SAGE researchmethods cases

From Replication to Further Scale Development: A Case Study of Incorporating an Eco-Friendliness Dimension in an Existing Consumer Experience Scale

Abstract

This case study describes how we developed a new dimension for measuring eco-friendliness in an existing brand experience measurement scale. We also replicated the original scale to test its generalizability in other countries. Extensions and replications of existing frameworks are important for validating current theories and testing their reliability, even though these research approaches are not yet commonly applied in marketing and business research. In this case study, we describe the research stages that we went through when replicating and extending the scale in two different countries: a small Western country and a large non-Western country. This case study also briefly presents the statistical methods we used (e.g., principal components analysis, confirmatory factor analysis, and structural equation modeling). Replications and extensions of existing frameworks with new dimensions for sustainability-related constructs could aid the field of business and management to further develop more theories on advancing sustainable business.

Learning Outcomes

By the end of this case, students should be able to

- Understand the importance of doing replication research in different countries and with different product types at different points in time
- Understand the research stages for replicating an earlier study in another context and country to verify the generalizability of the original research findings
- Understand how a scale can be further developed and extended for transdisciplinary sustainability research purposes
- Understand the relevance of utilizing research methods from marketing and business research in the field of sustainability research

Project Overview and Context

Recognizing the important role that brands can have in consumers' purchase decisions and the role of strong global brands in the markets, we believed it essential to study how global brands could influence markets' development toward a more sustainable and eco-friendly direction. This led to our research on the kinds of eco-friendly brand experiences consumers have. The research design was based on an existing brand experience measurement scale (BXS) (Brakus, Schmitt, & Zarantonello, 2009). We tested and replicated the BXS in two countries and cultural contexts other than that in which it was originally developed, including both a Western developed economy and a non-Western emerging economy. In the same countries, we also tested an extended version of the scale that included items for measuring the eco-friendliness of brand experiences.

Replication Research

Currently, replication research is underappreciated compared with the development of new frameworks and theories in business and marketing science; however, in the natural sciences, it is understood that making reliable generalizations requires studies to be verified and replicated in different settings and at different time points (Hubbard & Armstrong, 1994). When assessing the relevance and generalizability of research findings, criteria include the replicability, validity, and usefulness of the results (Armstrong, 2003). Raymond Hubbard and Murray Lindsay (2013a) have raised the aspect of significant sameness, which occurs when a research study is able to successfully replicate the results of an earlier study. However, replication studies that do not result in findings similar to those of the original study also offer important knowledge for the further development of associated theories (Hubbard & Lindsay, 2013b).

Successful replication research, in which original findings are supported by other studies in other contexts, verify the external and construct validity of original studies and support further generalizations (Armstrong, 2003; Hubbard & Lindsay, 2013b). External validity refers to the extent to which the findings of the research can be applied to other locations or points of time (i.e., how generalizable the findings are). Construct validity refers to the extent to which the measurement instrument used in the research measures the real-world phenomena it claims to measure. So-called conceptual replication studies use conceptual frameworks from earlier studies but may introduce new independent variables (Raman, 1994).

Although replication research at times may not be considered important, it is required for theory testing to verify equivalent effect sizes, which reinforce earlier studies' findings (Rossiter, 2003). In our study, we replicated the BXS to verify its validity and applicability for high-tech brands and hence contributed to strengthening the existing theories (Easley, Madden, & Dunn, 2000; Evanschitzky & Armstrong, 2013).

Testing a Western Scale in a Non-Western Country

Before using a measurement scale that has been developed in a Western national context in a non-Western country, the scale should be replicated and tested in the non-Western context; otherwise, the measurement items may have low validity and reliability in the non-Western setting. In business and marketing science, there have been calls for more international replication research and testing of frameworks (Evanschitzky & Armstrong, 2013; Madden, Easley, & Dunn, 1995; Uncles & Kwok, 2013).

In consumer research, the focus is often on single countries, most commonly Western countries, especially the United States (Askegaard & Linnet, 2011; Steenkamp & Baumgartner, 1998). Søren Askegaard and Jeppe Trolle Linnet (2011) have observed that the U.S. middle class is often used for sampling, resulting in a research bias that limits extensive generalizations based on findings. It should be considered that social and cultural aspects, especially, of Western and non-Western countries may differ to that extent that they can influence results (Askegaard & Linnet, 2011; Jafari, Fırat, Süerdem, Askegaard, & Dalli, 2012). Brand measurements made in Western and non-Western countries do not necessarily result in comparable

Page 3 of 19

From Replication to Further Scale Development: A Case Study of Incorporating an Eco-Friendliness Dimension in an Existing Consumer Experience Scale

frameworks and results, especially if the measures have been created and verified only in Western countries (Lehmann, Keller, & Farley, 2008).

Brand measures like brand trust (Hegner & Jevons, 2016) and branded product meanings (Strizhakova, Coulter, & Price, 2008) have previously been tested in cross-national contexts; however, as we started to test and extend the BXS, it had not yet been tested, replicated, or extended in a cross-national context. In our research, we focused on consumers in an advanced Western country (Finland) and a non-Western country with an emerging economy (India) (International Monetary Fund, 2013, 2016).

Why Conduct Research on How Consumers Experience Brands?

Brands encapsulate complex product information and, thus, support consumers' decision making in their purchasing processes by helping them recognize brands that they have earlier experienced and learned to trust (Baker & Hart, 2007). As consumers have numerous options from which to select in the market, many consumers have been found to base their selections on brands' trustworthiness and familiarity (Keller, 2013; Mohr, Sengupta, & Slater, 2010).

Research has shown that in both developed Western and developing non-Western countries, consumers' product selection is based on the quality reflected by the product's brand, which is one way of reducing the potential risks associated with a product (Erdem & Swait, 2004; Holt, Quelch, & Taylor, 2004; Steenkamp, Batra, & Alden, 2003; Strizhakova, Coulter, & Price, 2011; Zhou, Su, & Bao, 2002). Although brands are important for consumers globally, previous research has not yet extensively studied whether the experiences brands evoke among consumers in Western and non-Western contexts are the same. Thus, in our research, we set out to test whether the findings from earlier studies can be replicated in different national contexts and whether the results are comparable.

Consumers select and purchase brands not only for functional and emotional reasons but based on their perceptions of how responsible the brands' companies are (Kotler, 2011). Brand companies benefit, therefore, from considering consumers' brand experiences related to eco-friendliness when planning new eco-friendly products. Sustainable consumption and sustainable production require new measures for tracking consumers' views. Sustainable consumption and production have been listed as one of the sustainable development goals (SDG) of the United Nations under the heading SDG 12 "Ensure Sustainable Consumption and Production Patterns" (United Nations, 2018). In this context, sustainable consumption and production refer to advancing the consumer markets so that products are manufactured resource- and energy-efficiently in sustainably managed conditions with the target to reduce future economic, environmental, and social costs (United Nations, 2018). Definition and development of constructs for measuring the eco-friendliness of consumers' brand experiences had not been explored extensively in the business and marketing literature.
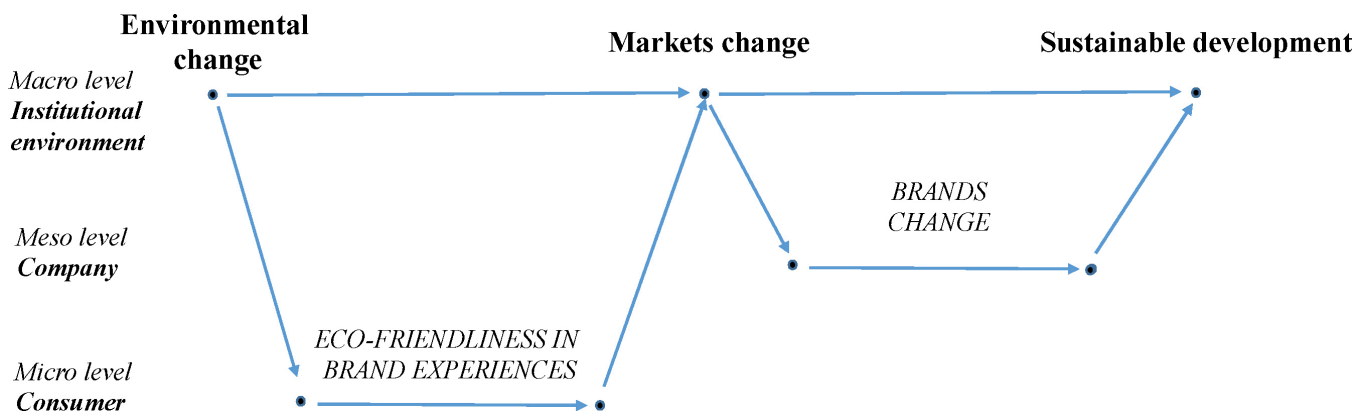
Hence, in the start of our replication study, we had a topic that was under-developed (replication of BXS and lack of measures for eco-friendliness) and which we considered important. We also recognized early on

that there may be challenges when linking micro-, consumer-level research findings to higher, meso-level processes in individual companies and macro-level processes involving the development of the business environment. This requires a transdisciplinary perspective to link consumer research and market research findings to the business and sustainability literature. This multi-level theorizing was an extension of a pure replication study and in combination with transdisciplinary nature drove us to collaborate with researchers representing various fields.

In the paper published in *Sustainability*, we theorized that the linkage of micro-level development to meso- and macro-level developments presents how eco-friendly brands could drive sustainable development (Saari, Baumgartner, & Mäkinen, 2017). In this article, we reported our findings concerning the replication of the extended version of the BXS we tested in Finland and India. In our study, we concentrated on consumers and illustrated through James S. Coleman's (1987) theoretical microfoundations model how a consumer-level brand experience measure could provide input to create sustainable development initiatives in companies. By measuring the eco-friendliness of consumers' brand experiences, companies and policy makers could track how markets are changing and how sustainable development is progressing. Company-level initiatives could then be made visible on the macro-level (see Figure 1). When selecting the countries for the study, we consciously focused on selecting countries that are very different from each other; thus, we selected Finland, which is a small Western country, and India, which is a very large country with a non-Western culture.

Figure 1. Consumer behavior can influence market change and further sustainable development, starting from the micro-level and influencing the meso- and macro-levels.

Source: Adapted from Saari et al. (2017).

Research Design

The objective of our study was to verify and further develop the BXS by testing the scale in two countries that differed from that in which the BXS was developed. In addition, we wanted to extend the measuring of eco-friendliness and measure the differences in how Indian and Finnish consumers experience the eco-

Page 5 of 19

From Replication to Further Scale Development: A Case Study of Incorporating an Eco-Friendliness Dimension in an Existing Consumer Experience Scale

friendliness of global smartphone brands. Our study continued the work that Ulla Saari (2016) did in her PhD research.

The nature of the research was empirical, and the approach was quantitative. We collected data with a web survey conducted in Finland and India and analyzed it with statistical methods. In the web survey, we used the existing set of questions from the earlier research by J. Joško Brakus et al. (2009) and the additional measures of eco-friendliness that Ulla designed in her PhD research on brand experiences. In the next section, we describe in more detail how we progressed in the different stages of the research.

Research Stages for Replicating and Extending a Scale in Another Context

We conducted the research in six clearly defined stages so that we would be prepared for each stage beforehand. This forced us to scrutinize the tasks needed in each stage and consider the theoretical and methodological choices we would be making. We also had a rough estimate of the timeline for the different stages; even though we could not always keep the actual target dates, it helped to keep the momentum going in the project. The stages listed here are described in more detail in the following subsections:

1. development of the extended scale
2. piloting
3. sampling
4. replication of the original scale and testing of the extended scale in two different countries
5. data analysis
6. validation and finalization of the scale to fulfill the model-fit indices

Development of the Extended Scale

First, we decided on the scale items describing eco-friendliness in brand experiences, keeping in mind the key objective of our research, which was to explore whether consumers perceive eco-friendliness as an element of their overall brand experiences. The items were already developed and tested with a Finnish sample in Ulla's doctoral thesis (Saari, 2016). The four dimensions included in the original BXS (Brakus et al., 2009) were used as a starting point for formulating the items for the eco-friendliness construct because the original items in the BXS had already been shown to represent the brand experiences of consumers in one country. In the study reported in *Sustainability*, we focused on testing whether the extended BXS scale with the eco-friendliness construct was valid for brand experiences in both a Western country (Finland) and a non-Western country (India).

When developing a measurement scale for human respondents, one must take certain principles into account in the design of the questionnaire. Not only do the items added in the scale need to be theoretically logical

Page 6 of 19

From Replication to Further Scale Development: A Case Study of
Incorporating an Eco-Friendliness Dimension in an Existing Consumer
Experience Scale

and based on previous research, they also need to be sensible so that the respondents understand them. For finding the theoretical basis for the eco-friendliness dimension in brand experiences, Ulla did extensive literature searches from the business, branding, marketing, and psychometrics literature and studied how previous research had covered eco-friendliness of consumer products and measuring of consumers' perceptions of eco-friendliness.

As the BXS scale measures attitudes of individual consumers toward brands, the techniques and theories related to the scale design are very similar to psychological measurements. To get more insight from the scientific work conducted in psychometrics, Ulla started by reading the classical book of Jum Nunnally and Ira Bernstein (1994) that offers a deep dive into the field of psychometrics. It is a good introduction to the topic and worth glancing through to get the essentials of the challenges in measuring human subjects.

To ensure the validity of the added construct of eco-friendliness, the guidelines proposed by Gilbert A. Churchill (1979) and John C. Mowen and Kevin E. Voss (2008) proved to be helpful. Churchill (1979) focuses specifically on the challenges associated with developing marketing constructs, and Mowen and Kevin (2008) on the validity of measures—that is, that the measures are truly measuring the phenomena they have been designed for.

To ensure that the eco-friendliness construct did not overlap with existing constructs, we tested the eco-friendliness construct in two ways: embedded in the original BXS scale and as an independent construct (Saari, 2016). First, we tested the scale with a pilot and later again with the collected data in the data analysis phase (see Step 5, Data Analysis).

Piloting

The measurement items in the original scale, which were written in English, were used in their original form for India and translated into Finnish for Finland. As required for translations used in studies, the Finnish translations were checked by a native speaker. In the development stage of a survey measure, it is important to test the measure with respondents from the survey target groups.

Before beginning the large-scale data collection via a web survey, we piloted the questionnaire face-to-face with selected Indian and Finnish participants. At that time, Ulla was working at in Information Technology department with both Indian and Finnish employees, so she randomly selected participants for the pilot testing by inviting people one at a time to test the questionnaire in the office kitchen area. The questionnaire was in paper-and-pen format, and while the participants filled out the questionnaire, Ulla stayed close by to collect feedback from the pilot respondents and observe their actions. The feedback from the pilot participants confirmed that the measurement items were acceptable, and only some of the wordings were edited or clarified based on the feedback. For example, a few of the pilot respondents had a hard time answering a negatively worded statement. Based on this feedback and Ulla's observations, we changed this statement into the affirmative format.

Page 7 of 19

From Replication to Further Scale Development: A Case Study of Incorporating an Eco-Friendliness Dimension in an Existing Consumer Experience Scale

Sampling

Our sample comprised 502 respondents in India and 506 respondents in Finland. The participants' demographics were carefully tracked to ensure that the sample reflected the studied populations as closely as possible. In Finland, the respondents were distributed evenly across the country. In India, we decided to concentrate on the urban population, as city dwellers in India tend to be more knowledgeable about smartphone brands. When sampling in different countries, it is important to consider carefully the distribution of the demographics to ensure that the respondents are similarly exposed to the phenomena studied (in this case, brand experiences associated with smartphone brands). (See Saari & Mäkinen, 2017, for more details on the samples used in the research.)

Replication of the Original Scale and Testing of the Extended Scale in Two Countries

The data were collected simultaneously in Finland and India. This ensured that the global brand companies were measured at the same point in time. The extended brand experience scale consists of 16 items. The respondents assessed the items on a 7-point Likert-type scale with an additional eighth response option ("do not know"). The eighth option ensured that respondents did not use the midpoint option (4) when they did not know what to answer, thereby increasing the validity of the responses (Asun, Rdz-Navarro, & Alvarado, 2016).

Data Analysis

We conducted the analyses using Statistical Package for the Social Sciences (SPSS). We evaluated the four scale items measuring eco-friendliness and the four brand experience dimensions in the original BXS by conducting structural equation modeling (SEM) and confirmatory factor analysis (CFA) according to the same research design used by Brakus and colleagues (2009).

One of the requirements for a cross-national construct is that the factor loadings follow the same pattern for the data sets collected in the assessed countries (Steenkamp & Baumgartner, 1998). If the countries differ in terms of economic level and/or cultural background (i.e., Western vs. non-Western), it is particularly important to test the scale in both countries (Jafari et al., 2012); otherwise, the results will not be generalizable (Hubbard & Lindsay, 2002).

The SEM was conducted in two phases: measurement and structural. The model's fit with the data was tested by conducting SEM in SPSS and Amos Graphics. The indices we selected to analyze the fitness of the model included the Tucker–Lewis Index (TLI) and the comparative fit index (CFI), both of which are considered reliable indices. A model is considered to fit well when the values for these indices are larger than 0.90 (Janssens, Wijnen, de Pelsmacker, & Van Kenhove, 2008).

We conducted an exploratory factor analysis (EFA), also referred to as a principal components analysis (PCA), and a CFA on the hypothesized models. EFA is factor analysis method for conducting an initial assessment of the interrelations among the measurement items in a newly defined measurement scale. The suitability of a scale developed in one country can be analyzed using CFA to assess the relationships between the observed variables and the hypothesized constructs (Steenkamp & Baumgartner, 1998). We used SPSS software to evaluate the relationships between the items. The eighth option ("do not know") was treated as missing data, and the data analysis applied the maximum likelihood method with the missing-at-random (MAR) approach (Bentler, 2010).

Later, when we assessed the validity of the model using the collected data, we verified the content and criterion validity, which together form the construct validity (Cook & Beckman, 2006). Construct validity shows whether a scale measures what it has been designed to measure, which is critical when developing and testing theories (Churchill, 1979; Peter, 1981). The discriminant validity of a measure shows whether the measure is sufficiently different from other similar measures (Churchill, 1979). For the construct we designed for eco-friendliness, we tested the discriminant validity using CFA. The validity and reliability were checked using the results of the CFA, the factor loadings, and the average variance extracted (AVE) and composite reliability (CR) values (Hair, Black, Babin, & Anderson, 2010). Cronbach's alpha and CR were also estimated for the model.

We hypothesized two different conceptual models in our studies. We first tested whether the original BXS model could be replicated in both Finland and India. Then, we tested how the eco-friendliness construct fit the revised, extended version of the BXS.

Validation and Finalization of the Scale to Fulfill the Model-Fit Indices

A valid scale should fulfill certain model-fit indices supported by the literature. To assess the model-fit indices, researchers should do extensive reading and use several references to back up their results. We used several indices to analyze and compare the models, including the root mean square error of approximation (RMSEA), the CFI, and the TLI. As models differ with respect to their complexity, variables, sample sizes, and data distributions, it is not possible to define general cutoff values for all models (Markland, 2007). It should be noted that the chi-square is especially sensitive to sample size, so a model that fits well can produce a chi-square close to the sample size (Iacobucci, 2010).

The model-fit indices can be considered acceptable when the TLI and CFI values are greater than 0.90 (Tomarken & Waller, 2005). The literature reports different cutoff values for the RMSEA, so that cutoff values can range from .05 to .06 and the fit can be considered acceptable when the value is .08 (Janssens et al., 2008). It has been found that the RMSEA is not reliable and that the fit worsens as the number of variables in a model increases (Iacobucci, 2010).

When we tested the model fit of the original BXS with the data collected from Finland, the fit indices were mostly acceptable; however, the data collected from India did not fit the model, as indicated by unacceptable
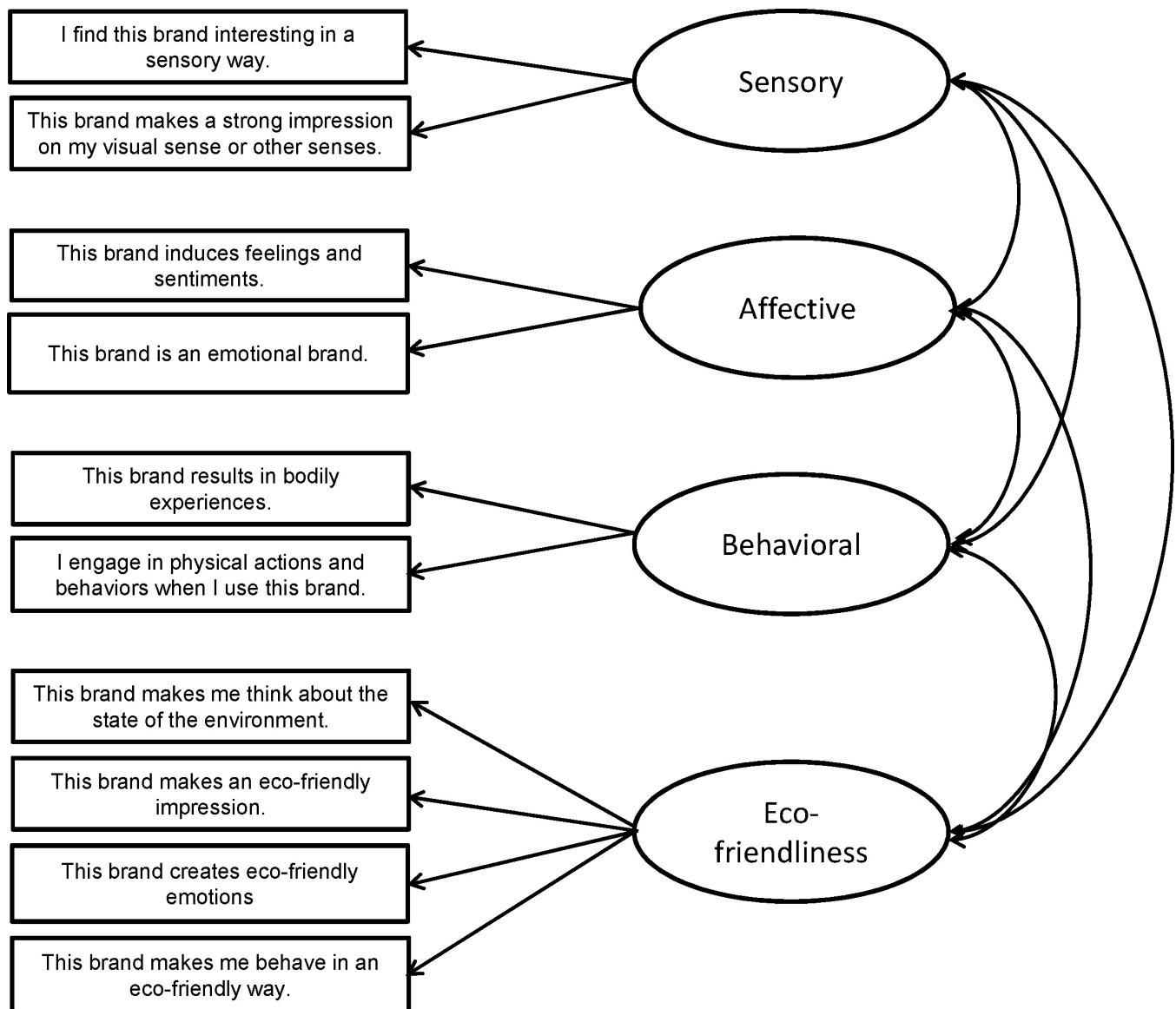
From Replication to Further Scale Development: A Case Study of
Incorporating an Eco-Friendliness Dimension in an Existing Consumer
Experience Scale

fit indices (Saari & Mäkinen, 2017). As a result, we started reducing the original scale by eliminating items. We did this by checking the individual model items that did not load on the common components. The Kaiser–Meyer–Olkin (KMO) measure was first checked to verify that the EFA could be conducted. The KMO was clearly over the recommended threshold of .50 (Malhotra et al., 2012). In addition, Bartlett's test of sphericity was checked.

The EFA resulted in a three-factor model, but the negatively worded items raised some questions. The data from India fit the model slightly differently than the data from Finland, and the negatively worded items loaded differently again. This indicated that the negatively worded items did not fit in the same model as the other items. When the negatively worded items were taken out of the model, the fit of the model for the shortened scale was acceptable for both countries; however, we went further with the testing to eliminate an additional construct (the one measuring the intellectual dimension), which caused some issues with the model fit for the data from India.

Then, we added the eco-friendliness construct to the successfully replicated and reduced model. This resulted in a model in which the fit indices were acceptable for both countries. As the factor loadings in both countries were all above 0.50, the variables were aligned with the factors (see Figure 2) (Bagozzi & Yi, 2012).

Page 10 of 19

From Replication to Further Scale Development: A Case Study of Incorporating an Eco-Friendliness Dimension in an Existing Consumer Experience Scale

Figure 2. The four-factor model for the reduced BXS with the eco-friendliness construct.

Source: Adapted from Saari et al. (2017).

Practical Lessons Learned

Here we would like to share with you some of the challenges and key learnings from the various stages of the project.

Scale Development

Page 11 of 19

From Replication to Further Scale Development: A Case Study of Incorporating an Eco-Friendliness Dimension in an Existing Consumer Experience Scale

In the scale development phase, when we added the eco-friendliness dimension to the BXS scale, we tried to design the items so that the style of the questions would be similar to the ones in the original questionnaire. However, at the same time we had to make sure that they were clearly differentiated from the other questions so that the respondents would not be confused and so that we were truly measuring the phenomenon of eco-friendliness. This balancing between fit in the overall measurement scale and being distinctly different did not at first seem that complicated, as it looked like we could just add the reference to the environment and eco-friendliness in the scale and that's it. However, in the actual phrasing of the statements, we had to consider the nuances and meanings of all synonyms in the statements. For example, how does one understand eco-friendliness as an emotion? Can one say that a brand induces eco-friendly feelings or sentiments? Are these concepts already used? Finally, we decided to keep the word "emotions" in the statement ("This brand creates eco-friendly emotions"), as it is close to one of the original items, which meant that we could use it also to verify whether there is a distinct difference between this item and the original scale item measuring brand-related emotions. We performed several iterations of all of the items used for measuring eco-friendliness before we were convinced that they were ready for piloting.

Piloting

In the piloting phase, the key challenge was interpreting the pilot respondents' reactions to the questionnaire. Ulla instructed the respondents to think aloud when filling in the questionnaire, especially for questions where they seemed to stop to think, so she could get instant and true feedback on the understandability of the items and the respondents' reactions to the newly added items on eco-friendliness. In addition, she looked at the way respondents used their pen to select options, whether they paused in silence before answering, and whether their facial expressions revealed uncertainty about the question's meaning. The respondents also gave feedback on the statements in the original BXS scale, which raised the question, should we also refine the original scale? However, as we had decided to do a replication of the original scale, we could not change any items in it at this stage. The key point here is to observe closely respondents who do the pilot survey and be open-minded to detect their reactions to survey measures.

Sampling

For the sampling, the biggest challenge was to ensure that in India we could also find a sample that would have enough background information on smartphones and would know the brands that were the focus of the study. Based on market research in India, we learnt that the younger urban age groups had the most smartphones, thus we selected the top five metropolitan areas in India (Mumbai, Delhi, Kolkata, Chennai, and Bangalore) as target cities for this survey (Saari & Mäkinen, 2017).

Replication and Testing

The actual testing of the extended scale with data collected in two countries was challenging due to the

From Replication to Further Scale Development: A Case Study of
Incorporating an Eco-Friendliness Dimension in an Existing Consumer
Experience Scale

various response styles we noted in the results. In India, the respondents used the positive end of the scale more often (over value 4), so that the response means are nearly 5 for most of the statements. In Finland, the responses were closer to the value of 4 or even 3, so very near to the middle point on a 7-point Likert-type scale. This did not come as a total surprise to us, as in Finland people generally do not exaggerate or use too many positive superlatives to describe their feelings or opinions, and we knew that in Indian culture stronger views can be expressed. Thankfully, this topic has been covered in the literature, so we found good references with which we could make sense of the results (see Baumgartner & Steenkamp, 2001; Johnson, Kulesa, Cho, & Shavitt, 2005).

Data Analysis, Validation, and Finalization

The data analysis, validation, and finalization of the scale all merged to an iterative process, whereby we checked the data and then validated it in several cycles to finalize the scale so that the model-fit indices in the SEM literature supported it well enough statistically. In the data analysis phase, we did not rush to any conclusions, but took the time to interpret the data and used it to validate the model options. The main challenge was to fit the data from both of the countries in the replicated original BXS scale. As we describe in the article published in the *Journal of Brand Management* (Saari & Mäkinen, 2017), we had to drop from the BXS scale some items to find a version of the scale that fit the data for both the countries. As a result, we learnt concretely that the replication of scales that have been tested only once and in one country do not necessarily apply automatically for other countries. At this point, this result felt like a true finding that made the whole replication exercise worthwhile.

Reporting Findings

When reporting the findings of a research, it is always important to understand and identify the relevance and implications of your findings. The results of our cross-national testing of the extended measurement scale have theoretical, managerial, and policy implications. We have highlighted in our findings that one important approach in business and social sciences is measuring micro-level phenomena, such as consumers' brand experiences, to analyze macro-level phenomena (Coleman, 1987; Raub, Buskens, & Van Assen, 2011). From a consumer viewpoint, sustainable production is the responsibility of companies and has an impact on sustainable development in society. By measuring individual consumers' perceptions and experiences of how brand companies manage their corporate social responsibility (CSR) and environmental responsibility, as well as how consumers perceive eco-friendliness on the product level, BXS measurements could be used to drive market changes on the macro-level by, for example, policy makers (Saari et al., 2017).

The links between theories created on the macro- and meso-levels have not been tested or operationalized on the micro-level in business and marketing publications. Thus, we decided to publish our research in *Sustainability* to present a balanced approach to the governance of environmental change and sustainable development in which individual consumer behavior and technological infrastructures help to guide industry transformation (Keller, Halkier, & Wilska, 2016; Spaargaren, 2011) as presented in Figure 1. The connections

Page 13 of 19

From Replication to Further Scale Development: A Case Study of Incorporating an Eco-Friendliness Dimension in an Existing Consumer Experience Scale

among all three levels influence sustainable development, and the micro–macro connection is particularly critical for executing sustainability strategies (Liedtke, Buhl, & Ameli, 2013).

Conclusion

We would like to highlight a few points as encouragement and recommendations for readers who might consider conducting a replication and extension study of an existing scale. The development of the extended scales is a worthwhile exercise that allows you to test and extend existing measurement scales with some new theoretical input while building theory based on earlier findings. This approach also helps to build more accurate generalizations in science (Hubbard & Lindsay, 2002). In addition, it is important to conduct business research in a more global context and go beyond Western countries (Jafari et al., 2012).

We also recommend taking the piloting phase seriously and reserve enough time for it. It was truly beneficial for us to get feedback from the pilot respondents and thus verify the quality of the questionnaire before starting a full-scale survey. When testing a replicated or extended scale cross-nationally in two or more countries, we recommend spending the time to verify the results to interpret them correctly. We think that the data analysis is one of the most exciting and important parts in the research project, even though at some points it may seem tedious. When analyzing the data, one can concretely see how the model and theory that one has worked on for a lengthy part of the project actually behaves with the real data.

Finally, we would like to note that the validation of the model fit is not necessarily as easy and straightforward as it may initially seem by reading books on statistical analysis for market research. Many aspects need to be taken into account in the validation phase, especially for SEM models. At this point, it is worthwhile to do extensive reading into guidelines on how to interpret the cutoff values. Cutoff values are like beacons guiding interpretation of the results and need not be taken always as strict rules, although if deviating from them, clear explanations and reasons for interpretation are warranted. Furthermore, the size of the sample, the number of constructs, the number of items, and complexity of the model all need to be carefully considered in the final analysis.

Now, after going through this learning journey into replication research and extending a measurement scale, we encourage you to also consider this as an option. With this approach, you can build more rigorous measurement scales that are applicable and usable in different markets and countries both in Western and non-Western parts of the world. And replication of earlier findings is the core of developing our body of knowledge further, thus truly "we stand on the shoulders of giants" (Sir Isaac Newton, 1676, cited in Adler, 2010), also in *Business & Management*.

Exercises and Discussion Questions

1. Discuss the rationale for conducting replication research, especially in the case of

Page 14 of 19

From Replication to Further Scale Development: A Case Study of
Incorporating an Eco-Friendliness Dimension in an Existing Consumer
Experience Scale

measurement scales.

2. List the various steps in replicating and extending a scale in two new national contexts. Comment on the rationale behind the procedure.

3. What specific aspects should be carefully considered when extending an existing scale?

4. In the development of a new construct, what is the relevance of first studying the generalizability of the existing construct and then creating an additional construct with which one can extend the scale?

5. Reflect on the issues that could result from overlapping content items and negatively worded items in a scale.

6. Why is it important to verify the validity of a scale from different perspectives?

References

Adler, N. J. (2010). Going beyond the dehydrated language of management: Leadership insight. *Journal of Business Strategy*, 31, 90–99. doi:http://dx.doi.org/10.1108/02756661011055230

Armstrong, J. (2003). Discovery and communication of important marketing findings: Evidence and proposals. *Journal of Business Research*, 56, 69–84. doi:http://dx.doi.org/10.1016/S0148-2963(02)00386-7

Askegaard, S., & **Linnet, J.** (2011). Towards an epistemology of consumer culture theory: Phenomenology and the context of context. *Marketing Theory*, 11, 381–404. doi:http://dx.doi.org/10.1177/1470593111418796

Asun, R. A., **Rdz-Navarro, K.**, & **Alvarado, J. M.** (2016). Developing multidimensional Likert scales using item factor analysis: The case of four-point items. *Sociological Methods & Research*, 45, 109–133. doi:http://dx.doi.org/10.1177/0049124114566716

Bagozzi, R., & **Yi, Y.** (2012). Specification, evaluation, and interpretation of structural equation models. *Journal of the Academy of Marketing Science*, 40, 8–34. doi:http://dx.doi.org/10.1007/s11747-011-0278-x

Baker, M., & **Hart, S.** (2007). *Product strategy and management*. London, England: Pearson.

Baumgartner, H., & **Steenkamp, J.** (2001). Response styles in marketing research: A cross-national investigation. *Journal of Marketing Research*, 38, 143–156. doi:http://dx.doi.org/10.1509/jmkr.38.2.143.18840

Bentler, P. M. (2010). SEM with simplicity and accuracy. *Journal of Consumer Psychology*, 20, 215–220. doi:http://dx.doi.org/10.1016/j.jcps.2010.03.002

Brakus, J. J., **Schmitt, B. H.**, & **Zarantonello, L.** (2009). Brand experience: What is it? How is it measured? Does it affect loyalty? *Journal of Marketing*, 73, 52–68. doi:http://dx.doi.org/10.1509/jmkg.73.3.52

Churchill, G. A., Jr. (1979). Paradigm for developing better measures of marketing constructs. *Journal of Marketing Research*, 16, 64–73. doi:http://dx.doi.org/10.2307/3150876

Page 15 of 19

From Replication to Further Scale Development: A Case Study of Incorporating an Eco-Friendliness Dimension in an Existing Consumer Experience Scale

Coleman, J. S. (1987). Microfoundations and macrosocial behavior. In **J. C. Alexander**, **B. Giesen**, **R. Münch**, & **N. J. Smelser** (Eds.), *The micro–macro link* (pp. 153–173). Berkeley: University of California Press.

Cook, D. A., & **Beckman, T. J.** (2006). Current concepts in validity and reliability for psychometric instruments: Theory and application. *The American Journal of Medicine*, 119, 166.e7–166.e16. doi:http://dx.doi.org/10.1016/j.amjmed.2005.10.036

Easley, R. W., **Madden, C. S.**, & **Dunn, M. G.** (2000). Conducting marketing science: The role of replication in the research process. *Journal of Business Research*, 48, 83–92. doi:http://dx.doi.org/10.1016/S0148-2963(98)00079-4

Erdem, T., & **Swait, J.** (2004). Brand credibility, brand consideration, and choice. *Journal of Consumer Research*, 31, 191–198. doi:http://dx.doi.org/10.1086/383434

Evanschitzky, H., & **Armstrong, J.** (2013). Research with in-built replications: Comment and further suggestions for replication research. *Journal of Business Research*, 66, 1406–1408. doi:http://dx.doi.org/10.1016/j.jbusres.2012.05.006

Hair, J. F., **Black, W. C.**, **Babin, B. J.**, & **Anderson, R. E.** (2010). *Multivariate data analysis: A global perspective* (7th ed.). Upper Saddle River, NJ: Pearson.

Hegner, S. M., & **Jevons, C.** (2016). Brand trust: A cross-national validation in Germany, India, and South Africa. *Journal of Product & Brand Management*, 25, 58–68. doi:http://dx.doi.org/10.1108/JPBM-02-2015-0814

Holt, D., **Quelch, J.**, & **Taylor, E.** (2004). How global brands compete. *Harvard Business Review*, 82, 68–75.

Hubbard, R., & **Armstrong, J.** (1994). Replications and extensions in marketing: Rarely published but quite contrary. *International Journal of Research in Marketing*, 11, 233–248. doi:http://dx.doi.org/10.1016/0167-8116(94)90003-5

Hubbard, R., & **Lindsay, R. M.** (2013a). From significant difference to significant sameness: Proposing a paradigm shift in business research. *Journal of Business Research*, 66, 1377–1388. doi:http://dx.doi.org/10.1016/j.jbusres.2012.05.002

Hubbard, R., & **Lindsay, R. M.** (2013b). The significant difference paradigm promotes bad science. *Journal of Business Research*, 66, 1393–1397. doi:http://dx.doi.org/10.1016/j.jbusres.2012.05.004

Hubbard, R., & **Lindsay, R. M.** (2002). How the emphasis on "original" empirical marketing research impedes knowledge development. *Marketing Theory*, 2, 381–402. doi:http://dx.doi.org/10.1177/147059310200200408

Iacobucci, D. (2010). Structural equations modeling: Fit indices, sample size, and advanced topics. *Journal*

Page 16 of 19

From Replication to Further Scale Development: A Case Study of
Incorporating an Eco-Friendliness Dimension in an Existing Consumer
Experience Scale

of Consumer Psychology, 20, 90–98. doi:http://dx.doi.org/10.1016/j.jcps.2009.09.003

International Monetary Fund. (2013, April). *World economic outlook*. Retrieved from http://www.imf.org/external/pubs/ft/weo/2013/01/pdf/text.pdf

International Monetary Fund. (2016, April). *World economic outlook*. Retrieved from http://www.imf.org/external/pubs/ft/weo/2016/01/pdf/text.pdf

Jafari, A., **Fırat, F.**, **Süerdem, A.**, **Askegaard, S.**, & **Dalli, D.** (2012). Non-western contexts: The invisible half. *Marketing Theory*, 12, 3–12. doi:http://dx.doi.org/10.1177/1470593111424183

Janssens, W., **Wijnen, K.**, **de Pelsmacker, P.**, & **Van Kenhove, P.** (2008). *Marketing research with SPSS*. London, England: Pearson.

Johnson, T., **Kulesa, P.**, **Cho, Y. I.**, & **Shavitt, S.** (2005). The relation between culture and response styles—Evidence from 19 countries. *Journal of Cross-Cultural Psychology*, 36, 264–277. doi:http://dx.doi.org/10.1177/0022022104272905

Keller, K. L. (2013). *Strategic brand management: Building, measuring, and managing brand equity* (4th ed.). London, England: Pearson.

Keller, M., **Halkier, B.**, & **Wilska, T. A.** (2016). Policy and governance for sustainable consumption at the crossroads of theories and concepts. *Environmental Policy and Governance*, 26, 75–88. doi:http://dx.doi.org/10.1002/eet.1702

Kotler, P. (2011). Reinventing marketing to manage the environmental imperative. *Journal of Marketing*, 75, 132–135. doi:http://dx.doi.org/10.1509/jmkg.75.4.132

Lehmann, D., **Keller, K.**, & **Farley, J.** (2008). The structure of survey-based brand metrics. *Journal of International Marketing*, 16, 29–56. doi:http://dx.doi.org/10.1509/jimk.16.4.29

Liedtke, C., **Buhl, J.**, & **Ameli, N.** (2013). Microfoundations for sustainable growth with eco-intelligent product service-arrangements. *Sustainability*, 5, 1141–1160. doi:http://dx.doi.org/10.3390/su5031141

Madden, C. S., **Easley, R. W.**, & **Dunn, M. G.** (1995). How journal editors view replication research. *Journal of Advertising*, 24, 77–87. doi:http://dx.doi.org/10.1080/00913367.1995.10673490

Markland, D. (2007). The golden rule is that there are no golden rules: A commentary on Paul Barrett's recommendations for reporting model fit in structural equation modelling. *Personality and Individual Differences*, 42, 851–858. doi:http://dx.doi.org/10.1016/j.paid.2006.09.023

Mohr, J., **Sengupta, S.**, & **Slater, S.** (2010). *Marketing of high-technology products and innovations*. Upper Saddle River, NJ: Pearson.

From Replication to Further Scale Development: A Case Study of Incorporating an Eco-Friendliness Dimension in an Existing Consumer Experience Scale

Mowen, J. C., & **Voss, K. E.** (2008). On building better construct measures: Implications of a general hierarchical model. *Psychology & Marketing*, 25, 485–505. doi:http://dx.doi.org/10.1002/mar.20221

Nunnally, J. C., & **Bernstein, I. H.** (1994). *Psychometric theory* (3rd ed.). New York, NY: McGraw-Hill.

Peter, J. P. (1981). Construct validity—A review of basic issues and marketing practices. *Journal of Marketing Research*, 18, 133–145.

Raman, K. (1994). Inductive inference and replications—A Bayesian perspective. *Journal of Consumer Research*, 20, 633–643.

Raub, W., **Buskens, V.**, & **Van Assen, M.** (2011). Micro–macro links and microfoundations in sociology. *Journal of Mathematical Sociology*, 35, 1–25. doi:http://dx.doi.org/10.1080/0022250X.2010.532263

Rossiter, J. R. (2003). Qualifying the importance of findings. *Journal of Business Research*, 56, 85–88. doi:http://dx.doi.org/10.1016/S0148-2963(02)00387-9

Saari, U. (2016). *Eco-friendliness in the brand experience of high-tech products* (PhD thesis), Tampere University of Technology, Tampere, Finland.

Saari, U. A., **Baumgartner, R. J.**, & **Mäkinen, S. J.** (2017). Eco-friendly brands to drive sustainable development: Replication and extension of the brand experience scale in a cross-national context. *Sustainability*, 9, 1286. doi:http://dx.doi.org/10.3390/su9071286

Saari, U. A., & **Mäkinen, S. J.** (2017). Measuring brand experiences cross-nationally. *Journal of Brand Management*, 24, 86–104. doi:http://dx.doi.org/10.1057/s41262-016-0022-4

Spaargaren, G. (2011). Theories of practices: Agency, technology, and culture: Exploring the relevance of practice theories for the governance of sustainable consumption practices in the new world-order. *Global Environmental Change*, 21, 813–822. doi:http://dx.doi.org/10.1016/j.gloenvcha.2011.03.010

Steenkamp, J., **Batra, R.**, & **Alden, D. L.** (2003). How perceived brand globalness creates brand value. *Journal of International Business Studies*, 34, 53–65. doi:http://dx.doi.org/10.1057/palgrave.jibs.8400002

Steenkamp, J. B. E., & **Baumgartner, H.** (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research*, 25, 78–90. doi:http://dx.doi.org/10.1086/209528

Strizhakova, Y., **Coulter, R. A**, & **Price, L. L.** (2011). Branding in a global marketplace: The mediating effects of quality and self-identity brand signals. *International Journal of Research in Marketing*, 28, 342–351. doi:http://dx.doi.org/10.1016/j.ijresmar.2011.05.007

Strizhakova, Y., **Coulter, R. A.**, & **Price, L. L.** (2008). The meanings of branded products: A cross-national scale development and meaning assessment. *International Journal of Research in Marketing*, 25, 82–93. doi:http://dx.doi.org/10.1016/j.ijresmar.2008.01.001

From Replication to Further Scale Development: A Case Study of Incorporating an Eco-Friendliness Dimension in an Existing Consumer Experience Scale

Tomarken, A., & **Waller, N.** (2005). Structural equation modeling: Strengths, limitations, and misconceptions. *Annual Review of Clinical Psychology*, 1, 31–65. doi:http://dx.doi.org/10.1146/annurev.clinpsy.1.102803.144239

Uncles, M., & **Kwok, S.** (2013). Designing research with in-built differentiated replication. *Journal of Business Research*, 66, 1398–1405. doi:http://dx.doi.org/10.1016/j.jbusres.2012.05.005

United Nations. (2018). *Sustainable development knowledge platform—GOAL 12: Ensure sustainable consumption and production patterns*. Retrieved from https://www.un.org/sustainabledevelopment/sustainable-consumption-production/

Zhou, K. Z., **Su, C. T.**, & **Bao, Y. Q.** (2002). A paradox of price-quality and market efficiency: A comparative study of the US and China markets. *International Journal of Research in Marketing*, 19, 349–365. doi:http://dx.doi.org/10.1016/S0167-8116(02)00096-4

Page 19 of 19

From Replication to Further Scale Development: A Case Study of Incorporating an Eco-Friendliness Dimension in an Existing Consumer Experience Scale