) showed that these types of hands-on exercises produced better understanding of concepts related to scale development, higher confidence in the topic overall and a better classroom atmosphere [22,28]

4.2.1 Timing and activities

Despite the virtual format, we felt a high level of energy and commitment to the task, reinforced by the tight time constraints of the product coolness scale challenge. More time might allow the students to do the task more thoroughly, yet our observations show that some stages might be discouraging if they stretch in time. This is the case for the

dimension selection or the item generation stages, which both involved some debates that challenged the group dynamics at times. Additionally, the learning objective is not to end up with a perfect scale but to discover in a compact format some of the stages of scale development and associated theoretical concepts. Obviously, we made it clear to the students that the activities conducted were a very simplified way of developing a scale, with some steps being unconventional for the sake of exploring a concept (e.g., step 1 moodboard of coolness to discuss construct validity). The 1.1 activity is for instance inspired by curated collections of cool artefacts (like the cool-tools.org by Kelly [13]), and the Cool Wall technique used by Fitton et al. [6] asking teenagers to sort out products in 4 categories: seriously un-cool, un-cool, cool, sub-zero. To simplify and gain time, we only asked our students to find cool products and reflects on what makes them cool. Looking at what is the opposite of cool is another interesting way we could have suggested to students to define and scope the construct. We later discussed how to use literature to define a construct.

4.2.2 Structure and guidance

Mandernach and Hackathorn [20] briefly report on an activity where students develop a quiz to measure mastery of learning objectives within their course. The learning objective of this exercise is similar to ours: introducing students to the construction of effective assessment instruments. The authors however did not provide any guidance to their students, simply a requirement to incorporate three different item formats. It is unclear what previous knowledge their (psychology) students had about scale construction, but we believe that in our course - with a student population not having any background skills on assessment - this might have been too overwhelming or disturbing.

In addition to the step-by-step guidance, we recognized the added value of our template slides. These support the documentation of the decisions made (we instructed the students to strike the items not selected without deleting them) and created the pitch presentation automatically, in a form that was comparable between groups. No time was therefore wasted on creating a presentation, and the class could focus on the content and the discussions.

4.2.3 Unveiling Bad Practices

In our teaching activity, step 3 consists of creating the items of the coolness scale. It has been reported by students as the most challenging and intense step. Overall, they did not generate a number of items way higher than the one they retained for the final scale. As a comparison, the pool of items of the COOL questionnaire [2] initially included 165 items. From our perspective, this addresses a concern that students typically underestimate the complexity of writing effective items, which is reflected in the lack of rigor often seen in ad hoc questionnaires created by undergraduates students. We believe that students realized after this exercise that robust scale development involves multiple studies and iterations, and that their "homemade" scales were often not valid measures of the constructs they intended to measure in their design projects. For their future design processes, they also understood that there is no one size fits all scale, that standardized scales do not always fit [17] and that designers need to think critically about the evaluation methods they use [16,17]. We also mentioned additional common issues related to the use of standardized scales, such as translation or cherry picking of items. A majority of students mentioned that they were not aware that these practices would put the quality of the measure at risk and that they would pay more attention in the future to factors that can hinder the validity or reliability of a measure.

4.2.4 Follow-up Activities

Our teaching activity stops at the face validity stage, and moves on to the reflection questions with the class and the wrap-up lecture. Similar to Webster [27], we could have extended the activity by asking students to pilot test the data

through cognitive interviewing, before collecting data and analyzing the results. Alternatively, students could take the role of end-users and test out each scale by rating the coolness of a given product (as done in the self-exploration of methods by [16]). Both approaches surely would have brought additional learning outcomes. Our activity was however part of a bigger "User evaluation method" course and the timeframe did not allow for such an assignment. Mandernach and Hackathorn [20] tackled this by including a peer-evaluation stage in their activity; students exchanged their scales with another and gave constructive feedback to improve the instrument. After an iteration, students distributed their scale to ten classmates and completed an analysis of the results before writing a report that critically analyzed the effectiveness, value and relevance of their instrument. The
authors do not specify how much time this required, but it remains realistic to include within a single session providing that some healthy breaks are planned. In our case, step 5 could also encompass, without too much effort or time, a demonstration of cognitive interviewing used in the context of scale development to assess the clarity of items and contribute to face validity [3,8].

In this Teachable Moment contribution, we have shared a 2-hours interactive teaching activity on the topic of scale development and introduction to psychometric concepts. This activity can be adapted to any similar course, with an audience of students unfamiliar with the basics of scale development. We are enthusiastic to see how other lecturers will appropriate and adjust the format of our teaching activities to their own practices.

ACKNOWLEDGMENTS

We thank the students of the DDB140 User Evaluation Methods Elective (Eindhoven University of Technology, Industrial Design Department) who participated in this teaching activity and provided valuable feedback on it.

REFERENCES

[1] John Brooke. 1996. SUS: A "Quick and Dirty" Usability Scale. In P.W. Jordan, B. Thomas, B.A. Weerdmeester, and A.L. McClelland, eds., *Usability Evaluation In Industry*. Taylor and Francis, London. DOI:10.1201/9781498710411-35

[2] Anders Bruun, Dimitrios Raptis, Jesper Kjeldskov, and Mikael B. Skov. 2016. Measuring the coolness of interactive products: The COOL questionnaire. *Behaviour and Information Technology* 35, 3: 233–249. DOI:10.1080/0144929X.2015.1125527

[3] Miguel Castillo-Díaz and José Luis Padilla. 2013. How Cognitive Interviewing can Provide Validity Evidence of the Response Processes to Scale Items. *Social Indicators Research* 114, 3: 963–975. DOI:10.1007/s11205-012-0184-8

[4] Alma Leora Culén and Andrea Alessandro Gasparini. 2012. Situated techno-cools: Factors that contribute to making technology cool in a given context of use. *PsychNology Journal* 10, 2: 117–139.

[5] Carol Farnsworth, Karen Holtzblatt, Shantanu Pai, et al. 2014. Measuring product "coolness." *Extended Abstracts of the 2014 CHI Conference on Human Factors in Computing Systems (CHI EA '14)*, ACM New York, NY, USA, 893–896, DOI:10.1145/2559206.2559968.

[6] Dan Fitton, Janet C. Read, Matthew Horton, Linda Little, Nicola Toth, and Yukang Guo. 2012. Constructing the cool wall: A tool to explore teen meanings of cool. *PsychNology Journal* 10, 2: 141–162.

[7] Daniel Fitton, Matthew Horton, Janet C Read, Linda Little, and Nicola Toth. 2012. Climbing the cool wall: exploring teenage preferences of cool. *Proceedings of the 2012 CHI Conference on Human Factors in Computing Systems*, ACM New York, NY, USA, 2093–2098.

[8] Floyd J. Fowler, Stephanie J. Lloyd, Carol A. Cosenza, and Ira B. Wilson. 2016. Coding Cognitive Interviews: An Approach to Enhancing the Value of Cognitive Testing for Survey Question Evaluation. *Field Methods* 28, 1: 3–20. DOI:10.1177/1525822X14549921

[9] J. Ronald Gentile. 2000. An exercise in unreliability. *Teaching of Psychology* 27: 210–212.

[10] Marc Hassenzahl, Michael Burmester, and Franz Koller. 2003. AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. 187–196. DOI:10.1007/978-3-322-80058-9_19

[11] Jeffrey D. Holmes. 2008. Teaching Traits and Constructs through Scale Construction: A Need and a Strategy. *Teaching of Psychology* 35, 2: 86–90. DOI:10.1080/00986280801978392

[12] Karen Holtzblatt. 2011. What makes things cool? Intentional design for innovation. *Interactions* 18, 6: 40–47. DOI:10.1145/2029976.2029988

[13] Kevin Kelley. 2013. *Cool Tools*. KK.

[14] David A. Kolb. 2014. *Experiential learning: Experience as the source of learning and development*. FT Press.

[15] Theodoros A. Kyriazos and Anastasios Stalikas. 2018. Applied Psychometrics: The Steps of Scale Development and Standardization Process. *Psychology* 09, 11: 2531–2560. DOI:10.4236/psych.2018.911145

[16] Carine Lallemand. 2021. Creative Pedagogical Activities for User Evaluation Methods Courses. In *3rd Annual Symposium on HCI Education (EduCHI'21), May 15, 2022, Virtual*. ACM, New York, NY, USA, 11 pages.

[17] Carine Lallemand and Vincent Koenig. 2017. "How could an intranet be like a friend to me?" - Why standardized UX scales don't always fit. *Proceedings of ACM ECCE conference (ECCE '17)*, Association for Computing Machinery New York, NY, USA, 9–16, DOI:10.1145/3121283.3121288.

[18] Carine Lallemand and Vincent Koenig. 2020. Measuring the Contextual Dimension of User Experience: Development of the User Experience Context Scale (UXCS). *Proceedings ofthe 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society (NordiCHI '20)*, Association for Computing Machinery New York, NY, USA, DOI:10.1145/3419249.3420156.

[19] Bettina Laugwitz, Theo Held, and Martin Schrepp. 2008. Construction and evaluation of a user experience questionnaire. In A. Holzinger, ed., *Lecture Notes in Computer Science*. Springer-Verlag, Berlin Heidelberg, 63–76. DOI:10.1007/978-3-540-89350-9_6

[20] Jean Mandernach and Jana Hackathorn. 2011. Activities for engagement in a psychometric coursee. *Best practices in psychology (Student Engagement vol. 2)*, Society of Teaching of Psychology, 125–129.

[21] Dimitrios Raptis, Jesper Kjeldskov, and Mikael Skov. 2013. Understanding "cool" in human-computer interaction research and design. *Proceedings of the 25th Australian Computer-Human Interaction Conference (OzCHI'13)*, 53–62, DOI:10.1145/2541016.2541032.

[22] Robert C. Reinehr. 1991. Demonstrating Personality Scale Validation Procedures. *Teaching of Psychology* 18, 4: 241–242. DOI:10.1207%2Fs15328023top1804_14

[23] Klaus R. Scherer. 2005. What are emotions? and how can they be measured? *Social Science Information* 44, 4: 695–729. DOI:10.1177/0539018405058216

[24] S. Shyam Sundar, Daniel J. Tamul, and Mu Wu. 2014. Capturing "cool": Measures for assessing coolness of technological products. *International Journal of Human Computer Studies* 72, 2: 169–180. DOI:10.1016/j.ijhcs.2013.09.008

[25] Caleb Warren and Margaret C. Campbell. 2014. What makes things cool? how autonomy influences perceived coolness. *Journal of Consumer Research* 41, 2: 543–563. DOI:10.1086/676680

[26] David Watson, Lee A. Clark, and Auke Tellegen. 1988. Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology* 54, 6: 1063–1070. DOI:10.1037//0022-3514.54.6.1063

[27] Sandra K Webster. 2001. Teaching Psychometrics in South Korea through a Reunification Attitude Scale Class Project. *Annual Meeting of the American Psychological Association*.

[28] Richard Wesp and Sussie Ehsun. 2005. Teaching the principles of test validation in introductory psychology. *Teaching of Psychology* 32: 234–236.

[29] Steven R. Wininger. 2007. Measuring marbles: Demonstrating the basic tenets of measurement theory. *Teaching Statistics* 29: 57–59.