Examination of the Effect of Congruity in Marketing

by

Qian Deng

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Marketing

Faculty of Business

University of Alberta

ABSTRACT

Congruity is a critical research construct in marketing, whose application is ubiquitous, including brand extensions, product designs, advertising, celebrity endorsements, brand alliances, event sponsorships, and others. It refers to the similarity, fit, or consistency between two objects (brands, persons, events, organizations, etc.). In the current marketing literature, however, many key questions on this topic are still left to answer. The three essays of my dissertation aim to address three different issues related to the effect of congruity.

First, when companies design their brand extension strategies, they already realize the importance of congruity (i.e., choosing a new product category fitting well with the parent brand). However, the current marketing research on this topic shows contradictory predictions and results, in terms of the magnitude of the congruity effect together with other relevant factors to determine the success of brand extensions. Therefore, in order to address this research issue, the first essay of my dissertation provides a conceptual and meta-analytical review of the congruity literature for brand extensions. First, the congruity hypothesis effect pattern (consumers respond better to congruity than to moderate incongruity, which they respond better to than to extreme incongruity) is found to be the dominant effect pattern of congruity for brand extensions, and the overall mean effect size of congruity between a brand extension category and the parent brand is positive and of small to medium size. Second, the main theoretical mechanism proposed in this literature is shown to focus on categorization and affect-transfer theories. Third, various theoretical moderators and methodological moderators of the congruity effect are also found to be significant. Overall, congruity is significant and very relevant from a managerial perspective, but it is far from being the sole determinant of success for brand extensions.

Second, although a sizable literature on brand extensions argues that the congruity between a parent brand and an extension product has a positive effect on consumer reception of the extension, the application of this literature is limited because of a lack of understanding of what "brand-extension fit" really is. The second essay of my dissertation develops a measurement scale of Brand Extension Fit (BEF) consisting of two core dimensions, *engineering-based* and *market-based congruity*, each measured by three items. The proposed scale represents a synthesis and extension of past work on congruity measurement, and is further validated with two separate datasets. Unlike uni-dimensional measures of fit – such as similarity, fit, consistency – used in most of the brand extension literature, the proposed Brand Extension Fit scale provides guidance for opportunity identification, idea generation, understanding the pros and cons of various alternatives, and the building of a marketing plan around a chosen alterative.

Third, the last essay of my dissertation focuses on the causal link between extreme incongruity and weirdness – two under-studied but very important topics in marketing. This essay, built on the extreme incongruity literature in product designs, investigates the research question: what makes products weird? This essay, via a series of experiments, presents convergent evidence that extreme incongruity between a product design and its own product category schema is a key antecedent of weirdness, because of a failed sense-making process. Furthermore, facilitating information for sense making can significantly decrease the weirdness perception. Moreover, although extremely incongruent products are weird and consumers like less and are less willing to buy them, those same consumers are more willing to share information about these products than about regular ones.

PREFACE

This thesis is an original work by Qian (Claire) Deng.

The research project, of which the essay 2 of this thesis is a part, received research ethics approval from the University of Alberta Research Ethics Board, Project Name "Scale development of the congruity construct: direct and indirect measures," No. Pro00042337, September 12, 2013.

The research project, of which the essay 3 of this thesis is a part, received research ethics approval from the University of Alberta Research Ethics Board, Project Name "The Concept of Weirdness," No. Pro00068648, November 2, 2016.

My advisor Professor Paul Messinger supervised the study design and the data collection.

ACKNOWLEDGEMENTS

My six-year Ph.D. study in Alberta School of Business, at the University of Alberta is coming to the end. These six years not only equip me with many skills as a researcher and teacher, but also enrich my life experience in order to become a better person. I'd like to express my sincere gratitude to the following persons and institutions, who have generously offered their help to me during these this fulfilling and rewarding journey.

I owe my deepest gratitude to my supervisor, Dr. Paul Messinger and my supervisory committee members, Dr. John Pracejus, and Dr. Sarah Moore. I feel so lucky and blessed to have guidance and support from these excellent scholars. They helped me develop my preliminary research ideas into meaningful and operational research frameworks, overcome various obstacles I have met in this journey, and offered valuable comments and directions when needed. I also would like to express my gratitude to the members of my dissertation committee: Dr. Kyle Murray, and Dr. Theodore Noseworthy. Their invaluable comments and suggestions on my proposal help me avoid many mistakes and pave a very solid foundation for the following work. I am indebted to the Alberta School of Business for generous and abundant research support and resources. My special thanks also go to Dr. David Deephouse, Debbie Giesbrecht, Jeanette Gosine, Helen Wu in the Ph.D. office for their immediate and patient assistance.

I am grateful to have joined and grown with the Ph.D. student family at the University of Alberta; this group of people has made my life in Canada warm and colorful. I want to thank my dear fellow research students in our school, particularly Muyu (Sarah) Wei, Virginia (Gini)

Weber, Ding (Allen) Tian, Ke (Kristy) Tu, Soyoung Kim, Shuo (Shirley) Chen, and Katie Lafreniere. Friendship is the golden ribbon that ties us together to become this lovely big family.

Finally, I want to devote special thanks to my husband, Lun Li, for his continued and unfailing love, support and understanding during my pursuit of Ph.D. degree. Without you, finishing this Ph.D. journey would have felt impossible. There were many times I found it so hard to continue and wanted to give up, but you were always here for me, and helped me go through these difficult times and keep things in perspective. I feel that I am the luckiest person in the world to have such a lovely and caring husband, standing beside me with his love and unconditional support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Essay 1

Is Congruity Desirable for Brand Extensions?

A Conceptual and Meta-analytic Review

Abstract

This paper provides a conceptual and meta-analytical review of the congruity literature for brand extensions. We find that the overall mean effect size of congruity between a brand extension product and its parent brand on consumers' responses toward brand extensions (mostly attitudinal responses) is positive and significant ($\bar{R}$=0.211), but the variance is high, and, managerially, the effect magnitude is small to medium. Our literature review shows that the main theoretical mechanism proposed in this literature for the congruity effect focuses on categorization and affect-transfer theories. We also review and examine moderators of the congruity effect that include sets of factors that (a) increase elaboration of moderately incongruent extensions, (b) encourage flexibility of categorizing incongruent extensions, (c) nullify the congruity effect, (d) are associated with parent brand characteristics, as well as (e) methodological influences. In particular, moderately incongruent extensions can generate more favorable outcomes than congruent ones in the presence of factors related to (a) and (b). Overall, congruity is significant and very relevant from a managerial perspective, but it is far from being the sole determinant of success for brand extensions.

Keywords: Congruity/Moderate Incongruity; Brand Extensions; Meta-analysis; Categorization; Affect Transfer; Fit.

Introduction

A brand extension is a marketing strategy, in which a firm marketing a product with a well-developed image uses the same brand name in a different product category. Brand extensions are a widely-used means of reducing the risk and expense of new product introductions (McCarthy, Heath, & Milberg, 2001). To reduce the risk of indifferent consumer reception, a brand extension builds on positive consumer associations with the established brand. To reduce expense, there is also benefit from leveraging existing advertising and distribution channels. The costs of this process can be substantial – in excess of $150 million for fast-moving consumer goods (Tauber, 1981) and an order of magnitude higher for complex technology product or services. For these reasons, 81 percent of new products are introduced as brand extensions, according to a leading text in the 1990s (Keller, 1998); and, more recently in India, a Nielsen study found that extensions are five times more successful than new brands for fast moving consumer goods (Kaul & Naire, 2012). Indeed, success can be the lifeblood of a company seeking to keep a brand relevant.

A key question for practitioners is how to maximize the success of brand extensions, and an important part of this question involves choosing a product category for a brand to extend to. Since a groundbreaking paper by Aaker and Keller (1990), much research has focused on choosing an extension product category that "fits" with either its parent category or its parent brand. In this literature, "fit", "similarity", and "congruity" have been used as equivalent terms, depending on the authors. "Fit" has a managerial slant; "similarity" is a simple elocution; and "congruity" ties in with some key concepts in psychology.

In the subsequent literature on brand extensions, a significant positive effect of congruity has been persistently reaffirmed in many individual studies and in prominent literature

summaries (Bottomley & Holden, 2001; Volckner & Sattler, 2006). In particular, after analyzing a large data set, the latter authors find that the degree of 'fit' between a parent brand and an extension product category is the most important of ten 'success factors' for brand extensions, including, among others, marketing support, parent-brand conviction, retail acceptance, and parent-brand experience. Interpreting the importance of congruity in the many studies about the determinants of brand extension success, Keller and Lehmann (2006) suggest "consumers need to see the proposed extension as making sense" (p.748).

Nevertheless, differing views about the importance of congruity relative to other factors persist in the sizable literature that has grown since. And core questions remain to be answered.

First, there is a lack of agreement about the magnitude of the congruity effect. Some papers find small effects, some find large effects, and some even find null effects. This motivates one to ask: *How large is the congruity effect for brand extensions? And how much variation is there?* To answer these questions, it can be helpful to examine the distribution of effect sizes of congruity across the brand extension literature.

Second, many other factors have been found to influence brand extensions since 1990, including various contextual cues, individual level factors, and brand characteristics. But there is no consensus (or compendium) of what other factors have been documented to influence brand extensions together with congruity. This leads to the question: *What are the relevant factors, in addition to congruity, that influence the success of brand extensions? And how do they interact with congruity?* In the current literature, there is no recent paper that addresses these issues.

Third, a stream of work, grounded on schema congruity/incongruity theory (Mandler, 1982), suggests that brands should extend into moderately incongruent categories, because a moderate level of incongruity requires more elaboration, cognitive effort, and processing to

resolve inconsistencies; and this process of resolution can lead to higher intensity of positive affect response. This motivates one to ask: *Can moderate incongruity generate more favorable responses to brand extensions from consumers than congruity? And under what circumstances?*

In order to address these three sets of research questions, we provide a comprehensive theoretical review and a meta-analysis of the congruity literature for brand extensions.

Our theoretical review summarizes the work on congruity since 1990, as well as various factors studied that work together with congruity to influence consumer reactions to brand extensions. We provide a conceptual framework that synthesizes this work, extending the conceptual framework of Czellar (2003). We include studies since that publication and we rework that framework to put a focus on the theoretical mechanism of categorization theory and affect transfer, which has frequently been argued to underlie the congruity effect. This theoretical review provides necessary background for our empirical analysis.

Our empirical analysis consists of a meta-analysis to measure the impact of congruity (between a parent brand and an extension product category) on consumers' reactions in the brand extension research. This method is particularly useful here, because it can directly contrast and combine different study results via effect sizes, so that patterns, relationships, and even source of variation among studies' results can be identified.

By providing both the overall mean effect size and the variance across studies, the meta-analysis permits a direct answer to the first set of questions question above. By providing a compendium of factors other than congruity that influence brand extensions and examining how groups of these factors moderate the effect of congruity, our conceptual review and meta-analysis addresses the second set of questions above. And by examining the moderate

incongruity hypothesis and when it holds, our meta-analysis permits examination of the third set of questions.

Overall, we find a highly significant positive effect of congruity across this literature, but the mean effect size is only of small to moderate size ($\bar{R}$=0.211), and the variance is large. Congruity with such an effect size would typically not be enough, by itself, to determine the success of a brand extension. But together with other factors (summarized in our answer to the second question), congruity can tip the balance. Lastly, firms should not exclusively feel confined to extending only to the most similar product categories. In answer to our third question, consumers can accept moderately incongruous extensions, if they can flexibly categorize or engage in elaboration to make sense of the extension product. Overall, the work that we review in what follows is hoped to provide a resource for managers and researchers applying or extending this important research stream.

Role of Congruity in the Literature on Brand Extensions

We begin by summarizing, in table 1.1, the key experimental research that considers congruity or fit as a determinant of brand extension success, supporting theories relied on, and relevant moderators examined.

This stream of research goes back to a paper by Aaker and Keller (1990) who consider two early congruity hypotheses ("H_3: The fit between the two involved product classes has a direct positive association with the attitude toward the extension", and "H_2: The transfer of a brand's perceived quality is enhanced when the two product classes in some way fit together. When the fit is weak, the transfer is inhibited." p. 30). This paper acknowledged, added to, and empirically examined various ideas considered in an early workshop on the topic (University of

Minnesota Consumer Behavior Seminar in 1987). It is worth noting explicitly that this paper focused on the fit between the *parent category* and the extension product category.

Table 1. 1. Literature on congruity effect for brand extensions.

Research Paper	Supporting Theories	Moderators of the Congruity Effect								
		k	*n*	Brand-related	*k*	*n*	Informational cues	*k*	*n*	Individual level factors
1. Aaker and Keller (1990). JM	C, AT, ANWM, O	c	c	Congruity Hypothesis						
2. Boush and Loken (1991). JMR	C, AT, SAMM	1	6	Broad vs. narrow						
3. Park, Milberg, and Lawson (1991). JCR	C, AT	1	4	Prestigious vs. functional						
4. Keller and Aaker (1992). JMR	C, AT	1	6	Parent brand quality						
5. Broniarczyk and Alba (1994). JMR	C, AT	2	6	Brand associations						
6. Morrin (1999). JMR	C, AT	1	2	Dominant or not						
7. Barone, Miniard, and Romeo (2000). JCR	C, AT				1	12	Similarity prime	4	48	Mood
8. Lane (2000). JM	M				1	4	Repeated ad exposure			
9. McCarthy, Heath, and Milberg (2001). ML	C, AT				1	3	New brand name			
10. Maoz and Tybout (2002). JCP	M							2	10	Involvement
11. Zhang and Sood (2002). JCR	C, AT				1	2	Similarity prime	2	6	Age
12. Martin, Stewart, and Matta (2005). JAMS	C				1	4	Goal congruity			
13. Barone (2005). JCP	C, AT							1	12	Involvement & mood
14 Yeung and Wyer (2005). JMR	C, AT	1	4	Brand affect	1	4	Similarity prime	1	4	Mood
15 Kalamas, Cleveland, Laroche, and Laufer (2006). JSM	C, AT, M	1	6	Congruity Effect						
16. Yeo and Park (2006). JCP	C, AT, M							3	6	Regulatory focus
17. Nan (2006). P&M	C, AT	1	12	Brand attitude						
18. Monga and John (2007). JCR	C, AT							1	8	Country & analytical vs. holistic thinking
19. Shine, Park, and Wyer (2007). JMR	S	1	1	Multiple brand extensions						
20. Kim and John (2008). JCP	C, AT				1	1	Fit prime	3	6	Construal level
21. Fedorikhin, Park, and Thomson (2008). JCP	C, AT	2	6	Emotional attachment to brand						
22. Ahluwalia (2008). JMR	C, M				1	6	Independence vs. interdependence prime (self-construal)			
23. Hagtvedt and Patrick (2008). JCP	C, AT				1	1	Art present			
24. Oakley, Duhachek, Balachander, and Sriram (2008). JCR	C, AT	1	2	Order of entry into extension category						
25. Buil, de Chernatony, and Hem (2009). EJM	C, AT	1	12	Brand equity				1	12	Country
26. Kapoor and Heslop (2009). IJRM	C, AT	2	9	Extension strength	1	8	Competitive cue			
27. Milberg, Sinn, and Goodstein (2010). JCR	C, AT				1	6	Competitive context			
28. Gierl and Huettl (2011). IJRM	C, AT, M	2	21	Brand attitude	4	10	Similarity prime (1, 2); Brand portrayal (1, 2); Brand slogan (1, 2); Peripheral design cue (1, 4)			

Table 1.1. Literature on congruity effect for brand extensions. (continued)

Research Paper	Supporting Theories	Moderators of the Congruity Effect								
		k	*n*	Brand-related	*k*	*n*	Informational cues	*k*	*n*	Individual level factors
29. Yorkston, Nunes, and Matta (2010). JM	C, AT							1	4	Incremental vs. entity orientation
30. Bambauer-Sachse, Hüttl, and Gierl (2011). P&M	C, AT	1	4	Brand attitude	1	4	Fit prime			
31. Shen, Bei, and Chu (2011). P&M	C, AT				1	6	Case reminder			
32. Liu and Hu (2012). P&M	C, AT	1	4	Prestigious vs. functional	1	4	Private vs. public-consumption occasion			
33. Sood and Keller (2012). JMR	C				3	10	Family brand vs. sub-brand			
34. Mathur, Jain, and Maheswaran (2012). JCP	C, AT							3	6	Incremental vs. entity orientation
35. Milberg, Goodstein, Sinn, Cuneo, and Epstein (2013). JMM	C, AT	1	2	Parent brand quality						
36. Noseworthy, Muro, and Murray (2014). JCR	M							1	6	Arousal
37. Huang, Jia, and Wyer (2017). P&M	C, AT				1	6	Physical distance			
38. Dimitriu, Warlop, & Samuelsen (2017). EJM	C, AT							1	6	Purchase goal: attribute vs. overall performance
Total: k =37 , n = 269	C: 34; AT: 31; M:6; ANMM: 1; SAMM: 1; S: 1; O: 1.	21	107		21	85		25	140	

Note. ***k*** is number of studies; ***n*** is number of effect sizes present; c: correlational and/or descriptive; C: categorization theories (Cohen & Basu, 1987; Fiske, 1982; Fiske & Pavelchak, 1986; Meyers-Levy& Tybout, 1989; Sujan, 1985; Tversky, 1977); AT: affect transfer theory (Boush et al,. 1987; Wright, 1975); ANMM: associative network memory model (Anderson, 1983); SAMM: spreading activation model of memory (Collins & Loftus, 1975); M: schema congruity/incongruity theory (Mandler, 1982); S: brand synergy effect (Shine, Park, & Wyer, 2007); O: other related theories, including cognitive consistency (Heider, 1958; Osgood & Tannenbaum, 1955), stimulus generalization (Bierley, McSweeney, & Vannieuwkerk, 1985; McSweeney& Bierley 1984).

After a study by Broniarczyk and Alba (1994), however, the literature evolved to focus mostly on the fit between *the parent brand* and the extension product category. These authors argued that fit with specific parent brand associations may dominate the effect of fit with the parent category. "For example, Apple computer is associated with user friendliness, but this association is not strongly associated with other computer brands or with the product class as a whole" (p. 215). So Apple in 1994, as a user-friendly brand, was more congruent with (and hence more suited to extend to) some consumer categories than other computer brands were.

As table 1.1 indicates, several other papers from the 1990s considered various general features of brands: (a) does the brand carry a narrow or broad range of product categories; (b) is the brand prestigious or functional; (c) is the brand perceived as high or low quality; (d) does the

brand elicit positive affect, attitudes, or associations; and (e) is the brand dominant in its own category. These brand features, together with congruity, were found to influence consumers' reception to proposed brand extensions.

From 2000 to 2005, the papers branched off to examine moderators of the congruity effect, pertaining to (1) information cues in the selling environment, and (2) individual level factors. The information cues in these papers include (a) primes for consumers to notice the similarity of the extension category with the parent brand category; (b) advertisement exposure time, (c) the amount of product information provided, (d) whether the extension product uses the parent brand name or a new name; and (e) whether there is goal congruity between the extension products and the parent brand categories. The individual level factors in these papers include (a) mood, (b) involvement, and (c) age. These information cues and individual level factors were found to influence the effect of congruity on consumers' reception to proposed brand extensions.

From 2006 to 2017, the literature elaborated on new brand characteristics (brands that elicit an emotional attachment, parent brand equity, and strength of the extension product), new information cues (art, references to competitors, reminding consumers of similar brands in the extension category, whether the consumption occasion was public or private, whether the extension uses the parent/family brand name or a new name, and physical distance between consumers and the brand extension product), and new individual level factors (regulatory focus, country, abstract or concrete construal orientation, analytical or holistic thinking, and incremental or entity orientation).

During these same periods, correlational research confirmed the importance of congruity (well-cited examples include Smith & Park, 1992; Sunde & Brodie, 1993; and Völckner & Sattler, 2006), using different kinds of data (e.g., survey and panel data for consumers and

managers), and different dependent variables (e.g., brand quality, brand position, market share, advertising efficiency, and the likelihood of trying the extension). These correlational studies (see the detailed appendix 1.B) provide evidence, in addition to the experimental research of table 1.1, that congruity leads to better outcomes than incongruity.

We use the term *Congruity Hypothesis* to describe the predicted effect pattern where congruity leads to better consumer-related outcomes than moderate incongruity and extreme incongruity. We describe this effect pattern with the shorthand: Congruity > Moderate Incongruity > Incongruity, where ">" means that consumers' responses toward the former condition are more favorable than the latter conditions.

Underlying Mechanism of the Congruity Hypothesis

The consideration of congruity or fit in the marketing literature, generally, has its antecedents in social psychology, going back to Gestalt psychology, which suggests that people prefer to perceive the environment in simple and coherent ways (Kohler, 1929). Various cognitive consistency theories have arisen from this perspective, including the strain toward symmetry (Newcomb, 1953), congruency theory (Osgood & Tannen-baum, 1955), and the affective-cognitive consistency model (Rosenberg, 1956). Particularly relevant is that the presence of congruity can be associated with (a) mental comfort (Festinger, 1957), (b) a balance state (Heider, 1958), (c) ease of processing or categorizing different objects (Fiske & Pavelchak, 1986), and (d) affect or image transference via an established memory link (Anderson, 1983; Shimp, 1981; and, in marketing, Keller, 1993; Gierl & Huettl, 2011; Smith, 2004).

Although the above ideas are frequently mentioned as antecedents, the work summarized in table 1.1 has been remarkably consistent about particularly relying on categorization theories

and affect transfer theories as a justification for their hypotheses about congruity as an important determent of positive consumer responses to brand extensions.

Categorization theories

Essential to understanding the effect of congruity on consumers is the *categorization process* (reviewed by Loken, Barsalou, and Joiner in 2008), where "consumers use categorical representations," which refer to information stored in the cognitive system for a consumer category, "to assign a particular product or service to a consumer category, so that they can understand and draw inferences about it" (p.133). Category inferences can happen for brand extensions, if there is a high level of congruity between a parent brand's representations and an extension category's representations. The majority of the papers (34 out of 37 papers) in table 1.1 specifically cites and builds on categorization theories (e.g., Cohen & Basu, 1987; Fiske, 1982; Fiske & Pavelchak, 1986; Sujan, 1985; Tversky, 1977).

Affect transfer theory

As the result of the successful categorization process, inferences are created whereby affect toward one object (the reference object concerning which consumers already have set attitudes) will be transferred to the other object (the target object) (Shimp, 1981). Here again, the bulk of the papers (31 out of 37 papers) in table 1.1 specifically rely on affect transfer theory (Boush et al., 1987; Wright, 1975). Two key characteristics of affect transfer should be noted. First, both positive and negative affect toward the reference object (e.g., a parent brand) can be transferred to the target object (e.g., an extension product). Second, affect transfer is bi-directional. Although positive affect from a parent brand can be transferred to its extension product, negative affect toward an extension can also be transferred to its parent brand, which is called reverse negative affect transfer and is recognized in managerial (e.g., Diamantopoulos,

Smith, & Grime, 2005) and academic literature (e.g., Loken & John, 1993; Mathur, et al., 2012; Martinez & Pina, 2003; Milberg, Park, & McCarthy, 1997; Sood & Keller, 2012).

CAT synthesis

The heavy reliance on categorization and affect transfer theories in the brand extension literature motivates us to formulate a synthesis, a two-staged process, which we refer to as CAT (categorization and affect transfer). The congruity hypothesis built on the CAT process requires (1) successful categorization of the extension category (i.e., target), facilitated through perceived congruity, as being closely related to the parent brand (i.e., reference), (2) positive affect or attitude toward the parent brand, and (3) some transfer of affect or attitude from the parent brand to the extension product. This synthesis formalizes what was anticipated in the early work of Aaker and Keller (1990, particularly their H_2, mentioned earlier). In summary, CAT theory is the primary argument used to support the congruity hypothesis. Our meta-analysis examines the extent to which this hypothesis holds in this body of literature, including the distribution of effect sizes.

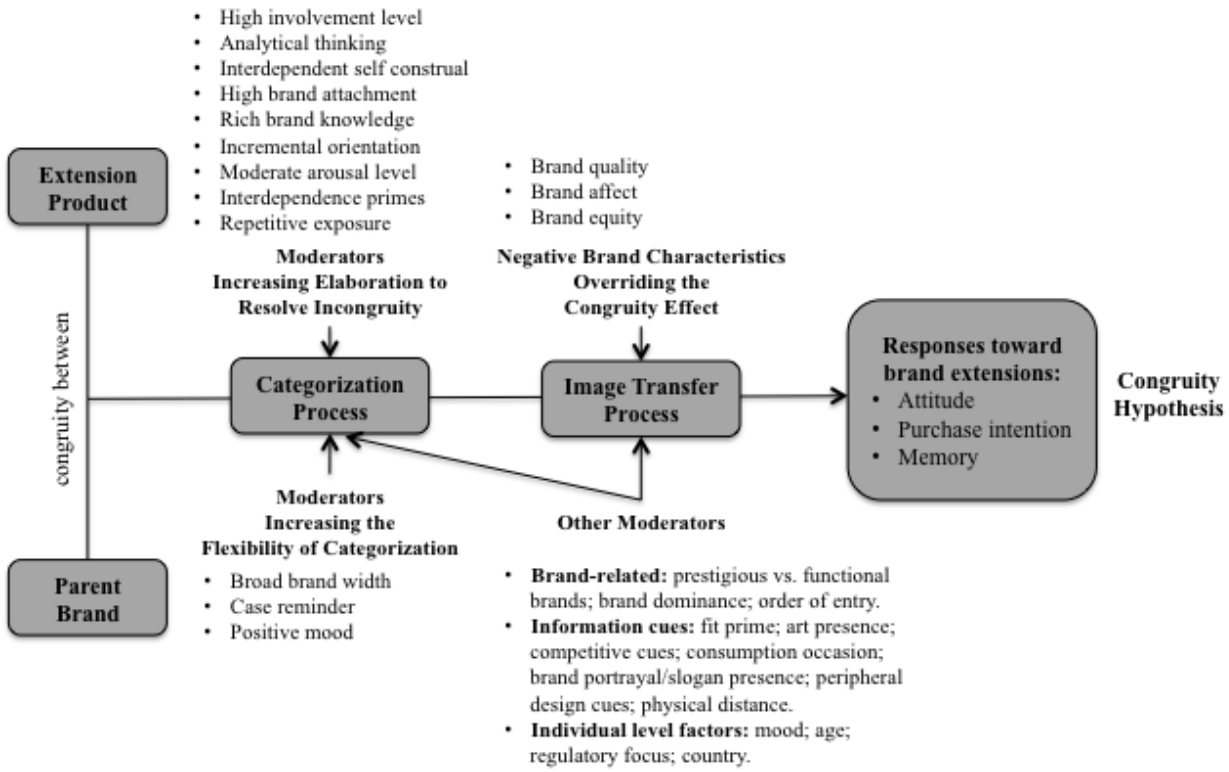
Moderators of the Congruity Effect

In order to identify the key relevant factors that interact with congruity to influence the success of brand extensions (our second research question), and to investigate when moderate incongruity can generate more favorable responses than congruity (our third research question), this section reviews and synthesizes various studied moderators of congruity in the brand extension literature, specifically considering moderators influencing the underlying mechanism of the congruity hypothesis (CAT mechanism) and different types of dependent variables.

Moderators influencing the CAT mechanism

We accordingly reviewed the conceptual arguments in the papers surrounding the moderators summarized in table 1.1 and created the framework summarized in figure 1.1. This framework builds on the previous conceptual framework of Czellar (2003), and goes beyond it by adding the underlying mechanism of categorization and affect transfer, and grouping moderators around the CAT mechanism. From a conceptual standpoint, we categorize the moderators influencing the CAT mechanism into three major groups: (1) moderators influencing the categorization process; (2) moderators influencing the affect transfer process, and (3) other moderators (see table 1.1 and figure 1.1, not reviewed in detail in this section).

Figure 1. 1. Moderators of the congruity hypotheses.



Moderators influencing the categorization process. The categorization process is usually the first step that consumers engage in, when they are evaluating any brand extension. Our literature review finds that there are two broad ways the categorization process is influenced.

First, a number of factors can influence the ***resolution of incongruity***, which is an important part of the categorization process. Mandler's (1982) schema congruity/incongruity theory describes this part of the process very specifically. According to Mandler (1982), moderate levels of incongruity can be resolved via assimilation or accommodation to lead to positive consumer reaction with greater affective intensity than congruity. But when incongruity is severe, deeper structural changes for the existing schema are needed to accommodate the new information, and outcomes become uncertain, often negative. The predicted effect pattern associated with this hypothesis is that a moderately incongruent extension can lead to higher intensity of positive affect responses than a congruent one, but an extremely incongruent extension usually leads to negative affect responses (we name this effect pattern as the *Moderate Incongruity Effect Pattern*, and describe it with the shorthand: Moderate Incongruity > Congruity> Incongruity). Application of this hypothesis started in marketing in the new product design literature (Campbell & Goodstein, 2001; Myerslevy & Tybout, 1989; Noseworthy & Trudel, 2011; Peracchio & Tybout, 1996) – looking at whether a new product was congruent with consumers' perceived schema for the product type, and later was introduced into the brand extension area.

One key determinant of the resolution of moderate incongruity is elaboration. Such elaboration can be inherently triggered (e.g., in the presence of analytical thinking, interdependent construal level, brand attachment, brand knowledge, or incremental orientation). Elaboration can also be externally induced (e.g., with the arousal level, involvement level, interdependence primes, or repetitive exposure). In either case, moderate incongruity is more easily resolved than extreme incongruity, and the resolution of moderate incongruity can help generate positive affective responses with higher intensity.

In the brand extension literature, elaboration draws on various theoretical mechanisms, involving more devoted cognitive effort/resources (Lane, 2000; Mandler, 1982; Maoz & Tybout, 2002), central route of processing (Petty & Cacioppo, 1986), relational processing (Ahluwalia, 2008), analytical processing (Monga & John, 2007), and process-focused processing (Mathur, et al., 2012). According to all of the accounts, the moderate congruity effect pattern is more likely to emerge than the congruity hypothesis effect pattern. We summarize these ideas as follows:

H1: In the presence of moderators that increase elaboration to resolve incongruity, the moderate incongruity effect pattern will emerge.

Second, many factors can change the ***flexibility of the categorization process***. The basic argument is that these factors promote flexibility in categorization and increase the chance of categorizing the moderately incongruent extension product into the parent brand category. As a result, a moderately incongruent extension product will be viewed as more congruent with the parent brand, which further facilitates affect transfer. In other words, a moderately incongruent brand extension may receive an equally or even more favorable outcome than a congruent brand extension. Those factors that facilitate the flexibility of the categorization process and moderate the congruity effect to the moderate incongruity effect pattern include characteristics of the parent brand (e.g., brand breadth by Boush & Loken, 1991), individual level factors (e.g., mood by Barone, et al., 2000), and information/contextual cues (e.g., case reminders by Shen, et al., 2011). We summarize this idea as follows:

H2: In the presence of moderators that increase the flexibility of categorization, the moderate incongruity effect pattern will emerge.

Moderators influencing the affect transfer process. According to the CAT synthesis, after successful categorization, the positive affect toward the parent brand will be transferred to

the extension product. If the parent brand does not elicit positive affect, then, even if categorization is successful, there will be no positive affect transfer to the extension product. This argument illustrates that characteristics of the parent brand play a key role in influencing the affect transfer process.

Characteristics of the parent brand. In the brand extension literature, many researchers (Bambauer-Sachse, et al., 2011; Gierl & Huettle, 2011) find that congruity leads to more favorable evaluations than incongruity, only when consumers have a favorable initial attitude toward the parent brand; but the level of congruity does not matter when attitude toward the parent brand is neutral or negative. Other factors studied in the literature leading to analogous effects on affect transfer include brand quality (Aaker & Keller, 1990; Keller & Aaker, 1992; Milberg, et al., 2013), brand affect (Yeung & Wyer, 2005), and brand equity (Buil, et al., 2009). Interestingly, a negative image of the parent brand overrides or dominates the positive effect of congruity. This could be viewed as a negativity effect/bias, which means "when of equal intensity, things of a more negative nature have a greater effect on one's psychological state and processes than do neutral or positive things" (Baumeister, et al., 2001; Kanouse & Hanson, 1972). Furthermore, perhaps, congruity makes transfer possible, but what is transferred depends on the affect toward the reference object relative to the target object. We summarize these ideas as follows:

H3: In the presence of negative information associated with the parent brand, the congruity effect will be attenuated.

Considerations related to the dependent variable

Lastly, the experimental studies in table 1.1 mostly focus on three types of dependent variables: attitudes, behavioral intentions, and memory-related outcomes. Attitudes and

behavioral intentions are common dependent variables for the congruity effect in many social psychology studies. Attitudes and behavioral intentions are particularly relevant for brand extension studies that rely on CAT theory, because this body of papers relates to affect transfer toward an extension product, and consumer attitudes and purchase intentions are natural, easily measured, outcomes of affect transfer. Therefore, we predict:

H4: The congruity effect pattern will emerge for attitudinal and behavioral intention dependent variables.

We note, however, that the underlying mechanism of the congruity effect on memory-related variables (e.g., assisted and unassisted memory recall, and reaction time) is very different from the CAT process. We can see how memory-related variables depart from the Congruity Hypothesis as follows. According to the associative network memory model (Anderson, 1983) and Hastie's memory model (1980), congruent stimuli may help establish memory links, but as soon as the link has been established, information about a stimulus will not stay in the working memory. On the other hand, moderately or extremely incongruent stimuli are difficult to comprehend and will stay in people's working memory longer. In the meantime, an inter-episode associative link between the two stimuli will be established. When previously stored information is retrieved and interacts with new information in the working memory, more inter-episode links can be generated. This increases the probability of being recalled or recognized later on. According to this argument, the more incongruent the stimulus is, the more likely the stimulus will be recalled or recognized. Consistent with this argument, some recent psychology research (e.g., Na & Kitayama, 2011) finds that schema incongruent stimuli evoke stronger brain activities than congruent stimuli. For these reasons, a different predicted effect pattern of the congruity effect may be relevant for memory-related measures. We name this alternative

prediction as the *Incongruity Effect Pattern*, according to which consumers can respond more favorably to extreme incongruent extensions than to moderately incongruent or congruent ones, and more favorably to moderately incongruent extensions than to congruent ones for memory-related outcome variables (we describe this pattern with the shorthand: Incongruity > Moderate Incongruity > Congruity). Therefore, we predict:

H5: The incongruity effect pattern will emerge for memory-related dependent variables.

To summarize our discussion of figure 1.1, we find it meaningful to group studies exploring the boundaries of the Congruity Hypothesis according to whether the papers study moderators that (a) increase elaboration to resolve incongruity, (b) increase the flexibility of categorization, or (c) override the congruity effect. A further boundary applies when (d) the dependent variable is a memory-related variable. Earlier, in our discussion of table 1.1, we grouped the papers according to those that study (i) brand characteristics, (ii) information cues, or (iii) individual level factors. The grouping (i) - (iii) situates moderating factors either with the brand, itself, the environment, or the consumer. The grouping (a) – (d) identifies factors that influence each stage of the CAT process and how. These are entirely different groupings, but together they provide a useful conceptual overview: *The body of research about brand extensions identifies moderating factors about the brand, the environment, and the consumer that variously affect the categorization stage, the affect transfer stage, or the outcome stage of figure 1.1.*

The subtext of this discussion is simply that there has been much work studying congruity for brand extensions from different perspectives and with different conclusions. Given this work, and the associated divergent conclusions about the Congruity Hypothesis, it is time to take stock.

Meta-Analysis Methodology

We use meta-analysis to consider three things: (1) how large is the effect size of congruity for brand extensions and how much variation is there across the studies; (2) how do other relevant variables interact with congruity to affect consumer response to brand extensions; and (3) when does the moderate incongruity hypothesis hold. Our methodology includes (1) literature search, (2) data screening, (3) inclusion criteria, (4) coding procedure, and (5) meta-analysis methods.

Literature search

In order to identify relevant research, a keyword search (brand extension, congruity, incongruity, fit, congruence, match-up, similarity) in the abstract was conducted in two main sources: three large electronic databases and ten highly-ranked marketing journals (see appendix 1.A.1 for details of selection of the databases, journals, intentional redundancy, and in-process learning extension of the search). A total of 764 references were identified, including journal articles, theses, conference papers, cases, book chapters, and business reviews.

Data screening

We excluded unrelated studies arising from keywords with multiple meanings. All identified references were screened according to the content of the abstracts, whereby a paper was retained if it was judged by the authors to be theoretically relevant to the topic of congruity. Furthermore, a number of theses, book chapters, and conferences papers were redundant with journal papers, so in this meta-analysis, we decided to focus on papers published in academic journals (with peer review). As a result, 101 papers were identified. Among them, 32 are correlational research, 42 are experimental research, and the others are qualitative, literature reviews, or general conceptual or descriptive discussions.

Inclusion criteria

We decided to include only experimental studies in this meta-analysis for three reasons. First and most importantly, most past correlation research only estimates a linear relationship between congruity and consumers' responses, whereas experimental studies manipulate different levels of congruity, which makes it possible for us to distinguish moderate incongruity from congruity or extreme incongruity. This is essential for one of our key objectives: to identify when the moderate incongruity hypothesis (a nonlinear relationship between congruity and consumers' responses) effect happens. Second, a causal relationship between congruity and consumer response is most clearly observed in well-designed experiments. Third, comparable research designs, which are used in these experimental studies, facilitate meta-analysis. In particular, from the 101 identified research papers on the topic of congruity for brand extensions, we coded 42 experimental papers for further meta-analysis.

Coding procedure

Following standard practice in meta-analysis, in our coding process, we calculated and recorded as the measure of effect size, Cohen's *d*, which is obtained by dividing the difference in means for the two experimental conditions by the pooled standard deviation (Cooper, Hedges, & Jeffrey, 2009). A positive *d* score indicates that the more congruent condition results in a more favorable consumer response, and a negative *d* score means the less congruent manipulation is associated with a more favorable response.

In order to obtain the information necessary for the calculation of effect sizes, there are some complications, which need to be dealt with. (1) Different researchers defined their conditions differently. Therefore, in order to accurately test the three key research hypotheses discussed earlier, we used information from a study's pretests and manipulation checks to create

a consistent label for the different levels of congruity examined (see the appendix 1.A.2: Correction of Inconsistent Labeling Problems). (2) Some studies do not report all the required information to calculate the effect sizes. As a general principle, we excluded studies when there was a preponderance of missing information (5 experimental papers were excluded from the meta-analysis due to incomplete information for effect size calculations), and we used approximations when most of the needed information was present (see appendix 1.A.3). (3) For studies that report significant interaction effects between congruity and other variables (e.g., familiar vs. unfamiliar focal brand), the effect sizes were calculated separately for each moderator condition to retain the interaction effect (following Scheibehenne, Greifeneder, & Todd, 2010). (4) For studies using within-subject designs, the effect sizes were corrected following Cooper et al. (2009) (see appendix 1.A.4).

A number of variables that could moderate the effect sizes of congruity were further coded. (1) Some methodological factors, which might moderate the effect size of congruity, were coded, including the journal in which the paper was published, whether the effect size was estimated, the study design used (between- or within-subject), the population type (students or general consumers), the use of a pretest, and the success of the manipulation check. (2) Three key moderators, increased elaboration to resolve incongruity, increased flexibility of categorization, and negative information associated with the parent brand, on the CAT process are dummy coded as proposed in figure 1.1. For example, for negative information associated with the parent brand dummy variable, we coded 1 for all effect sizes of congruity associated with negative parent brand (e.g. negative brand image, association, quality, equity), and 0 for other effect sizes as control. (3) The type of dependent variables (attitude, behavioral intention, memory-related), and the use of real parent brands were also further coded (1if real and 0 if a

hypothetical brand was used). (4) We also conducted a survey (see appendix 1.A.5) to augment our dataset to include information about several brand characteristics of the real parent brands used in our meta-analysis dataset: brand familiarity, brand attitude, brand quality, brand breadth and the length of brand history (we recognize that use of *ex post* measurements introduces measurement error, so we separate our analysis of these variables from our main analysis).

Meta-analysis methods

The unit of analysis in this meta-analysis is an effect size, which reflects the difference of the dependent variable (i.e., consumer responses) between two experimental manipulations of the independent variable (i.e., congruity). Two related statistics were used to describe the effect size, Cohen's *d* and the correlation coefficient, *r*. Cohen's *d* was recorded in the coding process, and we used a standard transformation to calculate the correlation coefficient, *r*, from Cohen's *d* (Cooper et al., 2009, p.234). We report statistics about the effect size in terms of the correlation coefficient, *r*, because this measure has a fixed range (from -1 and 1), which is commonly interpreted in a similar fashion across applications (Cohen's guideline: 0.1: small; 0.3: medium; 0.5: Large; Cohen, 1992, p.157). The mean effect size in this meta-analysis is the weighted mean effect size, where each effect size is weighted by the pooled sample size (of the two experimental conditions used to calculate the effect size). Apart from reporting mean effect sizes, confidence intervals, credibility intervals (Higgins & Thompson, 2002), and *Q* statistics (Cochran, 1954) were also considered. All the computations related to the mean effect size were performed by using MetaExcel (Steel, 2006).

Results

Our meta-analysis quantifies that the mean effect size pattern supports the congruity hypothesis, generally, and, more subtly, shows how the conclusion is influenced by various moderators. We begin with a summary of the data.

Descriptive analyses

The data concerning the congruity effect in brand extensions came from 37 journal articles in 11 academic journals in fields related to marketing or business. In total, 269 data points, each of which represents an effect size arising from the comparison between two levels of congruity, were coded.

Table 1. 2. Descriptive analysis of the meta-analysis data.

Journal	*k*	*j*	Design	*k*	*j*
European Journal of Marketing	18	2	Between-subject	185	29
International Journal of Research in Marketing	30	2	Within-subject	84	8
Journal of Consumer Psychology	61	7	**Sample**	***k***	***j***
Journal of Consumer Research	58	7	Student	237	29
Journal of Marketing	8	2	General consumers	32	8
Journal of Marketing Management	2	1	**Manipulation check**	***k***	***j***
Journal of Marketing Research	47	8	Yes	167	24
Journal of Strategic Marketing	6	1	No or failed	102	13
Journal of the Academy of Marketing Science	4	1	**Pretest**	***k***	***j***
Marketing Letters	3	1	Yes	230	30
Psychology & Marketing	32	5	No	39	7
			SD Estimation	***k***	***j***
			Yes	93	15
			No	176	22
Total	**269**	**37**	**Total**	**269**	**37**

k is the number of effect size measures. *j* is the number of journal articles.

Table 1.2 shows descriptive characteristics of the data. 68.8% of the coded data originate from studies using the between-subject experiment design. The majority of the observations are from studies using student samples (88.1%), while only 11.9% of them are from studies sampling from general consumers. 62.1% of data come from studies using manipulation checks. In addition, 85.5% of data come from studies with pretests. Furthermore, about a third of the

effect sizes (34.6%) were calculated based on an estimation of the standard deviations (which were not reported explicitly in some papers) from other information included in the papers.

To help visualize key topics examined in the literature, figure 1.2 shows a word cloud for all keywords occurring in all the papers we coded. The font size of each keyword in the word cloud is proportional to its frequency in our database. The word cloud reveals some interesting patterns. First, in the brand extension literature, there are many equivalent terms used to refer to the congruity, such as fit, similarity, and congruence. Second, categorization theory appears to be the most cited theory for the examination of congruity. Third, attitude evaluation is the most common outcome measure examined, and consumers' memory-related outcomes are also in evidence. Fourth, there is a very wide range of moderators, such as contextual factors and individual differences.

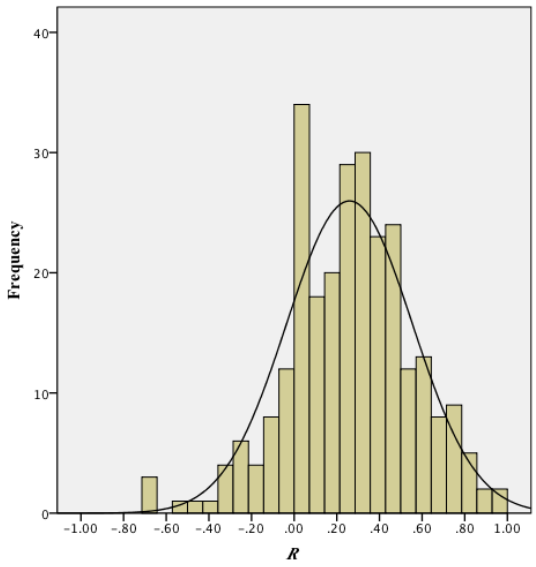
Figure 1. 2. Word cloud of keywords.

Mean effect size of congruity

The mean effect size for our whole sample of 269 data points helps us answer the question: overall, how big is the effect size of congruity for brand extensions?

Overall mean effect sizes. The frequency of the effect size (correlation coefficient scores, *r*) for our whole sample is shown in figure 1.3. This includes all treatments in our sample (e.g., pooling treatments labeled as comparing Congruity with Incongruity, Congruity with Moderate Incongruity, and Moderate Incongruity with Incongruity, and including all dependent variable types). We observe from figure 1.3 that the data display a basic normal distribution, and there are a large number of observations around zero, which shows that our dataset did not leave out the null effect sizes of congruity.

Figure 1. 3. Frequency distribution of effect sizes (*k*=269).



The weighted mean effect size ($\bar{R}$) is 0.211. The total variance of the weighted *r* scores (adjusted by *k/(k-1)* where *k* is the number of effect-size data points) is 0.067 with a 95% confidence interval of (0.18, 0.24), and the associated standard deviation is 0.259. The sampling error is 0.011, which accounts for 16.41% of total variance. The residual variance is 0.056, which accounts for 83.58% of the total variance. This indicates that there are still large amounts of variance unexplained and that there might be some moderator variables that can explain some portion of this variance. The credibility interval is (-0.26, 0.68) at the level of 95%, which

crosses the zero point, suggesting that the sign of the effect size of interest could be negative in some situations. The narrow confidence interval, wider credibility interval, and the frequency distribution (figure 1.3) together suggest the existence of moderators. Nevertheless, as a summary of all 269 effect sizes in the sample, we conclude that the more congruent the stimulus, the better the consumers' response, although the average effect size is not very large in magnitude. While perhaps not surprising, this quantifies a benchmark effect. We now refine the analysis by decomposing the effect sizes according to different comparison groups.

Table 1. 3. Overall mean effect sizes of congruity.

Comparison Groups	*R*	*k*	Variance	Residual variance	*Q* statisitcs	*p* value	Confidence interval*	Credibility interval*
C-IC	0.259	84	0.084	0.075	975.043	0.000	(0.20, 0.32)	(-0.28, 0.80)
C-MI	0.193	100	0.060	0.048	567.762	0.000	(0.15, 0.24)	(-0.23, 0.62)
MI-IC	0.174	85	0.052	0.041	453.062	0.000	(0.13, 0.22)	(-0.22, 0.57)
Total	0.211	269	0.067	0.056	2044.968	0.000	(0.18, 0.24)	(-0.26, 0.68)

Note. Congruity: C; Moderate Incongruity: MI; Incongruity: IC; *: 95%.

Overall mean effect sizes for three comparison groups. Table 1.3 shows the mean effect size and other relevant summary measures for each of these comparison groups three comparison groups (Congruity - Moderate Incongruity: C-MI, Congruity - Incongruity: C-IC and Moderate Incongruity - Incongruity: MI-IC). We observe that congruity leads to more favorable consumer responses than both moderate incongruity and incongruity ($\bar{R}_{\text{C-MI}}$=0.193 > 0, $\bar{R}_{\text{C-IC}}$=0.259 > 0), and moderate incongruity yields more favorable response than incongruity ($\bar{R}_{\text{MI-IC}}$ =0.174> 0), which supports the *Congruity Hypothesis* (Congruity > Moderate Incongruity > Incongruity). This finding is consistent with our predicted effect pattern of congruity for brand extensions.

Effect size patterns for different dependent variables. After decomposing the effect sizes of congruity according to different dependent variables, table 1.4 identifies the following results. (1) For the attitudinal dependent variable, the mean effect sizes support the *Congruity*

Hypothesis (Congruity > Moderate Incongruity > Incongruity). (2) For the behavioral intention variables, due to the small sample size, we cannot identify the effect pattern with statistical significance. However, the magnitudes of the mean effect sizes match a weak version of the *Moderate Incongruity Hypothesis* (Congruity ≈ Moderate Incongruity > Incongruity, where "≈" that consumers' responses are similar between the former and the latter conditions). (3) For the memory-related dependent variables, there are only four data points, and the data only represent the comparison between congruity and incongruity, so we lack statistical power to draw any conclusion here. To summarize, these results only provide partial evidence for H4, because the congruity hypothesis effect pattern is identified only for attitudinal dependent variables, and our meta-analysis failed to provide enough evidence to accept or reject the hypotheses regarding the behavioral intention (H4) or memory-related variables (H5).

Table 1. 4. Mean effect sizes of congruity for different dependent variables.

Dependent Variables	All data			Congruity-Incongruity			Congruity-Moderate Incongruity			Moderate Incongruity-Incongruity		
	R	***k***	***N***	***R***	***k***	***N***	***R***	***k***	***N***	***R***	***k***	***N***
Attitude	0.21	240	21,032	0.243	73	7,332	0.202	89	6,964	0.181	78	6,736
Behavioral intentions	0.222	25	1,914	0.265	7	658	0.096	11	620	0.099	7	636
Memory-related variables	0.178	4	208	0.178	4	208	-	-	-	-	-	-
Total	0.211	269	23,154	0.259	84	8,198	0.193	100	7,584	0.174	85	7,372

R is mean effect size. ***k*** is the number of effect size measures. ***N*** is the number of experimental participants.

Moderator analysis

To further explain the variation among the effect sizes and to test how various proposed factors influence the effect of congruity, a moderator analysis (see table 1.5) was performed. The dependent variable used in this analysis was the effect size of an experimental manipulation, measured by the Cohen's *d* score, because the *d* score has a range of -∞ to +∞, which matches the range of a normally distributed error term. We used weighted least squares regression (with the sample size of each effect size as the weight). This estimation method has been criticized by Chernev, Böckenholt, and Goodman (2015), because it ignores the correlation between data that

Table 1. 5. Analysis of moderators of the congruity effect.

Moderators		Model 1		Model 2		Model 3	
		Estimate	Sig	Estimate	Sig	Estimate	Sig
(Intercept)		1.618	***	1.386	***	0.945	
Comparison groups	C-MI	-0.309	**			-0.263	*
	MI-IC	-0.334	***			-0.279	*
DV Type	Behavioral Intention	-0.165					
	Memory-related	-0.239					
Increased Elaboration	Elaboration:C-MI			-0.4487	*		
	Elaboration:MI-IC			-0.081			
	Elaboration:C-IC			-0.313			
Increased Flexibility of categorization	Flexibility:C-MI			-0.413			
	Flexibility:MI-IC			0.143			
	Flexibility:C-IC			0.682	*		
Negative brand characteristics	Negative:C-MI			-0.523	*		
	Negative:MI-IC			-0.479	**		
	Negative:C-IC			-0.263	.		
Characteristics of Real Parent Brands	Real Brand	-0.150					
	Familiarity					0.154	*
	Attitude					0.229	*
	History					-0.234	.
Methodological Moderators	European Journal of Marketing	-0.214		-0.268		-0.201	
	International Journal of Research in Marketing	-0.062		0.016		-0.023	
	Journal of Consumer Psychology	-0.067		0.086		0.053	
	Journal of Consumer Research	0.361	*	0.391	*	0.238	
	Journal of Marketing	0.333		0.296		0.485	
	Journal of Marketing Management	-0.346		-0.071		-0.266	
	Journal of Strategic Marketing	0.562	*	0.426	.	0.829	*
	Journal of the Academy of Marketing Science	-0.441		-0.278		-0.165	
	Marketing Letters	-0.857	**	-0.620	.	-1.083	*
	Psychology & Marketing	0.010		0.084		0.216	
	Design: Between-subject	0.012		-0.117		0.148	
	Student Sample	-0.334	*	-0.295	*	-0.244	
	Manipulation check: yes	0.506	***	0.360	**	0.323	*
	Pretest: yes	-0.750	***	-0.708	***	-1.068	***
	SD Estimation: yes	-0.181	.	-0.195	.	-0.054	
k		269		269		181	
R^2		0.357		0.389		0.390	
Adjusted R^2		0.306		0.329		0.305	

Note. Estimation method: Weighted Least Square. ***k*** is the number of effect size measures.
Intercept: *Journal of Marketing Research*, Comparison group: C-IC.
C: Congruity; MI: Moderate Incongruity; IC: Incongruity.

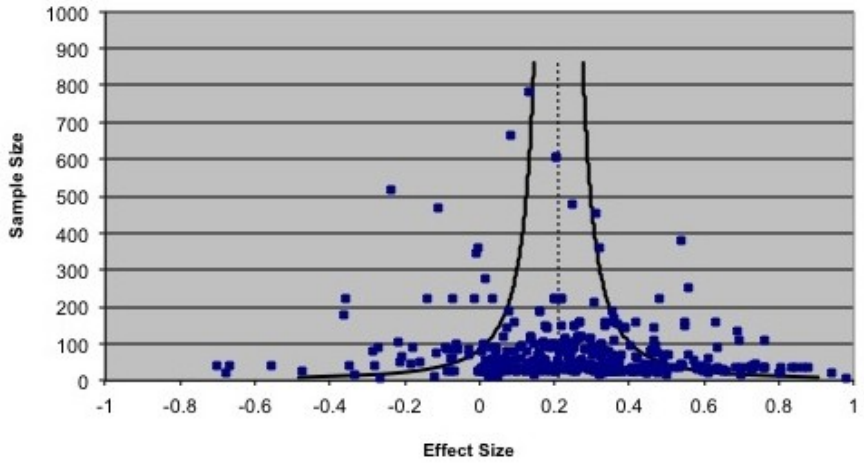
were generated from the same paper and the same study. In order to address this issue, we also estimated a model with generalized least squares using weighted data. The average correlation among observations from the same study turns out to be virtually zero, however, and the

estimated results are very similar. We present the results from the weighted least squares method in table 1.5.

Model 1 of table 1.5 reaffirms the *Congruity Hypothesis* and provides several further methodological conclusions. The model accounts for a moderate amount of total variance (R^2 = 0.357). We observe:

- The coefficients for the C-MI and MI-IC dummy variables are significantly negative at the .001 level. So the mean effect sizes for the C-MI and MI-IC comparisons are both significantly smaller than the mean effect sizes for the C-IC comparison, which is consistent with the *Congruity Hypothesis*.
- The coefficients on the behavioral intention and memory-related dummy variables are not significant (due to small sample sizes for these variables, noted earlier).
- The effect sizes from experiments using real parent brands (k =181) and fictitious parent brands (k = 88) were not statistically different.
- *Publication bias.* There is little evidence of publication bias. First, most journals have effect sizes not significantly different from the reference journal (*Journal of Marketing Research*). *Marketing Lett*ers, across three model estimations, shows a significant negative coefficient, but this result arises from only 3 effect sizes from one research paper in this journal. Second, to test whether the mean effect size is influenced by our screening procedure, we performed the fail-safe N test. This test indicates the number of non-significant, unpublished, or missing studies needed to nullify the meta-analysis results (Rosenthal, 1979). We find that 5,523 non-significant effects (r or d= 0.00) would be required to bring the congruity effect to a virtual zero level (r = 0.01 or d= 0.02). This result suggests that the conclusion of a significant positive mean effect size is unlikely to be affected by the selection criteria used in this meta-analysis. Third, we examine a funnel plot (figure 1.4), which shows a symmetric inverted funnel shape. This scatterplot of the treatment effect against a measure of study precision is used primarily as a visual aid for detecting bias or systematic heterogeneity. These do not appear to be problems.

Figure 1. 4. Funnel plot of effect size.



- Studies using between-subject experiment designs and within-subject designs indicate similar effect sizes.
- Studies using student samples report significantly smaller effect sizes than studies using other types of participants (general consumer samples or unknown sampling).
- Effect sizes are significantly larger when a manipulation check is successful than when a manipulation check is not used or unsuccessful. But, for reasons that elude us, studies with pretests report significantly smaller effect sizes than studies without pretests.
- The effect sizes reported in papers with incomplete information (on the measurement of standard deviations) are not significantly different from those with complete information.

Model 2 of table 1.5 tests the key moderators of the congruity hypothesis discussed earlier in figure 1.1, specifically, the increased elaboration to resolve incongruity (H1), the increased flexibility of categorization (H2), and negative information associated with the parent brand (H3). The model accounts for a moderate amount of total variance ($R^2 = 0.389$). We find:

- When the elaboration is increased/induced to resolve incongruity, the moderate incongruity effect pattern (Moderate Incongruity > Congruity > Incongruity) emerges. The congruity effect sizes for the comparison group C-MI becomes significantly smaller, which means congruity is much less preferred over moderate incongruity when the elaboration to resolve incongruity increases. Therefore, H1 is supported.

- When the flexibility of the categorization process is increased (bigger brand width, more positive mood, and case reminder present), some partial evidence for the moderate incongruity effect pattern (Moderate Incongruity ≈ Congruity > Incongruity) emerges. When the flexibility of the categorization process is increased, the congruity effect sizes for the comparison group C-IC become significantly larger, which suggests congruity is preferred to a greater extent over extreme incongruity. This can be because extreme incongruity is still hard to be accepted even with information increasing the categorization flexibility, and the negative responses toward extreme incongruity become even worse. However, the congruity effect sizes for the comparison group C-MI does not change significantly. Therefore, H2 is only partially supported.
- When there is negative information associated with the parent brand, the effect sizes for all three comparison groups become significantly smaller, which can be interpreted as the overall congruity effect is attenuated. Therefore, H3 is supported.

In model 3 of table 1.5, we further examine how various characteristics of the parent brands moderate the effect of congruity. This model is based on (181) effect sizes from studies that used real brands in their experimental designs. As additional moderators of the congruity effect sizes, we obtained data about brand characteristics by surveying consumer perceptions of the real brands in our dataset (see appendix 1.A.5). Using ex post surveys about consumer brand perceptions introduces measurement error, but if related coefficients are nevertheless significant, this is a conservative indication of effect. Model 3 shows improved model fit ($R^2 = 0.390$). We further note the following:

- *Brand attitude*. The effect sizes of congruity are significantly larger for brands with higher consumer brand attitude. This finding is consistent with existing arguments and results in the literature (e.g., Broniarczyk & Alba, 1994; Bambauer-Sachse, et al., 2011; Gierl & Huettl, 2011; Nan, 2006; Yeung & Wyer, 2005), and with our discussion of factors influencing affect transfer (p. 17).

- *Brand familiarity*. The effect sizes of congruity are found to be significantly larger for brands consumers are more familiar with. Although brand familiarity has not been explicitly discussed in brand extensions, it has been examined in product designs (i.e., Peracchio and Tybout's research on the moderating effect of brand knowledge on the congruity effect in 1996). If brand knowledge is extensive for familiar brands (limited for unfamiliar brands), consumers may be able (unable) to accommodate moderately or extremely incongruent brand extensions, which could facilitate (inhibit) affect/image transfer.
- *Brand history*. Brand history is another factor not previously examined relating to the congruity effect in the brand extension literature. Interestingly, we find a marginally significant negative effect of brand history on the congruity effect size. In other words, the effect sizes of congruity are smaller for brands with longer history than those with shorter history. This suggests that consumers may make certain inferences based on the length of a brand's history, which could influence the flexibility of the categorization process. As a result, consumers may be less rigid in their preference for congruent extensions when brands have a long history.

This last set of empirical results for brand characteristics (model 3 of table 1.5) add to past literature on brand extensions. This concludes our review of congruity theory for brand extensions, which began with the seminal work of Aaker and Keller (1990).

Conclusions & Discussion

The literature on brand extensions has come a long way since 1990. The main idea is that consumers will more readily accept an extension of a brand if the extension product category is congruent (or fits) with the parent brand (or category). The papers summarized in table 1.1 alone account for around 16,000 citations. This is a powerful idea that has stood the test of time.

With growth of this literature, researchers have pushed the boundaries of research on the congruity effect by discussing and examining various moderators of the congruity effect, and various alternative effect patterns (e.g.. the moderate incongruity effect pattern, the incongruity

effect pattern). Given the range of results from this large body of literature, a systematic and comprehensive review and synthesis can help organize the various findings to date to facilitate future work and to provide a resource for practitioners.

This paper addresses three sets of research questions: (1) *How large is the congruity effect for brand extensions? And how much variation is there?* (2) *What are the relevant factors, in addition to congruity, that influence the success of brand extensions? And how do they interact with congruity?* (3) *Can moderate incongruity generate more favorable responses to brand extensions from consumers than congruity? And under what circumstances?*

As a conceptual synthesis and as the first meta-analytical review of the congruity literature for brand extensions, our findings help answer some fundamental questions about this topic. (1) For the first set of research questions, our literature review of the key literature (table 1.1 and appendix 1.B) shows a consistent positive effect of congruity on consumers' responses, and our meta-analysis reveals that the overall mean effect size of congruity ($\bar{R}$=0.211) is of small to medium size, that there is support for the *Congruity Hypothesis* effect pattern (Congruity > Moderate Incongruity > Incongruity), particular when the dependent variable is consumer attitudes toward the brand extension, but that the variation of the congruity effect sizes across studies is large. (2) For the second set of research questions, our conceptual review of moderators of the congruity effect (figure 1.1) identifies and proposes some key factors that may interact with congruity to influence consumers' responses. Our meta-analysis finds a very significant positive effect of congruity for attitudinal dependent variables (H4), that negative information about the parent brand nullifies the congruity effect (H3), and that parent brand attitude and brand familiarity positively interact with congruity. (3) For the third set of research questions, our conceptual review and meta-analysis together identify two important factors that

may moderate the congruity effect into a moderate incongruity effect pattern: increased elaboration to resolve incongruity (H1), and increased flexibility of the categorization process (H2).

Discussion

There are several managerial, theoretical, and methodological implications of the above results.

First, it is worth commenting on the ramifications of our first set of conclusions that the overall mean effect size of congruity is significantly positive, but of small to medium size ($\bar{R}$ =0.211), with a high variance, and that there is support for the *Congruity Hypothesis* (Congruity > Moderate Incongruity > Incongruity). Although these results are consistent with the results in the literature, this is the first paper in the literature that summarizes the average magnitude of the congruity effect. This directly speaks to the relative importance of congruity. Although congruity is significant and very relevant from a managerial perspective, it is far from being the sole determinant of success for brand extensions.

Second, the high variation of the congruity effect sizes raises the important question of when the congruity effect can be increased and when it will be turned off. Researchers and practitioners care about what to do and what not to do to ensure the success of brand extensions. Our conceptual review and synthesis suggests how the congruity positively affects consumers' evaluation of brand extensions is via the categorization and affect transfer process (CAT mechanism). Our meta-analysis, especially the moderator analysis in table 1.5, shows that some brand characteristics are important factors interacting with congruity to influence consumers' responses. Specifically, when consumers are more familiar with and have a more positive attitude toward a parent brand, the effect size of congruity is enlarged. In other words, for

familiar brands and brands with positive brand attitudes, they should obey the "more congruent, the better" strategy, and utilize their current brand associations to the maximum extent by extending to similar product categories. In addition, when consumers have negative perceptions of a parent brand, the congruity effect is overridden or nullified (H3). In order words, if the parent brand is perceived negatively, extending to congruent or incongruent categories does not particularly matter, and consumers usually respond negatively to its extensions.

Third, our conceptual synthesis and meta-analysis also identify two situations where moderate incongruity can lead to more or equally favorable results, specifically, when incongruity can be resolved via increased elaboration (H1) and when the flexibility of the categorization process can be increased (H2). These results provide very important implications for practitioners. Other than the simple "more congruent, the better" strategy, there is an alternative strategy that a company could extend to a moderately incongruent product category, and employ marketing campaigns designed to increase flexibility of categorization or induce greater elaboration about any perceived incongruity between the brand and its extension product. It is also worth mentioning for future theoretical consideration that we think one key difference between the influencing elaboration and categorization flexibility is that they may happen at different stages of the categorization process. Moderators influencing the flexibility of categorization usually affect consumers' initial judgment of congruity, but moderators influencing elaboration could happen or be induced later to update the initial judgment. For managerial purposes, this could provide implications for the timing to use certain marketing techniques to induce consumers' elaboration, or increase consumers' flexibility of their categorization process.

Fourth, our meta-analysis has methodological implications for research examining the effect of congruity on consumers' evaluation of brand extensions. In particular, the use of student samples may lead to under-estimation of the congruity effect sizes, which suggests that general consumer samples may lead to a more realistic estimation of the congruity effect. Moreover, the use of manipulation checks may be an important element in research designs. As previous research suggests (e.g., Barone, 2005), the order of these manipulation check questions may be even more important, because such questions before the final dependent variable, such as similarity measures between a parent brand and its extension product, may serve as a reminder or prime of the importance of congruity.

Fifth, the methodological paradigm of this paper could provide useful guidance for future experimental meta-analyses that aim to consider possible nonlinear effect size patterns. Unlike traditional meta-analyses of experimental research, which typically compare two experiment levels and estimate the effect size, the current paper considers three experimental levels (i.e., congruity, moderate incongruity, and extreme incongruity), and overcomes difficult coding issues (e.g., inconsistent labeling) to examine the overall effect pattern.

Lastly, one important question for brand extensions concerns the reverse effect of how a brand extension influences consumers' responses to the original brand (which was not the focus of the current paper). In our literature search and review, we identified some research investigating this question (e.g., Loken & John, 1993; Mathur, et al., 2012; Martinez & Pina, 2003; Milberg, et al., 1997; Sood & Keller, 2012), and they suggest that similar extensions could help the parent brand, but dissimilar ones could hurt the parent brand by decreasing the brand image and diluting the brand belief/personality. We coded these papers and identified the overall mean effect size of congruity on consumers' responses toward the original parent brand is 0.130

($\bar{R}$) from 16 effect sizes. This result is very interesting, because it suggests that the image transfer process can happen in reverse, but the magnitude of the transfer may be different (the congruity effect is positive on consumer responses to the original parent brand, but the magnitude of the effect size is small), which is definitely an avenue for future research.

Future research

Although the literature on the congruity effect for brand extensions is sizable, there are still many issues that merit future consideration.

First, because of a limited number of data points, the current meta-analysis did not identify a clear effect pattern for behavior intention and memory-related outcome variables (H4, H5). This suggests that the past literature put too much focus on attitudinal variables, and ignored other outcome variables that may also be critical for marketing, such as consumers' behavioral intention, real purchase behavior, memory recall and recognition, and so on. Future research should further explore the effect of congruity on these outcome variables.

Second, in addition to congruity (and its moderators), other factors (e.g., marketing support, parent-brand conviction) can also be important for brand extensions (Volckner & Sattler, 2006). Future research should continue to identify and examine the effect of such factors, and compare the relative impact of these factors. Future reviews should build on and go beyond the scope of the current meta-analysis, and try to identify and code enough papers of each factor's effect, and synthesize them using other methods (e.g., meta-analysis using structural equation modeling).

Third, congruity has been considered in many application areas in marketing, including product designs, brand alliances, celebrity endorsement, event sponsorship, cause-related marketing, and even consumer reception of particular linguistic forms. Some of these application

areas specifically consider a categorization and transfer process as the underlying mechanism. Other application areas deal with contexts where some incongruity is expected (e.g., new products), and there is greater focus on elaboration and the moderate incongruity effect pattern. Still other application areas involve contexts where completely different theories apply. A challenge for future research is to compare and synthesize the effect of congruity across various marketing application areas and understand the different consumer assessment processes at work. The largest application area of congruity research, thus far, has been brand extensions, but other application areas are rapidly growing and are ripe for further consideration (e.g., new product designs).

Appendix 1.A.

Methodological Detail Appendix

A.1. Selection of the databases and journals

A keyword search (congruity, incongruity, congruence, match-up, similarity, fit) in the abstract was conducted in two main sources: (1) databases, including the *Web of Science*, *Business Complete*, and *Proquest Dissertations*; and (2) ten highly-ranked marketing journals: *Journal of Marketing, Journal of Marketing Research, Management Science, Marketing Science, Journal of Consumer Research, Journal of Retailing, International Journal of Research in Marketing, Quantitative Marketing and Economics, Journal of Advertising,* and *Journal of Advertising Research*. As background, the first five of the ten listed journals here are the marketing journals in the current Financial Times Top 45 list for Management Journals; the first seven are A+ and the other three are A in "ABCD Journal Quality List 2013" used in Australia and New Zealand; the first five are in the category are "Widely Recognized/Elite Journals for Management Scholarship" and the next three are in the "Other Distinguished Journals" category in the "Consortium List of Top Management Journals (June 2012)" used by full-service research universities in Canada. The last two journals in the list cover topics closely related to congruity.

There is redundancy of coverage, which helps insure identification of key papers. At the same time, some of these ten journals did not include many papers on the topic, while the databases, which include other marketing journals, and many journals in other areas of management, psychology and other social sciences, picked up several other important journals, such as *Journal of Consumer Psychology.* The papers that ultimately met the inclusion criteria for this meta-analysis came from eleven journals, as shown in the table 1.2.

We emphasize that our screening process covered many dozens of journals outside of marketing, in the social sciences and elsewhere, and in other areas of business; but after review of the abstracts, the relevant papers ultimately came from the above journals.

A.2. Correction of inconsistent labeling problems

Researchers use different scales to measure congruity in their pre-tests or manipulation checks. Some of them use single-item scale (e.g., congruent-incongruent); while others use multiple-item scale (e.g., congruent-incongruent, consistent-inconsistent, and similar-dissimilar). Despite the difference in the number of items used, they usually calculate an index based on their scales, and use this index to represent the level of congruity. However, how they label their experimental stimuli is subject to their own judgment. Some researchers are only interested in the relative comparison between two significantly different levels of congruity, while others are more interested in the accurate levels of congruity. In this case, for the same stimuli (e.g., one's pretest or manipulation check test result lies on the congruent extreme of the scale; the other one's lies in the middle range of the scale), researchers in the former situation will name them as congruity and incongruity conditions, while those in the latter situation will name them as congruity and moderate incongruity conditions. We refer to this problem as inconsistent labeling.

In order to correct this problem, we followed the correction scheme as below.

1. If the study reports both pre-test information and manipulation check information, check the pretest information and make corrections according to the table below.
2. If the study reports only manipulation check or pretest information, check this information and make corrections according to the table below.
3. If the study did not report either pretest or manipulation check information, use the original labels named by the researcher.

Table 1. A. 1. Correction of mislabeling problem.

Scale	Incongruity	Moderate Incongruity	Congruity
A to B point scale A: incongruity; B: congruity	A to ((B-A)/3-0.01)	(B-A)/3 to ((B-A)*2/3-0.01)	(B-A)*2/3 to B
1 to 7 point scale	1 to 2.99	3 to 4.99	5 to 7
Examples 1 to 5 point scale	1 to 2.32	2.33 to 3.65	3.66 to 5
0 to 100 point scale	0 to 33.32	33.33 to 66.65	66.66 to 100

A.3. Coding procedure for studies without complete information for effect size calculation

One complication in the coding process is that some studies do not report all the required information to calculate the effect sizes. As a general principle, we excluded studies when there was a preponderance of missing information, and we used approximations when most of the needed information was present. In particular, we excluded studies that failed to report the group means, the standard deviations, and any other statistical tests (*t* value, *F* value, *p* value). We also excluded studies that did not report any statistical information, but only verbally mentioned that there was no significant difference between two experimental conditions. For studies that only report group means but no standard deviations, the standard deviations for each group are deduced from other related information (typically, reported statistical tests) in that study. If there is not enough information available to deduce the standard deviation, the mean scores are recorded first, and the standard deviations were estimated from other studies. For studies that do not report sample sizes for each condition, but reported the total sample size, we assumed equal sample size across various conditions. For studies that report significant interaction effects between congruity and other variables (e.g., familiar versus unfamiliar focal brand), the effect sizes were calculated separately for each moderator condition to retain the interaction effect (following Scheibehenne, et al., 2010).

A.4. Correction for effect sizes from within-subject studies

For studies using within-subject designs, the effect sizes were corrected by using the formula:

$$d_{IG} = d_{RM}\sqrt{2(1-\rho)}\,,$$

where ρ is the correlation between the repeat measures; d_{IG} is the effect size of independent groups; and d_{RM} is the effect size of repeat measures. Since the articles in our dataset typically do not provide sufficient data to calculate ρ, following Cooper et al. (2009) we used $\rho = 0.7$ in those cases. We also did sensitivity analysis, available on request, with $\rho = 0.875$ and 0.5, with similar results to those reported in this manuscript.

A.5. Data collection for characteristics of real parent brands

In brand extensions, some studies use fake parent brands, while others use real ones as their experimental stimuli, which gives us the ability to test the potential moderating effect of some brand characteristics on the congruity effect. From our database, we identified that 45 real parent brands are used as experiment stimuli. In order to collect information on consumers' perceived characteristics of these brands, a survey was conducted on Mturk. As a result, 204 participants in US were recruited, and were assigned to evaluate 15 brands (randomly selected from the 45 brands). For each brand, they were asked to answer the following questions: (1) *Brand familiarity***:** "To what extent are you familiar with Brand X?" (1: never heard of; 7: extremely familiar); (2) *Brand attitude***:** "How do you like/dislike Brand X?" (1: extremely dislike; 7: extremely like); (3) *Brand quality:* "What do you think of the quality of the service/products of Brand X?" (1: extremely low-quality; 7: extremely high-quality); (4) *Brand history:* "To your understanding, how long is the history of Brand X?" (1: extremely short; 7: extremely long); and (5) *Brand breadth:* "To your understanding, how wide/narrow is the business of Brand X?" (1: extremely narrow; 7: extremely wide).

Appendix 1.B.

Key Correlational Research On Congruity In Brand Extension

Table 1. B. 1. Key correlational research on congruity in brand extension.

Authors	Key Findings
Aaker and Keller (1990). JM	This paper found that (1) attitude toward the extension was higher, when there was both a perception of "fit" between the two product classes along one of three dimensions and a perception of high quality for the original brand or the extension was not regarded as too easy to make; and (2) potentially negative associations can be neutralized more effectively by elaborating on the attributes of the brand extension than by reminding consumers of the positive associations with the original brand.
Smith and Park (1992). JMR	The findings indicate that brand extensions capture greater market share and realize greater advertising efficiency than individual brands. The relative effect of brand extensions on market share is not moderated by the degree of similarity between the extension and other products affiliated with the brand. Advertising efficiency effects, however, are elevated when similarity is high, but only when it is based on intrinsic attributes.
Sunde and Brodie (1993). IJRM	Consumer acceptance of a proposed brand extension will tend to be higher if: (1) the perceived quality of the brand is high, (2) there is perceived fit between the two product categories, especially in terms of the transferability of the skills and the complementarily of the two products, (3) the extension is in a category which is difficult to make
Bijmolt, Wedel, Pieters, and DeSarbo (1998). IJRM	This paper provides empirical insight into the way consumers make pairwise similarity judgments between brands, and how familiarity with the brands, serial position of the pair in a sequence, and the presentation format affect these judgments.
Garlin and McGuiggan (2002). P&M	This study empirically tests the congruity between movie-content preference and choice on television, video, and at the cinema; and the impact of consumer involvement on this relationship. Involvement with movie choice was found to vary by medium, yet it did not unequivocally change the nature of preference–choice congruence.
Echambadi, Arroniz, Reinartz, and Lee (2006). IJRM	This paper find that, although the simple effects of neither parent brand quality nor measures of fit affect evaluations of brand extensions, the interaction effects of parent brand quality with fit are important determinants of brand extension evaluations.
Pina, Martinez, de Chernatony, and Drury (2006). EJM	This paper found that the extent of perceived fit between the corporate brand and the service extension influences the perceived quality of the extension, which in turn affects corporate image, especially for corporate brands that originally had highly rated images.
Völckner and Sattler (2006). JM	The paper find that (1) fit between the parent brand and an extension product is the most important driver of brand extension success, followed by marketing support, parent-brand conviction, retailer acceptance, and parent-brand experience, (2) several important structural relationships among the investigated success factors; (3) the interaction terms of fit with the quality of the parent brand and with parent-brand conviction are significant, albeit of relatively low importance.
Völckner and Sattler (2007). IJRM	The authors investigated the empirical generalizability of existing brand extension research results. Using a comprehensive data set compiled from two large-scale consumer samples and panel data, they address the generalizability of empirical findings.
Martinez, Polo, and de Chematony (2008). IMR	The purpose of this paper is to propose and test a model that shows how extending a brand affects the overall brand image. Results show that the brand extension strategy dilutes the brand image, and that brand image before extension and fit has positive effects on brand image after extension.
Voelckner, Sattler, and Kaufmann (2008). ML	This paper examines the issue of image feedback effects and potential drivers of these effects by using real brand extensions and a longitudinal field study. It's found that even for successful extensions, negative image feedback effects can occur, particularly when the perceived quality of the extension fails to meet the quality level of the parent brand. Strong brands tend to be more vulnerable to negative image feedback effects because consumers have a higher reference level for their extensions than for those of weaker brands. The likelihood of negative feedback effects decreases as the level of perceived fit and consumers' perceptions of the general extendibility of the parent brand increases.
Martínez Salinas and Pina Pérez (2009). JBR	The results verify that extension attitude influences brand image, whereas initial brand associations and perceived fit between the new product and either the remaining products (category fit) or the brand image (image fit) are able to strengthen consumer attitude. The study also explains the role of consumer innovativeness as a moderating factor.
Arslan and Altuna (2010). JPBM	The results show that brand extensions affect the product brand image negatively, whereas the fit between the parent and extension brands decreases the negative effect. The drop of image as a result of extension is greater when the perceived image and quality of the parent brand are higher. Perceived quality of the brand, consumers' brand familiarity, fit perceived by the consumer, consumers' attitudes towards the extension have a positive effect on the product brand image after the extension.
Sichtmann, Schoefer, Blut, and Kemp (2017). EJM	This paper investigates extension category effects on service brand extensions, and the extension category's influence on brand/consumer-level success drivers and the perceived quality of the extension.The findings indicate a general and consistent extension category-dependent effect that moderates the importance of brand extension success drivers.

Essay 2

Dimensions of Brand Extension Fit:

A Measurement Scale of Congruity for Brand Extensions

Abstract

A sizable literature on brand extensions argues that the congruity or fit between a parent brand and an extension product (category) has a positive effect on consumer reception of the brand extension. However, the application of this literature is very limited, due to a lack of understanding of what "brand-extension fit" really is.

The current paper develops a measurement scale of Brand Extension Fit (BEF) consisting of two core dimensions, *engineering-based* and *market-based congruity*, each measured by three items. Our proposed scale represents a synthesis and extension of past work on congruity measurement. We validate the proposed measurement model with two separate datasets obtained from general consumers consisting of judgments about extensions for (a) fictitious parent brands and (b) real parent brands.

Unlike uni-dimensional measures of fit – such as similarity, fit, or consistency– used in most of the brand extension literature, our scale provides guidance for opportunity identification, idea generation, understanding the pros and cons of various alternatives, and the building of a marketing plan around a chosen alterative.

Keywords: Congruity; Brand Extensions; Dimensionality; Formative Measurement Scale; SEM-Partial Least Squares.

Introduction

A brand extension is a marketing strategy, in which a firm marketing a product with a well-developed image uses the same brand name in a different product category. Indeed, brand extensions are indispensable to brand renewal and growth, because they are a widely-used means of reducing the risk and expense of new product introductions (McCarthy, Heath, & Milberg, 2001). The general idea is to harness a brand's popularity in one area to facilitate positive consumer reception in a new area. Nevertheless, brand extensions are still risky. According to Völckner and Sattler (2006) and Torelli and Ahluwalia (2012), for many fast-moving consumer goods, the failure rate of brand extensions can be over 80%.

To address this critical managerial activity, academics and practitioners recognize that the fit between an extension product and a popular parent brand helps yield favorable consumer responses. Summing up the extensive literature that has developed, Völckner and Sattler (2006) identified five key factors driving brand extension success and found that the fit between the parent brand and an extension product is the most important driver (others include marketing support, parent-brand conviction, retailer acceptance, and parent-brand experience). Similar conclusions come from the practitioner side: for example, John Parham, president of the brand extension agency Parham Santana, summarizes three pillars for successful brand extensions: fit, leverage, and opportunities (Klara, 2013).

But, although an abundance of research shows that higher levels of congruity are associated with higher likelihood of brand extension success, after almost 30 years of academic research and more than 60 years of practice, the specific guidance for managers is limited and sometimes contradictory.

To illustrate the point, imagine a scenario in which the brand manager of a coffeehouse chain, Jaydon Lee, wants to make a decision in terms of which product categories his company should extend to. Jaydon will need advice in at least three stages of decision-making: (a) *Opportunity identification and idea generation*: Jaydon will need help identifying business areas (categories) and product ideas amenable for a brand extension: the, so called, "low hanging fruit." (b) *Selection from alternatives*: Once he has a list of alternatives, Jaydon will need to rank the alternatives from best to worst. When two alternatives have a similar fit, Jaydon will consider qualitative issues to break any ties. (c) *Development of a marketing plan:* Jaydon will need clear insight into what makes the selected brand extension idea work to help develop a marketing plan that effectively communicates with consumers. It would be desirable if academic research could help Jaydon with this process. The literature, however, provides limited guidance.

For *opportunity identification and idea generation*, most research fails to offer specific direction into kinds of new products, markets, or areas to consider. Since 1990, the literature has mostly studied congruity as a uni-dimensional construct (measured by reflective items – which are essentially synonyms – like "fit", "similarity", "congruity", or "consistency"). A uni-dimensional construct/scale is good for ranking alternatives, if one already has particular product ideas or prototypes under consideration, but not particularly helpful for providing guidance as to where to look for potential extension product ideas. The scale items like "fit", "similarity", "congruity", or "consistency", etc., are all too general to point to specific directions.

For *selection from alternative extension ideas*, a uni-dimensional congruity scale (measured by a single or multiple reflective items) does provide guidance for collecting assessments of different alternatives from a consumer panel and creating a ranking. This does aid Jaydon in decision-making. But if the top alternatives are very close in terms of overall perceived

congruity, a uni-dimensional scale is silent as to the trade-offs between the top-ranked alternatives. It is then useful to consider the qualitative nature of the fit of an extension with its parent brand. This is facilitated by considering fundamentally different aspects of congruity in the literature (e.g., product complementarity, product substitutability, and resource transferability by Aaker and Keller in 1990; feature and concept similarity by Park, Milberg and Lawson in 1991; attribute, benefits, and image similarity by Lefkoff-Hagius and Mason in 1993; and attribute, benefit, and value-level goal congruency by Martin, Stewart and Matta in 2005). This work helps Jaydon think about the pros and cons of different alternatives, but this literature is not definitive because the various authors have divergent points of emphasis, priorities, and conclusions, and there is little work that compares these different points of emphasis. This stream is in need of a better synthesis in offering guidance for Jaydon.

For *developing a marketing plan* around a selected alternative, there has been much research since roughly 1995 that considers the impact of various moderators (e.g., individual differences, marketing efforts, contextual factors) that can be used to leverage the positive effect of congruity (we review and summarize this research in the next section). This work can help in enhancing and fine-tuning a marketing plan. But to get to the essence of what to communicate, Jaydon must understand which aspects of congruity consumers are responding to. Jaydon is not enabled in this effort with a scale that measures a uni-dimensional latent construct, and would be better served by a multi-dimensional scale (composed of potentially different formative measurement dimensions/items).

To better help brand managers, like Jaydon, with these three stages of decision-making, there is a need for a clear conceptualization of the multi-dimensional nature of congruity as pertain to brand extension success. The main objective of this paper is, accordingly, to develop a

comprehensive measurement scale that incorporates the key dimensions of congruity for brand extensions. A secondary methodological objective of this paper (required for the first objective) is to advance and carry out a framework for formative measurement scale development.

To achieve these objectives, the current paper (a) considers multiple items that form and define congruity, (b) narrows them down to two core dimensions: *market-based congruity* and *engineering-based congruity* (explained and defined in the results section), (c) estimates and validates a measurement model for these two core dimensions, (d) shows that these two dimensions make up congruity as a whole with predictive power as good as a traditional reflective scale of congruity, and (e) provides some theoretical conjectures about why these two dimensions arise, with discussion of future research directions. We then explain how such a multi-dimensional model can better help brand managers pursue successful brand extensions.

Literature Review

Congruity in Brand Extensions

Since the seminal work of Aaker and Keller (1990), mentioned above, and the refocus on the fit of an brand extension with its parent brand (rather than with its parent category, as emphasized by Bronjarczyk and Alba 1994), much experimental and correlational research in brand extensions has consistently replicated the positive effect of congruity on consumers' responses toward brand extensions. Specifically, consumers respond more favorably to congruent brand extensions than to moderately incongruent ones, and, more favorably to moderately incongruent brand extensions than to extremely incongruent ones. The typical dependent variables are customer attitudes or behavioral intentions toward the extension product.

This literature consistently builds on the theories of categorization (Cohen & Basu, 1987; Fiske, 1982; Fiske & Pavelchak, 1986; Meyers-Levy & Tybout, 1989; Sujan, 1985; Tversky,

1977) and affect/image transfer (Boush et al., 1987; Shimp, 1981; Wright, 1975). In particular, congruity between a brand and its extension product promotes the categorization of the extension product with the brand. This facilitates positive affect/image transfer from the parent brand to the extension product.

A large number of studies have subsequently explored moderators of the congruity effect on consumers' responses. These moderators can be categorized into three groups: (a) *characteristics of the parent brand*, including brand breadth (Boush & Loken, 1991), brand quality (Keller & Aaker, 1992), brand affect (Yeung & Wyer, 2005), brand attitude (Gierl & Huettl, 2011; Nan, 2006), brand emotional attachment (Fedorikhin, Park, & Thomson, 2008), brand equity (Buil, de Chernatony, & Hem, 2009), and brand positioning (Liu & Hu, 2012); (b) *other information cues present in the environment*, including similarity/fit primes (Barone, Miniard, & Romeo, 2000; Zhang & Sood, 2002; Yeung & Wyer, 2005), ad exposure time (Lane, 2000), independence versus interdependence primes (Ahluwalia, 2008), the presence of art (Oakley, Duhachek, Balachander, & Sriram, 2008), competitive cues (Kappoor & Helson, 2009; Milberg, Sinn, & Goodstein, 2010), brand portrayals, brand slogans, and peripheral design cues (Gierl & Huettl, 2011), the consumption occasion (Liu & Hu, 2012), family-brands versus sub-brands (Sood & Keller, 2012), and physical distance (Huang, Jia, & Wyer, 2017); and (3) *individual differences*, including mood (Barone, Miniard, & Romeo, 2000; Yeung & Myer, 2005), involvement level (Barone, 2005; Maoz & Tybout, 2002), age (Zhang & Sood, 2002), regulatory focus (Yeo & Park, 2006), analytical versus holistic thinking (Monga & John, 2008), construal level (Kim & John, 2008), country of origin of the parent brand (Buil, de Chernatony, & Hem, 2009), incremental versus entity orientation (Mathur, Jain, & Maheswaran, 2012),

arousal level (Noseworthy, Muro, & Murray, 2014), and the nature of the purchase goal (Dimitriu, Warlop, & Samuelsen, 2017).

Some of these variables turn out to be nearly as important as congruity for the success of a brand extension, such as positive brand attitude, brand affect, and brand quality. Some of them nullify congruity's effect on consumers' responses, such as negative brand quality or associations, competitive cues, and very high or low arousal levels. Also, for some moderators (e.g., involvement level), a different effect pattern of congruity is identified: the moderate incongruity effect (Mandler, 1982). Specifically, moderately incongruent brand extensions receive more favorable responses from consumers than both of congruent ones and severely incongruent ones, and the congruent brand extensions lead to more favorable outcomes than extremely incongruent ones.

Dimensionality of Congruity

In the above-described literature, the most common approach to measuring congruity is treating congruity as a uni-dimensional reflective construct. According to this approach, congruity is defined as *consumers' overall perception of the similarity, fit or consistency between the parent brand (product category) and the extension product*. As shown in table 2.1, there are two slightly different uni-dimensional measurement methods for congruity. (1) Consumers' overall perceived congruity may be measured by a single item, such as "similar/dissimilar", "good/bad fit", or "consistent/inconsistent". Or (2) congruity may be measured by multiple reflective measurement items. For example, Shen, Bei, and Chu (2011) used five items: "fit," "reasonable," "connected," "associated," and "understandable." These five items are theoretically nearly identical, and methodologically interchangeable – adding or dropping any of them will not change the theoretical content/domain of the congruity construct.

Table 2. 1. Reflective measures of congruity in brand extensions.

Research Paper	Type of Research	Measure Used
Boush and Loken (1991). JMR	Exp	The similarity of each potential brand extension to products the brand currently makes (1 = dissimilar, 7 = similar).
Broniarczyk and Alba (1994). JMR	Exp	A 9-point scale that ranged from "not similar" (1) to "very similar" (9).
Bijmolt, Wedel, Pieters, and DeSarbo (1998) IJRM	Cor	A 7-point scale: 1: shighly dissimilar, 7: shighly similar..
Morrin (1999). JMR	Exp	A scale of 1 = "very bad fit" to 9 = "very good fit";
Barone, Miniard, and Romeo (2000). JCR	Exp	How similar the extension was to the current products marketed by [the parent brand] (1 = not at all similar; 7 = very similar)
Lane (2000). JM	Exp	Two seven-point scales anchored at zero and six— good fit/bad fit and extremely consistent/extremely inconsistent.
McCarthy, Heath, and Milberg (2001). ML	Exp	How well the target brand name seemed to fit with the 35mm camera category on a 9-point scale (1:not much, 9:very much).
Maoz and Tybout (2002). JCP	Exp	A single 9-point semantic differential scale.
Zhang and Sood (2002). JCR	Exp	A fivepoint scale (1 = Not at all similar, 5 = Very similar).
Barone (2005). JCP	Exp	A 7-point scale ranging from 1 (not at all similar) to 7 (very similar)
Yeung and Wyer (2005). JMR	Exp	Each product's relationship to airline services on a scale that ranged from -5 ("'not at all") to +5 ("very").
Nan (2006). P&M	Exp	(1). How much sense does it make for the (product category) brand (brand name) to introduce digital cameras? (2). How logical is it for the (product category) brand (brand name) to introduce digital cameras? (3). How do digital cameras fit with the (product category) brand (brand name)? (4). How surprised are you that the (product category) brand (brand name) will introduce digital cameras?
Yeo and Park (2006). JCP	Exp	Participants rated the perceived similarity between the parent brand and the extension along a scale of 1 (very dissimilar) to 7 (very similar).
Monga and John (2007). JCR	Exp	Respondents then evaluated brand extension fit on a scale from 1 (“inconsistent with the parent brand”) to 7 (“consistent with the parent brand”)
Shine, Park, and Wyer (2007). JMR	Exp	Participants rated perceptions of parent–extension similarity along a scale ranging from 1 (“very dissimilar”) to 7 (“very similar”).
Fedorikhin, Park, and Thomson (2008). JCP	Exp	Three-item 7-point scale anchored with agree/disagree: [the parent brand category] and [extension product] are very similar, [the parent brand category] and [extension product] go together really well, and [extension product] is a natural extension for a [the parent brand category] company.
Hagtvedt and Patrick (2008). JCP	Exp	The extent to which they thought the[parent brand] MP3 player had a close fit with each of the extension products (1=not at all, 7=very).
Kim and John (2008). JCP	Exp	Participants judged the fit of the brand extension with the brand on two 7-point scales (inconsistent/consistent and atypical/typical)
Martinez, Polo, and de Chematony (2008). IMR	Cor	(1) How similar or dissimilar are “new product” to the products usually offered by X? (1 – very dissimilar, 7 – very similar); (2) How inconsistent or consistent is the new product with X’s brand image? (1 – very inconsistent, 7 – very consistent)
Buil, de Chernatony, and Hem (2009). EJM	Exp	Respondents indicated the degree of similarity, on a seven-point Likert scale.
Kapoor and Heslop (2009). IJRM	Exp	Participants rate the potential brand extensions in terms of their fit with the parent brand. The 9-point scale (1 = no sense at all; 9 = a lot of sense)
Milberg, Sinn, and Goodstein (2010). JCR	Exp	The fit manipuladon is then measured on two scales (1 = very low fit, 7 = very high fit) and (1 = makes little sense, 7 = makes a lot of sense).
Bambauer-Sachse, Hüttl, and Gierl (2011). P&M	Exp	Perceived fit was measured using four items according to the recommendations of Boush and Loken (1991), Dawar and Anderson (1994), and Bridges, Keller, and Sood (2000).
Gierl and Huettl (2011). IJRM	Exp	Four response items: “The core product and this extension are very similar/not at all similar”; “The core product and this extension possess a very high/very low fit”; “I can understand the connection very easily/not at all”; and “The extension is logical and makes sense to a very high/very low degree”
Shen, Bei, and Chu (2011). P&M	Exp	Five items of consumers’ perceptions of extension fit (“fit,” "reasonable,” “connected,” “associated,” and “understandable”)
Liu and Hu (2012). P&M	Exp	Similarity was measured by a 7-point scale (1 = very dissimilar, 7 = very similar).
Mathur, Jain, and Maheswaran (2012). JCP	Exp	Respondents indicated their perceptions of extension fit on two seven-point items anchored by “is very dissimilar(1)/ similar(7) to the Cheerios brand,” and “has a low-fit (1)/has a high-fit (7).”

Table 2.1. Reflective measures of congruity in brand extensions. (continued)

Research Paper	Type of Research	Measure Used
Sood and Keller (2012). JMR	Exp	Three seven-point scales: "bad fit between company and product/ good fit between company and product," "not at all logical for company/very logical for company," and "not at all appropriate for company/very appropriate for company."
Milberg, Goodstein, Sinn, Cuneo, and Epstein (2013). JMM	Exp	Two seven-point fit scales (1 = 'very low fit', 7 = 'very high fit', and 1 = 'makes little sense', 7 = 'makes a lot of sense'
Noseworthy, Muro, and Murray (2014). JCR	Exp	Three items captured participants' perceived typicality (is common, is likely, matches expectations)
Huang, Jia, and Wyer (2017). P&M	Exp	A scale from 1 = inconsistent with [the parent brand] to 7 = consistent with [the parent brand].
Dimitriu, Warlop, & Samuelsen (2017). EJM	Exp	Two items about the perceived similarity between mobile phones and each of the product categories where the brands Scera and Myrto were established (1= "very dissimilar" / 7= "very similar"),

Note.Cor: correlational research; Exp: experimental research; JMR: *Journal of Marketing Researc*h; IJRM: *International Journal of Research in Marketing*; JCR: *Journal of Consumer Research*; JM: *Journal of Marketing*; ML: *Marketing Letters*; JCP: *Journal of Consumer Psychology*; P&M: *Psychology & Marketing*; IMR: *International Marketing Review*; EJM: *European Journal of Marketing*; JMM: *Journal of Marketing Management*.

The popularity of this approach can be attributed to the fact that overall perceived congruity (as a uni-dimensional reflective construct) is easy to measure and is able to predict some key variables that researchers and practitioners are interested in, such as consumers' overall attitudes and behavioral intentions toward the brand extension. However, the disadvantage of this approach is that it fails to provide information about how overall congruity is formed, or what causes the changes in the overall congruity.

An alternate measurement approach views congruity as a multi-dimensional formative construct. Unlike the uni-dimensional approach (with its focus on outcomes related to consumer reception of brand extensions), this stream of research focuses on how congruity is formed. We synthesize the perspective of this stream of research and develop a definition of congruity as a formative construct: congruity is consumers' overall perception of similarity or fit between the parent brand and the extension product*, formed by consumers' separate evaluations of congruity on relevant dimensions or aspects of the brand extension*. In other words, congruity is an outcome of an evaluation process, where consumers first decompose congruity into specific

aspects/dimensions, then evaluate them sequentially or simultaneously, and lastly form an overall perception of congruity. This process can be a single-directional process, or an interactive cycle, where consumers can always go back to any specific dimension/aspect and update their evaluation.

Some aspects/dimensions of congruity, on which consumers evaluate the brand extension, have been proposed and measured in the brand extension literature. Indeed, Aaker and Keller (1990) began this literature by focusing on three key aspects of congruity, measured by variables that they refer to as substitute, complement, and transfer (see table 2.2 for details). These three measures of congruity have been adopted in subsequent research (e.g., Echambadi et al., 2006; Kalamas, Cleveland, Laroche, & Laufer, 2006; Sunde & Brodie, 1993). Since then, other dimensions/aspects have been proposed as well. For example, Park, Milberg, and Lawson (1991) find that when evaluating brand extensions, consumers not only evaluate the product-level feature similarity, but also consider the concept consistency between the brand and the extension. Furthermore, Smith and Park (1992) examine congruity from both supply-side forces (e.g., manufacture skills, physical features) and demand-side forces (e.g., needs satisfied, usage situations) forces. As mentioned earlier, in 1994, Broniarczyk and Alba distinguish between category similarity and brand association similarity, and much subsequent work examines the congruity between the *parent brand* and the extension product or category (the earlier work starting with Aaker and Keller 1990 considered the congruity between the *parent category* and the extension category). In 2005, Martin, Stewart, and Matta investigate congruity from the goal hierarchy perceptive, specifically, whether the extension product is congruent with the parent brand's attribute-level, benefit-level, and value-level goals. All of this research significantly improved our understanding of the formation/causes of congruity.

Table 2. 2. Formative measures of congruity in brand extensions.

Research Paper	Type of Research	Measure Used
Aaker and Keller (1990). JM	Cor	(1) The extent to which the products were substitutes that they would select between in certain usage situations (SUBSTITUTE); (2) the extent to which the products were complements that they would be likely to use together in certain usage situations (COMPLEMENT); (3) Would the people, facilities, and skills used in developing, refining, and making the original product be helpful if the manufacturer were to make the product extension? (TRANSFER).
Park, Milberg, and Lawson (1991). JCR	Exp	The level of similarity on the feature similarity and concept consistency.
Keller and Aaker (1992). JMR	Exp	Overall similarity (1 = not at all similar, ..., 7 = very similar); how helpful the people, facilities, and skills used in developing, refining, and making the first product would be if the manufacturer were to make the second product (Aaker and Keller 1990).
Smith and Park (1992). JMR	Cor	How similar they believed the focal product was to each of the other products affiliated with the brand in terms of (1) the type of needs they satisfy, (2) the situation in which they are used, (3) skills required to manufacture them, and (4) their physical features.
Sunde and Brodie (1993). IJRM	Cor	(1) TRANSFER, the usefulness of manufacturing skills and resources in the original product class for making the extension product; (2) SUBSTITUTE, substitutability of the original and extension products in use ; (3) COMPLEMENT, complementarily of the original and extension product classes in use (Aaker and Keller 1990)
Broniarczyk and Alba (1994). JMR	Exp	(1) 9-point scale used in the pretest that ranged from "not similar" to "very similar". (2) the relevance of the brand associations in each of the potential extension categories on a 9-point scale ranging from "not at all relevant" to "very relevant."
Martin, Stewart, and Matta (2005). JAMS	Exp	(1) Feature-based perceived similarity measures: (a) Overall Perceived Similarity: "How similar\typical is Benetton leather shoes and Benetton clothing?"; (b) Manufacturing Similarity: "What is the ability of Benetton to manufacture and produce dress leather shoes,clothing?" (2) Usage Similarity:(a) "How similar are Reebok athletic shoes and Reebok dress leather shoes in terms of how\when they are used?" (b) "How likely are you to use Reebok athletic shoes and Reebok dress leather shoes together?" (c) "How appropriate is it to use Reebok athletic wear to exercise?" (3) Goal-derived categorization measures: Goodness-of-fit: (a) "How well does Benetton's leather shoes fit with the goal of wanting high quality, colorful clothing?" (b) "How consistent is Benetton's dress leather shoes with the goal of wanting high quality, colorful clothing?" (c) "How well does Benetton's dress leather shoes exemplify the goal of wanting high quality, colorful clothing?"
Echambadi, Arroniz, Reinartz, and Lee (2006). IJRM	Cor	Three fit variables: (1) transferability, (2) complementarity and (3) substitutability, (Aaker and Keller 1990)
Kalamas, Cleveland, Laroche, and Laufer (2006). JSM	Exp	7-point scaled items: fit, substitute, complement, and transfer (Aaker and Keller 1990)
Pina, Martinez Salinas, de Chernatony, and Drury (2006). EJM	Cor	(1) The company's perceived capacity to offer the extension (transfer fit), (2) the possibility of jointly using the extension and the current services (complement fit) . (Aaker and Keller 1990)
Völckner and Sattler (2006). JM	Cor	Perceived fit: (1) the overall similarity of the brand extension to the parent brand, (2) Would the people, facilities, and skills used in making the original product be helpful if the manufacturer were to make the extension product? (Aaker and Keller 1990), and (3) the relevance of the brand-specific associations in the extension product category (Broniarczyk and Alba 1994).

Table 2.2. Formative measures of congruity in brand extensions. (continued)

Research Paper	Type of Research	Measure Used
Völckner and Sattler (2007) IJRM	Cor	(1) Global similarity: global similarity between the parent brand and the extension product (Boush and Loken 1991); (2) Brand concept consistency: [Brand name] and [extension product] have similar images (Bhat and Reddy 2001); (3) Relevance of the extended associations for the extension product: extent that parent brand-specific associations are relevant in the extension category (Broniarczyk & Alba 1994); (4) Symbolic value of the parent brand (i.e., image orientation of extended information): respondents indicate the extent that associations such as exclusivity or a particular brand image can be transferred to the extension product (Reddy et al. 1994); (5) Linkage of the utility of the parent brand to product attributes of the original product category: [Brand name] is closely tied to the attributes of the original product category (Rangaswamy, Burke, and Oliva 1993)
Ahluwalia (2008). JMR	Exp	Pretest: the level of similarity between each proposed extension product (1) the product categories, (2) attributes, (3) usage situations, and (4) target market associated with the parent brand of the extension; Main study: the perceived fit of the new product with the brand ("similar/dissimilar," "inconsistent/consistent"),
Oakley, Duhachek, Balachander, and Sriram (2008). JCR	Exp	How similar is the extension to the parent brand product (1) in terms of consumer needs being satisfied and (2) in terms of skills required for production.
Voelckner, Sattler, and Kaufmann (2008). ML	Cor	Five 7-point scales anchored by high/low global similarity, high/low similarity of the extension to products the brand currently makes, very logical/not at all logical for the company, high/low perceived ability of the company to make a product in the extension product class, and appropriate/inappropriate for the parent brand.
Martínez Salinas and Pina Pérez (2009). JBR	Cor	(1) Category fit :(a)The extension is similar to the brand's products, (b) The firm's resources are helpful to make the product extension. (2) Image fit: (a) The product extension fits with the brand image, (b) Launching the extension is logical for the company, (c) Launching the extension is appropriate for the company
Sichtmann, Schoefer, Blut, and Kemp (2017). EJM	Cor	(1) How does the picture you have of [brand name] fit [extension product]? (2) In your opinion, how does the [extension product] fit with the other products and services that are offered by [brand name]? (3) Would the people, facilities and skills of [brand name] used to deliver the original service be helpful if the service provider were to offer the following products and services? (Völckner and Sattler 2006)

Note. Cor: correlational research; Exp: experimental research; JM: *Journal of Marketing;* JCR: *Journal of Consumer Research*; JMR: *Journal of Marketing Research*; IJRM: *International Journal of Research in Marketing*; JAMS: *Journal of Academy of Marketing Science*; JSM :*Journal of Strategic Marketing*; EJM: *European Journal of Marketing*; ML: *Marketing Letters*; JBR: *Journal of Business Research*.

But despite the value of this work, there is much overlap among the aspects of congruity proposed by different researchers. For example, the transfer dimension by Aaker and Keller (1990), which measures "perceived ability of any firm operating in the first product class to make a product in the second product class" (p. 30), is highly similar to the skills dimension by Smith and Park (1992), which refers to the skills required for the parent brand to manufacture the extension product. Furthermore, it is not clear how the different aspects of congruity fit together.

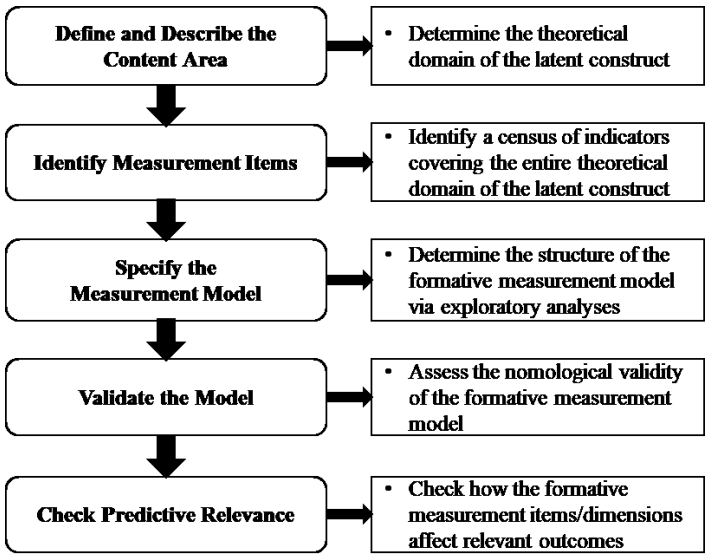
And there is no research that has systematically sorted through these various possible formative aspects of congruity (which the current paper addresses).

From the construct measurement perspective, all the above aspects/dimensions of congruity can be viewed as formative items, which measure different possible determinants of the brand-extension fit. The current paper identifies these items, removes redundant items, and uses these items to develop a formative measurement model of congruity. We also examine the effect of the resulting formative dimensions of congruity on consumers' responses.

Methodology

Since the main objective of this paper is to develop a formative measurement scale for the congruity construct, the traditional Churchill (1977) paradigm for reflective measurement scales is not directly applicable. Substantial effort has been made to develop procedures for formative measurement scale development (Chin, 2010; Coltman et al., 2008; Diamantopoulos & Winklhofer, 2001). But as yet, there is no well-established and widely accepted approach or framework. To help organize the work of the current paper, we start by suggesting a simple high-level framework in figure 2.1. We provide this figure as a conceptual overview of our analysis.

Figure 2. 1. Formative scale development framework



Define and Describe the Content Area

The first step is to specify the theoretical domain of the latent construct. This step is critical because it drives the researcher's assessment of the content validity of the latent construct. Specifically, a conceptual definition of the latent construct should be clearly stated to indicate the focus of conceptual interest. An operational definition of the latent construct should also be provided that (a) indicates how the latent construct is measured, and, since we are engaged in formative scale development, (b) makes clear that the underlying dimensions and measurement items jointly constitute the latent construct.

Identify Measurement Items

The second step is to identify as many distinct formative items as possible to cover the theoretical domain (i.e., the content) of the latent construct. Since the formative measurement items directly form and define the latent construct, these items should cover the domain of the latent construct (as much as possible). Ideally, a census of the formative items should be identified. However, from an empirical perspective, as long as all the items conceptually represent what most scholars believe are the key facets of the domain of interest, they can be regarded as adequate (Rossiter, 2005).

Another key issue to be examined at this step is the causal direction between the formative items and the latent construct. In particular, temporal precedence, co-variance of the item and the construct, and plausible alternative explanations should all be considered.

Specify the Measurement Model

The third step is to explore and develop a specification for the internal structure of the formative measurement model. The items that are included should cover different theoretical parts of the domain of the latent construct; in other words, these formative items should not be

interchangeable. A reasonable way to arrive at a model specification is to examine the relationship among all the formative items through exploratory analyses, in terms of whether these items are theoretically and statistically distinct from each other or have overlap.

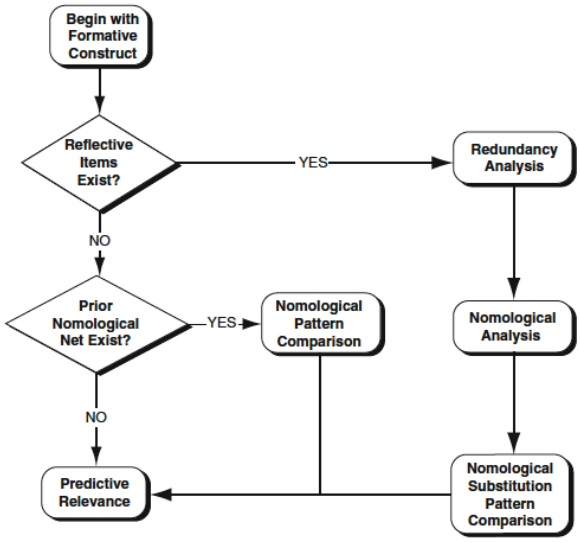
Different scenarios lead to different model specifications. (a) If all these formative items are statistically independent and theoretically different, it suggests that these items directly form the latent construct. Therefore, the measurement model should be specified as a first-order formative model. (b) If some of the formative items are statistically highly correlated and theoretically interchangeable, it suggests that these items could be reflective of the same theoretical domain of a latent construct. In other word, these items are reflective items of a common latent variable, and this latent variable is a formative dimension of the focal construct. Therefore, the measurement model should be specified as a first-order reflective, second-order formative model.

Validate the Model

The next step is to validate the formative measurement using empirical data. Due to the theoretical difference with reflective measures, the formative measures' validity cannot be evaluated by the traditional methods, such as factor loadings or average variance extracted. Instead, for the formative measure, its nomological validity, which refers to whether a construct behaves as it should within a system of related constructs (Liu, Li, & Zhu, 2012), should be assessed. A roadmap for formative construct validation proposed by Chin (2010) (figure 2.2) can provide useful guidance for this step. One can check for the convergent validity of the formative model with past reflective models through redundancy analysis (which we do shortly). One can compare the formative model with past (non-reflective) models (which we will not do, because we are not aware of past formative modeling effort in this area). And, if redundancy analysis is

used, one can consider the robustness of the model structure when removing various parts of the model, or substituting them with variations of those parts of the model (which we will do in the appendix 2.D). These various checks (and also that of the next subsection) establish nomological validity (that the model behaves as it should in as many ways as we can think of).

Figure 2. 2. Formative construct validation roadmap (by Chin 2010)



Check Predictive Relevance

Our last step (and the last step of figure 2.2) is to use the validated measurement model to predict relevant outcomes. Particularly interesting in this step is to use path analysis to understand how the formative components of the focal construct influence the outcomes of interest. We will use the model to predict consumer attitudes toward brand extensions. In the next section, we follow the methodology outlined in figure 2.1.

Analysis & Results

1. Define the Construct Content

For this paper, we use the following conceptual definition of congruity: *the similarity, consistency, or fit between the parent brand and the extension product*. The reason for our interest in this concept is that the literature on brand extensions supports the hypothesis that a

higher level of congruity between a parent brand and its extension category leads to more favorable consumer attitudes toward the extension product. Our operational definition of congruity elaborates on the conceptual definition, as follows: Congruity is consumers' overall perception of similarity or fit between the parent brand and the extension product, *formed by consumers' separate evaluations (of brand-extension fit) on relevant underlying dimensions or aspects of the brand extension*. From a measurement perspective, this operational definition emphasizes that we will use measurement items from consumers that identify the underlying dimensions or aspects of congruity. From a methodological perspective, this operational definition emphasizes that that underling dimensions (and measurement items) individually influence and jointly constitute the latent congruity construct (consumers' overall perception of the congruity between the parent brand and the extension product).

2. Identify Measurement Items

To help identify items that cover the theoretical domain of the congruity construct, we conducted a comprehensive review of the measurement items for congruity explicitly used or suggested in the brand extension literature. Table 2.D.1 of the technical appendix 2.D shows all the items used in the literature as the various authors named and operationalized them. It is apparent upon inspection that some of these items measure the same thing, and some measures are reflective of overall congruity. After grouping together the various items that measure the same thing, we focused on the potentially formative items. We identify six distinct formative items, shown in table 2.3 below, that we suggest cover the various formative measures used in the literature.

Table 2. 3. Six formative items of congruity.

Formative Items of Congruity	Definition the conguity, fit or consistency between …	Sources	Example
Feature-based	the specific features or attributes of the parent brand's main product and those of the extension product	Park, Milberg, and Lawson (1991); Smith and Park (1992); Martin, Stewart, and Matta (2005); Völckner and Sattler (2007); Ahluwalia (2008).	McDonald's - Frozen french fries
Resource-based	the resource required (e.g.,people, faicilities, skills) to develop and manufacture the parent brand's main product and the extension product	Aaker and Keller (1990; 1992); Smith and Park (1992); Echambadi, Arroniz, Reinartz, and Lee (2006); Kalamas, Cleveland, Laroche, and Laufer (2006); Pina, Martinez Salinas, de Chernatony, and Drury (2006); Völckner and Sattler (2006); Oakley, Duhachek, Balachander, and Sriram (2008); Voelckner, Sattler, and Kaufmann (2008); Sichtmann, Schoefer, Blut, and Kemp (2017).	Nike -Treadmills
Image-based	the abstract associations coming up in consumers' mind when thinking of the parent brand's main product and the extension product	Park, Milberg, and Lawson (1991); Broniarczyk and Alba (1994); Völckner and Sattler (2006; 2007); Martínez Salinas and Pina Pérez (2009); Sichtmann, Schoefer, Blut, and Kemp (2017).	BMW - Motorboats
Usage-occasion-based	the usage occasion of the parent brand's main product and that of the extension product	Aaker and Keller (1990; 1992); Smith and Park (1992); Martin, Stewart, and Matta (2005); Echambadi, Arroniz, Reinartz, and Lee (2006); Kalamas, Cleveland, Laroche, and Laufer (2006); Pina, Martinez Salinas, de Chernatony, and Drury (2006); Völckner and Sattler (2006); Ahluwalia (2008).	Westjet Airline - Suitcases
Function-based	the basic function of the parent brand's main product and that of the extension product	Aaker and Keller (1990); Smith and Park (1992); Sunde and Brodie (1993); Martin, Stewart, and Matta (2005); Echambadi, Arroniz, Reinartz, and Lee (2006); Kalamas, Cleveland, Laroche, and Laufer (2006); Pina, Martinez Salinas, de Chernatony, and Drury (2006); Völckner and Sattler (2007); Oakley, Duhachek, Balachander, and Sriram (2008).	Johnson-Johnson - Stuffed toys
Target-market-based	the target market of the parent brand's main product and that of the extension product	Ahluwalia (2008).	Cheerios - Waffles

This form of human-supervised consolidation of items involves subjective assessments. To be transparent, table 2.D.1 of the technical appendix 2.D explicitly shows our matching between the measurement items used in the literature and the six distinct formative measurement items that we list in table 2.3 above. It is worth emphasizing that we primarily rely on the past literature for generating candidate measurement items (rather than consulting various content experts individually or in groups). Our justification is that the literature is very large (tables 2.1 and 2.2 include many papers), and we are relying on these many authors as content experts (who themselves have devoted considerable effort to gain insight into the topic).

Subsequent to identifying the six items of table 2.3, a focus group (with ten students from a North American University) was then conducted for two purposes: (a) to evaluate the comprehensiveness of our list of formative items of congruity; and (b) to confirm that the formative items were perceived to have a causal effect on the overall latent congruity construct.

Other potential formative items covering distinct theoretical domains of the congruity construct were suggested; and some definitions in table 2.3 were modified to cover several of the suggestions. The focus group did agree that the formative items in table 2.3 should have a causal effect on the congruity construct.

Discussion. There are two key theoretical considerations that should be into taken account when developing measurement items for a formative latent construct. First, a causal relationship between the measurement item and the latent construct is required. These six formative items identified here, according to our literature review and focus group, appear to satisfy the causality direction criterion. Second, the formative measurement items collectively define the latent construct. Specifically, the measurement items do not need to share a common theme, and they are not interchangeable – adding and dropping any item will change the theoretical domain covered by the construct. In this step of the analysis, we endeavored to identify as many formative measurement items as possible to cover the theoretical domain of the congruity construct. In the next step of the process, we explore the internal structure of these six formative items of congruity with exploratory data analysis.

3. Specify the Measurement Model (Study 1: Exploratory Analysis)

To explore the internal structure of the six identified formative items of congruity, we began by compiling a list of brand extension manipulations in the literature. In particular, from an extensive review of the various papers, we compiled a total of 101 experimental manipulations of congruity in the literature (see appendix 2.A). Each manipulation involved a comparison of two different brand extensions (the papers were asking whether the change in congruity from one extension to another extension implied a change in some outcome variable).

We then constructed a dataset as follows. Two researchers independently coded these 101 brand extension manipulations in terms whether any, all, or some combination of the six formative congruity items in table 2.3 had been involved in the manipulation. For a particular congruity item, a value of 1 was coded if the manipulation increased this aspect of congruity, -1 was coded if the manipulation decreased this aspect of congruity, and 0 was coded if the manipulation did not change this aspect of congruity. We then added together the two coders' assessments. As a result, we generated six variables, each measured on a 5-point scale ranging from -2 to 2, where the larger positive (negative) value represents a stronger manipulation of congruity (incongruity) for the associated measurement item (see appendix 2.A for details). Then, we analyzed the data with three exploratory techniques: correlation analysis, cluster analysis, and exploratory factor analysis.

Correlation analysis. For a formative construct, there is no assumption about correlation (or lack thereof) among the measurement items. Nevertheless, the correlation results shed some light on the measurement model specification. We noted that some of these formative items are highly correlated with each other. For example, the correlation between feature and function is as high as .657. Theoretically, the high correlations may suggest that these six formative items have some overlap regarding their coverage of the theoretical domain of the latent construct.

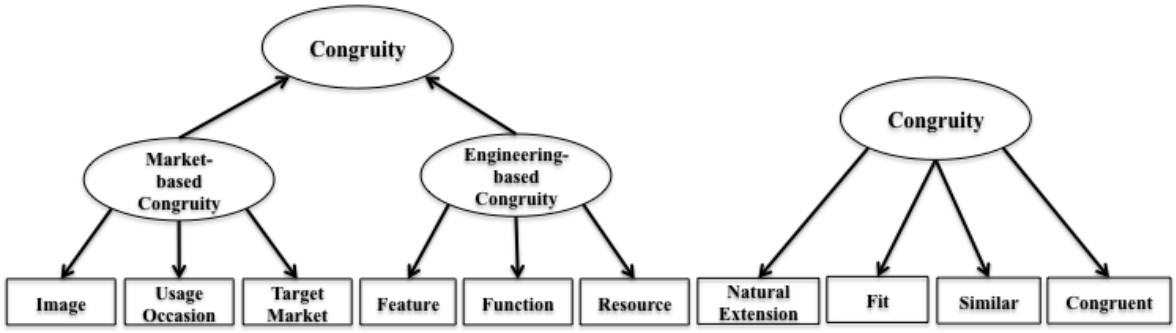
Cluster analysis. To visualize the correlation relationships among these variables, we carried out a hierarchical cluster analysis (average linkage) with the distance metric, $d = (1 - \widehat{\rho_{ij}})/2$, where $\widehat{\rho_{ij}}$ is the measured correlation between congruity item i and j. The cluster analysis very clearly showed that target-market, image, and usage-occasion-based congruity are grouped in one cluster, while function, feature, and resource-based congruity are grouped in another cluster (see technical appendix 2.D.2).

Factor analysis. Principal component factor analysis was then conducted in order to corroborate and quantify the results of the cluster analysis. Similar to the cluster analysis, the factor analysis results (see technical appendix 2.D.3) show that these six formative items of congruity are loaded on two distinct factors, which accounts for 75.81% of the total variance. The first factor, which we name as *market-based congruity*, includes image-based, target-market-based and usage-occasion congruity. The second factor, which we name as the *engineering-based congruity*, includes feature-based, function-based, and resource-based congruity.

Discussion. The exploratory analyses of study 1 provide insight into the internal structure of the six formative items of congruity. First, the correlation analysis suggests that these six items overlap in terms of their coverage of theoretical domain of the latent construct. Second, both cluster analysis and factor analysis suggest that these six items clearly fall into two groups, which we will henceforward refer to as *market-based congruity* and *engineering-based congruity*. Third, within each group, the measurement items are highly correlated with each other, which suggests that these measurement items may be reflective of the same intermediate latent construct. Fourth, these two groups are also statistically different from each other in factor analyses. It is worth mentioning that these two exploratory techniques are not constrained to produce meaningful or easily interpretable results, but they do. The former group of items pertains to how people perceive a brand (image-based, target-market-based, and usage-occasion-based congruity), whereas the latter group of variables pertains to aspects of technical similarity of the products under the same brand (function-based, feature-based, and resource-based congruity). The former latent dimension more resides in humans and their perceptions, and the latter latent dimension more resides in the products, themselves.

To sum up, the exploratory analyses suggest that the six items of congruity we identified fall neatly into groups that may constitute two underlying dimensions: *market-based* and *engineering-based congruity*. Therefore we propose that the internal structure of the six formative congruity items we identified can be described by the model of figure 2.3-a. Formally, market-based and engineering-based congruity constitute two formative dimensions of the higher-order latent construct, congruity. Using the terminology of structural equations modeling, figure 2.3-a is a first-order reflective, second-order formative measurement model – because image, usage-occasion, and target-market congruity are reflective of *market-based congruity*, and feature, function, and resource-based congruity are reflective of *engineering-based congruity*. Figure 2.3-b describes a measurement model with four reflective items of overall congruity (these four items were often used in the literature). We will subsequently use the reflective model as a check for the congruity construct coming from the formative model.

Figure 2. 3. Formative vs. reflective measurement models of congruity.



(a) First-order reflective, second-order formative measurement model

(b) First-order reflective measurement model

4. Validate the Model (Studies 2-a and 2-b)

In order to calibrate and validate the model specification of figure 2.3-a, we carried out two further studies. These studies asked consumers to evaluate brand extensions for fictional and real brands, respectively.

Study Design & Data Collection. In study 2-a, we selected and developed 36 brand extensions stimuli from previous literature that used fictional parent brands. Specifically, 12 fictional brands across different product categories, each of which extends to three different extension categories, were compiled (see appendix 2.B). We then recruited 199 Mturk participants. Each participant evaluated six brand extension stimuli (selected randomly out of the 36 brand extensions), in terms of their overall attitude toward the brand extension (one-item seven-point bi-polar scale: dislike- like), their perception on the six formative items of congruity (see the appendix 2.B for the measurement items used), and their perceptions of four reflective items of congruity (natural extension, fit, similar, and congruent – each measured with a seven-point Likert scale). After data cleaning (18 incomplete data points were deleted), 1,176 data points were generated.

Study 2-b used a similar study design. From previous literature, we compiled 36 brand extensions stimuli using real parent brands (see appendix 2.C). We then recruited 194 Mturk participants. Each participant randomly evaluated six brand extension stimuli (out of these 36 brand extensions) on the same measures of study 2-a. After data cleaning (40 incomplete data points were deleted), 1,124 data points were generated.

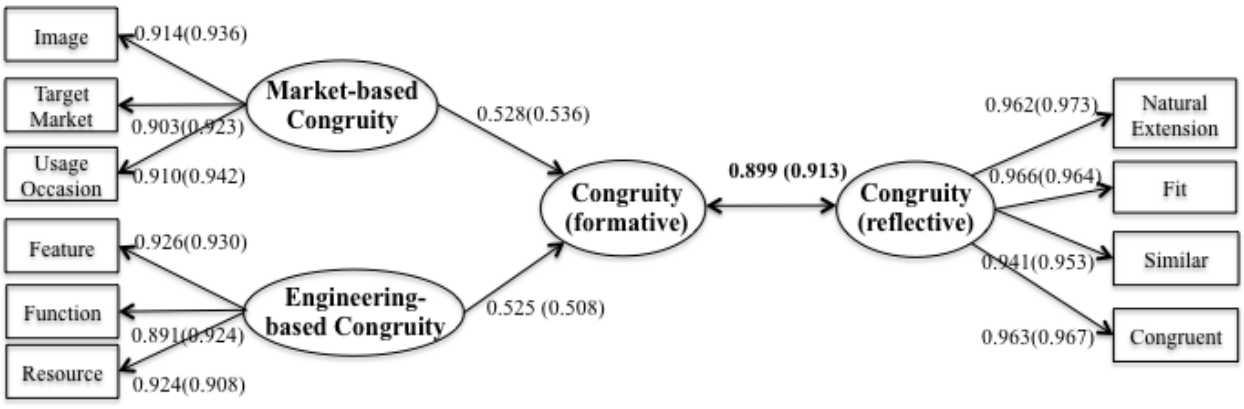
For both studies 2-a and 2-b, note that each study used 36 brand extensions from the previous literature (a subset of the brand extension manipulations from the previous literature, described in study 1). We considered coming up with completely different brand extension stimuli, but we believe that the benefits of well-tested stimuli outweigh the disadvantage of some overlap of the stimuli with the exploratory analysis. Furthermore, unlike most of the brand extension stimuli in the literature, which are described in a bipolar and relative fashion (high versus low congruity), we restricted our attention to stimuli pretested to have three levels of

congruity; so in our stimuli, we have 12 cases that were pretested to be congruent extensions, 12 cases of moderate incongruity, and 12 cases of extreme incongruity. In this way, we are able to construct a sample of brand extensions covering an approximately equally distributed range of congruity levels for each of our two studies.

Data Analysis Method: Structural Equation Modeling with Partial Least Squares (SEM-PLS). The major reason for the use of SEM-PLS is that we are examining and testing a formative measurement model of congruity (figure 2.3-a). The traditional covariance-based SEM (CB-SEM) has a key assumption that the measurement items used to measure a latent construct are reflective in nature. Estimating a formative model (such as that of figure 2.3-a) using CB-SEM introduces measurement model misspecification, which can lead to substantial estimation bias (Jarvis et al., 2003). We used the SmartPLS software for the data analysis in this research.

Figure 2.3-b specifies a reflective model for the same intended congruity construct as our formative model of figure 2.3-a. The reflective model will be used to check the validity of the formative measurement model.

Figure 2. 4. PLS results for a redundancy model (Studies 2-a & 2-b)



Note: first number is estimate from study 2-a (number in parenthesis is estimate from study 2-b).

Redundancy Model Results. Figure 2.4 presents estimates of a combined structure including both the formative and reflective models of figure 2.3-a and 2.3-b showing results for

both studies 2-a and 2-b (estimates for study 2-b are shown in parentheses). In this study, the inclusion of reflective (figure 2.3-b) and formative (figure 2.3-a) models permits a check on convergent validity.

From the perspective of interpretation, the reflective part of the model may tell us that participants see a brand extension as congruent, similar, a good fit, or a natural extension, but it does not tell us the reasons such congruity may be achieved; the formative part of the model directly addresses this issue by tracing congruity back to market-based or engineering-based causes (and going further back to the particular six items at the left of figure 2.4).

From the perspective of methodology, this form of model is known as a redundancy model because it measures both reflective and formative versions of the same intended construct (in this case Congruity), together with estimation of the correlation between these two estimated related latent constructs (congruity estimated from a reflective structure and congruity estimated from a formative structure). The rule of thumb is that a correlation of .80 or above suggests an adequate convergence between these two models. A correlation of .90 or above would indicate an extremely strong result (Chin, 2010).

Details of the estimation are provided in the technical appendix 2.D.4; in particular, table 2.D.3 provides the complete estimation results for both the formative and reflective parts of the model. All the coefficients are significant at the .001 level for both studies 2-a and 2-b. Table 2.D.4 provides details of various correlations, including the correlation between the formative and reflective congruity constructs.

We observe that the models fit very well in both studies. For the items reflective of the three latent variables (i.e., market-based congruity, engineering-based congruity, and reflective-measured congruity), the loadings of their measurement items in both studies are all around .9,

and significant at the .001 level. As for the reliability of all these reflective variables, their composite reliability scores are all above .93 (see table 2.D.3 in technical appendix 2.D.4). Regarding convergent validity, the high loadings of their reflective measurement items suggest that the items are descriptive of the associated latent variables. As for the discriminant validity, for any two latent variables, the average variance extracted (AVE) scores of both variables are greater than the squared correlation between these two variables (see table 2.D.4 in technical appendix 2.D.4), which suggests good discriminant validity among these latent constructs.

The formative measurement model of congruity shows that market-based congruity and engineering-based congruity have significant impact on the overall congruity. This suggests that the overall congruity construct is formed by these two latent constructs (themselves measured by three reflective items each). Furthermore, the weights for market-based and engineering-based congruity are similar across both studies, which is indicative of remarkable inter-sample reliability of the measurement process, especially recognizing that all the stimuli and the survey participants are different for the two samples. Moreover, from a substantive perspective, we observe that market-based and engineering-based congruity constructs have similar weights (in the .508 to .536 range). Each of these two dimensions appears about equally important for consumers' overall impression of congruity for brand extensions.

Lastly, for both studies 2-a and 2-b, we now observe that the correlations linking the formative and reflective measurement models of congruity are .899 and .913 respectively. This indicates a very strong convergence between the formative and reflective models. In other words, the formative congruity construct that we focus on this paper, composed of market-based and engineering-based components, does largely account for participants sense that an extension is

"congruent" and "a fit" (and that the brand and the extension product have "similarities" and are a "natural extension"). Overall, this is very strong support for the model specified in figure 2.3.

Nomological Analysis. In the brand extension literature, abundant empirical research consistently shows that the congruity between the parent brand and the extension product has a significant and positive effect on consumers' overall attitude of the brand extension. Therefore, we develop three structural models (in technical appendix 2.D.5) to help evaluate the nomological validity of the formative measure of congruity. Model (1) includes the reflective measure of congruity only and examines the link between reflective-measured congruity and consumers' overall attitude toward brand extensions (*reflective-only model* shown as figure 2.E.1). Model (2) adds the formative measure of congruity to model (1) as an antecedent of the reflective measure of congruity (*full model* shown as figure 2.E.2). Model (3) replaces the reflective measure of congruity in model (1) with the formative measure (*formative-only model* shown as figure 2.E.3).

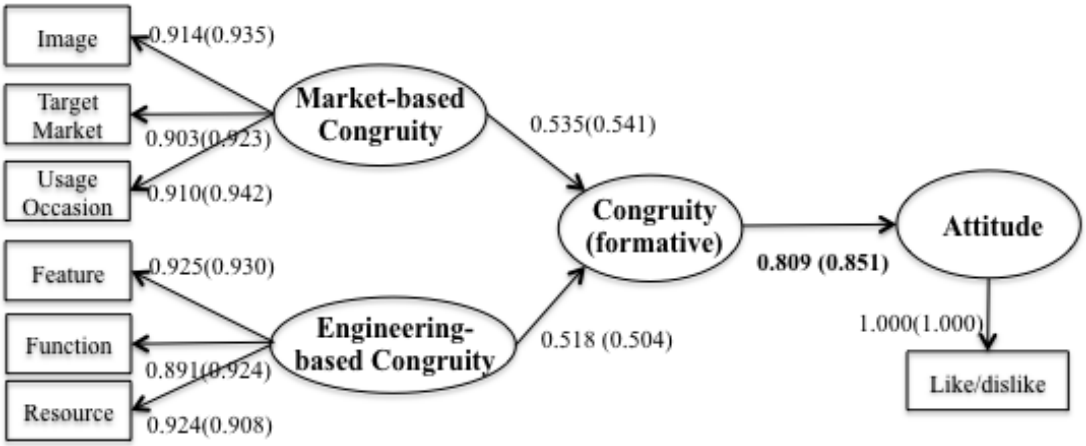
Several findings, via the model comparisons, are identified to support the nomological validity of the formative measure of congruity (see technical appendix 2.D.5). First, the formative measure of congruity works similarly to the reflective measure in affecting the key outcome variable, consumer attitudes toward brand extensions (model 1 vs. 3). Second, the formative measure has the additional advantage of affording predictions from changes in engineering-based versus market-based congruity (model 1 vs. 2). Third, the reflective measure of congruity is a full mediator of the effect of the formative measure of congruity on the brand extension attitude (model 2 vs. 3).

5. Check Predictive Relevance

In figure 2.5, we examine whether our formative model of congruity does a good job of predicting overall consumer attitude toward brand extensions (measured by a seven-point bi-polar "like-dislike" scale), which is a critical outcome variable in our two studies.

We confirm that congruity, as measured by our model, does indeed have a positive impact on consumer attitudes toward brand extensions, with path coefficients of .809 and .851 (both significant at the .001 level) for studies 2-a and 2-b, respectively. This is consistent with the literature, and provides further evidence of nomological validity of our formative model of congruity.

Figure 2. 5. Predictive relevance of the congruity model (Effect of congruity on attitude toward brand extension).



Note: first number is estimate from study 2-a (number in parenthesis is estimate from study 2-b).

It is worth noting the robustness of the other coefficient estimates in the analysis of figure 2.5, as compared to the estimates in figure 2.4. All the estimates are very similar (item weights around .9, significant at the .001 level, and path coefficients from the latent variables (market-based and engineering-based congruity) all just above .5 (and again all significant at the .001 level).

Lastly, we consider the model predictions afforded by changing the two formative dimensions, market-based and engineering-based congruity. Table 2.4 shows that the indirect effects of these two formative dimensions are respectively .433 and .419 (.460 and .428) for the formative-only model in study 2-a (study 2-b) (which are similar values to those of the full model).

Table 2. 4. Path analysis and model comparison (Study 2-a & 2-b).

Path	Study 2-a			Study 2-b		
	Reflective Model	Full Model	Formative Model	Reflective Model	Full Model	Formative Model
Direct Path						
Congruity ➔ Attitude	.893***	.893***	.809***	.903***	.903***	.851***
R square	.798	.797	.654	.816	.816	.723
GoF	.872	.866	.839	.887	.803	.869
Indirect Path						
Market Based Congruity ➔ Attitude	-	.423	.433	-	.442	.460
Engineering Based Congruity ➔ Attitude	-	.421	.419	-	.419	.428

Note. ***: p <0.001

Although this is a very simple result, it has important theoretical and practical meaning. Having a new extension product fit with a company's typical products in terms of their engineering characteristics is not enough – the company needs also to build from its goodwill with existing customer segments. However, even if the latter is true that an extension is targeted to the same customer segment, that customer segment should also believe that the engineering strengths of the company can be transferred to the extension product. Thus, the conclusion of this research is that two conditions promote successful brand extensions, particularly if they are both present: congruity on both engineering-based and market-based dimensions.

Conclusions & Discussion

Conclusions

This paper explores and identifies the key dimensions of congruity in the context of brand extensions, and develops a formative measurement scale of congruity (the BEF scale, short for Brand Extension Fit scale). Toward this end, we review the theoretical domain and the

measurement items used in the brand extension literature, and we identify six distinct items. We code the congruity manipulations in the experimental brand-extension literature in terms of the presence of each of these six items. We conduct exploratory analyses on this data (correlations, cluster analysis, exploratory factor analysis) to suggest a model specification. Our resulting specification decomposes the overall congruity construct into two dimensions: *engineering-based* and *market-based congruity*. The *engineering-based congruity* dimension is measured by the three items, feature-based, function-based, and resource-based congruity. The *market-based congruity* dimension is measured by the three items image-based, usage-occasion-based, and target-market-based congruity.

We validate our measurement model of congruity using SEM-PLS with two separate datasets obtained from general consumers consisting of their judgments about extensions of (a) fictitious parent brands and (b) real parent brands. Redundancy analysis for both samples confirms that the formative measures of congruity converge very well to the common reflective measures of congruity for brand extensions. Nomological validity is confirmed for broader models with and without the reflective model component. The results not only support the measurement model of congruity suggested by our exploratory analyses, but also provide evidence of excellent sample-resample reliability for fictional and real brands. Lastly, we found that our validated model of congruity is a good predictor of consumers' overall attitudes toward brand extensions, which is a key outcome of congruity. We observe equal relative weights (each close to a half) for market-based and engineering-based congruity concerning their impact on customers' overall attitudes toward brand extensions. This provides insight for managerial purposes.

Discussion

This paper contributes to the marketing and brand extension literatures theoretically, strategically, methodologically, and practically.

From a theoretical perspective, this paper adds to our understanding of brand-extension congruity as a multi-dimensional construct. In particular, via multiple datasets and various analysis methods, we consistently find two key dimensions of congruity: engineering-based and market-based congruity. These dimensions cover various different theoretical domains of the congruity construct discussed by past researchers in the literature.

From a strategic perspective, we find that the two formative dimensions of congruity have nearly equal positive effects on consumers' overall attitudes toward a brand extension, across various product categories. Therefore, both dimensions are important conditions for a successful brand extension. Brand managers must be aware that people generally respond to both market-based and engineering-based congruity in their brand extensions, in roughly equal measure, and that the most easily accepted extensions would have both forms of congruity.

From a methodological perspective, this paper also advanced a framework for formative measurement modeling (which we carry out in our analysis). This framework may be useful for other researchers doing formative scale development.

From a managerial perspective, we return to the scenario introduced in the introduction of this paper faced by Jaydon, a brand manager of a coffeehouse chain. Jaydon can use the BEF scale proposed in this paper (see appendix 2.B) to help with the following practical issues:

(a) *Stimulate extension ideas:* The proposed BEF scale provides direction for Jaydon to think of categories with similar functions (e.g., energy drink), features (e.g., coffee soda), resource-requirements (e.g., packaged coffee for supermarkets), usage occasions (e.g., cookies),

image (e.g., tea), and target market (e.g., French press coffee makers) – or, more generally, to consider new categories of products congruent with the engineering attributes of the coffee shop or the market-based features surrounding the brand.

(b) *Select from alternatives:* The proposed BEF scale can help Jaydon to give precedence to extension ideas high on measurement items (from the proposed six) that build on the parent brand company's desired value proposition.

(c) *Develop a marketing plan around the selected alternative:* The proposed BEF scale can help Jayden identify which key extension dimensions (and applicable items) to communicate in the campaign message and to recognize other aspects that may need further development to cover both engineering-based and market-based congruity. For example, suppose Jaydon's company decides to launch a brand extension product of fresh-baked cookies. Our formative measurement scale can show Jaydon that this extension product addresses existing customers' needs and usage-occasions to sell this product, but their customers may be concerned about whether it has the resource and ability to produce tasty cookies. This information can help Jaydon realize that he needs to put more emphasize on erasing any concerns of product quality in their marketing campaign.

Future Research

Overall, the literature has arrived at a good understanding of congruity theory since 1990. The current paper can help to operationalize the importance of congruity for brand extensions through development of a formative measurement scale of congruity (the BEF scale). In the process of developing this scale, we now recognize two robust dimensions of congruity. Nevertheless, there remain at least two important avenues for future research.

First, it would be desirable to get a sense of whether the various moderators studied in the literature work together with either engineering-based or market-based congruity, or both. It would be particularly useful to know how to use the moderators to enhance congruity on one dimension, when the other dimension is not present or present to a smaller degree (or to enhance the perceived congruity of the weaker dimension).

Second, congruity has been considered in many application areas in marketing, including product designs, brand alliances, celebrity endorsement, event sponsorship, and cause-related marketing. A natural extension of the present research would be to ascertain the dimensions of congruity and to explore developing similar scales for these other application domains. Dimensions of congruity analogous to engineering-based and market-based congruity might be applicable to some contexts (e.g., new product designs), while completely different dimensions of congruity might be applicable for other contexts. Brand extensions represent the largest application area of congruity research, thus far, but these other application areas are already very managerially relevant and are destined to grow in importance.

Appendix 2.A.

Data Generation of Study 1

Table 2. A. 1. Brand extension stimuli from previous literature.

No	Parent brand	Parent brand-category	Congruent Extension	Moderate Incongruent Extension	Extremely Incongruent Extension	References
1	McDonald's	Fast food		Frozen French fries for home cooking	Online community	Bambauer-Sachse, Huttl, & Gierl (2011).
2	Puma	Sports goods		Jeans	Camera	Buil, de Chernatony, & Hem (2009).
3	-	Sneaker	Shorts	Sunglasses		Fedorikhin, Park, & Thomson (2008)
4	-	Cell phones	MP3 players	DVD players		
5	McDonald's	Fast food	Frozen French fries		Pralines	
6	McDonald's	Fast food	Frozen French fries	Roasted coffee		
7	McDonald's	Fast food		Roasted coffee	Pralines	Gierl & Huettl (2011)
8	Katjes	Fruit gum	Jam		Tissues	
9	Haribo	Fruit gum	Jam		Tissues	
10	Apple	Electronic goods		Digital picture frames	Hair dryer	
11	Consul (fictitious brand)	MP3 players	Digital radio		Clothing	Hagtvedt & Patrick (2008)
12	Croscill	Soap dispenser	Towels		Cheese	
13	Xerox	Photo copying machines	Photo processing	Modems		
14	Xerox	Photo copying machines	Photo processing		Wristwatches	Kalamas, Cleveland, Laroche, & Laufer (2006).
15	Xerox	Photo copying machines		Modems	Wristwatches	
16	Kleenex	Tissues	Toilet paper			
17	Listerine	Mouthwash	Toothpaste			
18	Kodak	Imaging products	Camcorder			
19	Sony	Electronic goods		Laser printers		
20	Starbucks	Coffee		Ice cream		Kapoor & Heslop (2009)
21	Dell	Electronic goods		TVs		
22	Kleenex	Tissues	Toilet paper			
23	Listerine	Mouthwash	Toothpaste			
24	Dell	Electronic goods		TVs		
25	Nike/New balance	Sports goods	Comfort insoles	Treadmills		
26	MTV	Music	music downloading	Travel agency		Kim & John (2008)
27	New balance	Sports goods	Comfort insoles	Treadmills		
28	Haagen-Dazs	Ice cream	Chocolate gift		Popcorn	Liu & Hu (2012).
29	Yili	Ice cream	Chocolate drink		Ice tea	
30	BMW	Cars	Motorboat		Camera	
31	BMW	Cars	Motorboat	Lawnmower		Maoz & Tybout (2002)
32	BMW	Cars		Lawnmower	Camera	
33	BMW	Cars	Motorboat	Lawnmower		
34	Benetton	Clothing	Dress leather shoes		Cotton spandex athletic wear	
35	Reebok	Sports goods	Cotton spandex athletic wear		Dress leather shoes	Martin, Stewart, & Matta (2005)
36	Benetton	Clothing	Dress leather shoes		Cotton spandex athletic wear	
37	Reebok	Sports goods	Cotton spandex athletic wear		Dress leather shoes	
38	Cheerios	Cereals	Granola bars		Frozen dinner	
39	Guess	Clothing	Watch		USB	Mathur, Jain, & Maheswaran (2012).
40	Timberland	Outdoor sports goods	Tent		Sports drinks	
41	Optix	-	Camera		Sports goods	McCarthy, Heath, & Milberg (2001)
42	Sony	Electronic goods	Camera		Sports goods	
43	Nike	Sports goods			Camera	
44	Nike	Sports goods	Camping equipment		Camera	Milberg, Goodstein, Sinn, Cuneo, & Epstein (2013).
45	Nikon	Camera	Scanner		CD player	
46	-	Digital camera	Digital Camcorder		Breakfast cereals	
47	-	Digital camera	Digital Camcorder	TV		Nan (2006)
48	-	Digital camera		TV	Breakfast cereals	
49	CNN	News network	Weekly news magazine	Movie channel		Yeo & Park (2006)
50	Guess	Clothing	Casual suite	Ski gear		

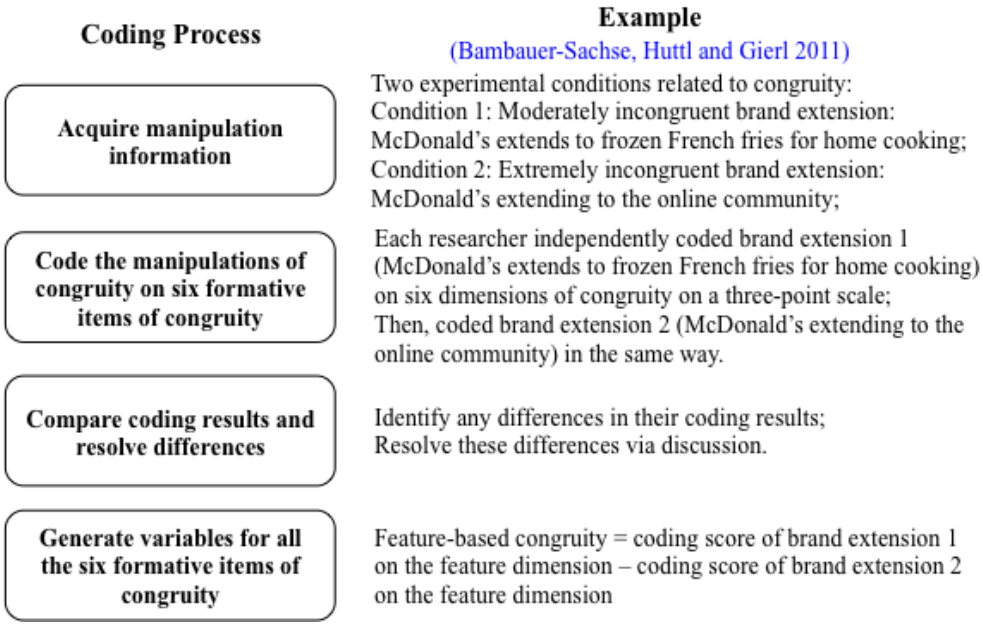
Table 2.A.1. Brand extension stimuli from previous literature. (continued)

No	Parent brand	Parent brand-category	Congruent Extension	Moderate Incongruent Extension	Extremely Incongruent Extension	References
51	Rolex	Wristwatch	Sunglasses		Handbag	
52	Rolex	Wristwatch	Sunglasses	Fountain pens		Shen, Bei, & Chu (2011)
53	Rolex	Wristwatch		Fountain pens	Handbag	
54	Johnson&Johnson	Baby lotion	Skin care lotion	Stuffed toys		
55	Johnson&Johnson	Baby lotion	Skin care lotion		Instant noodles	Ahluwalia (2008)
56	Johnson&Johnson	Baby lotion		Stuffed toys	Instant noodles	
57	Band A	TV	26-in. stereo-adaptable television	A microwave oven		
58	Band A	TV	26-in. stereo-adaptable television		Snow skis	Barone (2005)
59	Band A	TV		A microwave oven	Snow skis	
60	Electronic company A	Electronic goods	26-in. stereo-adaptable television	A microwave oven		
61	Electronic company A	Electronic goods	26-in. stereo-adaptable television		Snow skis	Barone, Miniard, & Romeo (2000)
62	Electronic company A	Electronic goods		A microwave oven	Snow skis	
63	Coppertone	Sunscreen	A magazine on how to properly apply skin lotion	A magazine on the benefits of Vitamin D		Noseworthy (2014)
64	Coppertone	Sunscreen	A magazine on how to properly apply skin lotion		A magazine on how to prepare steak	
65	Michelin	Car tires	Bicycle tires		Sports scandals	Lane (2000).
66	Nikon	Camera	Scanner		CD players	Milberg, Sinn, & Goodstein (2010)
67	Nike	Sports goods	Camping equipment		Camera	
68	Kodak	Camera		Greeting cards	Shoes	
69	McDonald's	Fast food		Omelet	Razor	Monga & John (2007)
70	McDonald's	Fast food	Onion rings		Razor	
71	McDonald's	Fast food	Onion rings	Omelet		
72	-	Partially baked pizza	Pizza Hut	Frito Lay		Oakley, Duhachek, Balachander & Sriram (2008)
73	Timex	Watches	Stopwatch		Smoke detector	Park, Milberg, & Lawson (1991)
74	Timex	Watches	Grandfather clock		Cologne	
75	Xerox	Photo copying machines	Digital Camera		A snowboard	Shine, Park, & Wyer (2007).
76	Pepsi	Cola	Vitamin-fortified cola	Sodium-free orange juice		Sood & Keller (2012).
77	Lufthansa	Airline		Suitcases	Backpacks	
78	Alaska Airlines	Airline		Suitcases	Backpacks	
79	Japan Airlines	Airline		Suitcases	Backpacks	Yeung & Wyer Jr (2005)
80	Air China	Airline		Suitcases	Backpacks	
81	Panasonic	Electronic goods		A computer mouse	A leather wallet	
82	Coca cola	Soft drinks		Ice tea	Toffee	
83	Wrigley's	Gum		Toffee	Ice tea	Zhang & Sood (2002)
84	Coca cola	Soft drinks	Ice tea	Toffee		
85	Wrigley's	Gum	Toffee	Ice tea		
86	Gleem	Toothpaste	Dental floss		Soft drinks	Morrin (1999).
87	-	Electronic products	Camera	Refrigerator		
88	-	Electronic products	Camera		Ballpoint pen	Boush & Loken (1991)
89	-	Electronic products		Refrigerator	Ballpoint pen	
90	Cheerios	Cereal	Oatmeal	Pastry		
91	Froot loops	Cereal	Oatmeal	Pastry		
92	Froot loops	Cereal	Hot cereal	Lollipop		Broniarczyk & Alba (1994).
93	Cheerios	Cereal	Hot cereal	Lollipop		
94	Apple	Computers	Machine reader	Mainframe		
95	Compad	Computers	Machine reader	Mainframe		
96	-	Potato chip	Cheese crackers	Cookies		Keller & Aaker (1992).
97	-	Potato chip	Cheese crackers		Ice cream	
98	Samsung	Electronic products		Smart watch	Power compact	Huang, Jia, & Wyer (2017).
99	Calbee	Snakes		Apple chips	Orange soda	
100	Subaru	cars		Jet ski	Motorized scooter	Yorkston, Nunes, & Matta (2010)
101	Tiffinay	jewlery		high-heel shoes	work flats	

Note. "-": no parent brand.

Two researchers independently coded the 101 pairs of brand extension stimuli recorded from previous experimental research in brand extensions (see table 2.A.1 above) on these six formative items of congruity according to the following procedures explained in the figure 2.A.1.

Figure 2. A. 1. Coding process for the six formative items of congruity.



(1) One researcher acquired the brand extension manipulation information of congruity from 36 experimental papers. When congruity is manipulated in experiments, usually two brand extensions are developed to represent two levels of congruity. Therefore, information about pairs of manipulation information was acquired. For example, in the paper by Bambauer-Sachse, Huttl and Gierl (2011), two levels of congruity were manipulated. A moderately incongruent brand extension is McDonald's extending to frozen French fries for home cooking, and an extremely incongruent brand extension is McDonald's extending to the online community.

(2) For each pair of brand extensions, two researchers independently coded each of the two manipulations of congruity on the six formative items of congruity. In the case of Bambauer-Sachse, Huttl and Gierl's (2011) paper, each researcher coded the two brand extension cases

separately (McDonald's extends to frozen French fries for home cooking; McDonald's extends to online community) on the six formative items of congruity (feature, resource, image, usage-occasion, function, and target-market-based congruity) on a three-point scale (-1: nothing similar between the parent brand and the extension category; 1: the parent brand and the extension category are almost the same; 0: intermediate level of similarity between the parent brand and the extension category).

(3) After the two coders independently finished all the coding, they compared their coding results across all seven marketing application contexts (their initial percentage of agreement is over 75% on average). Then, they resolved their differences via discussion.

(4) Based on the agreed coding results, six variables were generated, and each variable represents one formative item of congruity. For instance, in the context of brand extensions, we generated six variables: feature, resource, image, usage-occasion, function, and target-market-based congruity. Each variable is the difference between the coding results of the two manipulation levels of congruity on that formative item of congruity. For example, the variable called feature-based congruity in brand extensions is calculated as the difference score between the agreed coding results of two brand extension cases (e.g. McDonald's extends to frozen French fries for home cooking vs. McDonald's extends to online community) on the feature dimension of congruity. As a result, each generated variable ranges from -2 to 2, and the bigger value represents more successful and stronger manipulation of congruity on this formative item of congruity.

Appendix 2.B.

Six Formative Measurement Items

Table 2. B. 1. Six formative measurement items used in study 2-a and 2-b.

Formative Items of Congruity	**Measurement** (1: extremely dissimilar; 7: extremely similar)
Feature-based	How similar are [parent brand] and [extension product], in terms of their specific features and attributes (e.g. size, colour, smell, taste, etc.)?
Function-based	How similar are [parent brand] and [extension product], in terms of their basic functions?
Resource-based	How similar are [parent brand] and [extension product], in terms of the resources required to develop the products?
Image-based	How similar are [parent brand] and [extension product], in terms of their abstract images and concepts (i.e., associations, concepts or images that come to you mind when you think about the product)?
Usage-occasion-based	How similar are [parent brand] and [extension product], in terms of their usage occasions (i.e., where or when to ues them)?
Target-market-based	How similar are [parent brand] and [extension product], in terms of their target markets (i.e., consumers at which a product is aimed)?

Appendix 2.C.

Research Stimuli Used in Study 2-A and Study 2-B

Table 2. C. 1. Research stimuli used in study 2-a and study 2-b.

Study 2-a			Study 2-b			
No	**Parent brand category**	**Extension category**	**No**	**Parent brand**	**Parent brand category**	**Extension category**
1	Sneakers	BBQ grills	1	Alaska Airline		Suitcases
2		Shorts	2		Airlines	Flight socks
3		Sunglasses	3			Running shoes
4	Juice	Beer	4	Benetton		Dresses
5		Soft drink	5		Clothing	Leather shoes
6		Coffee	6			Cotton spandex athletic wear
7	Cell phones	Guitars	7	BMW		Motorboat
8		Speakers	8		Automobile	Lawnmower
9		Tablets	9			Camera
10	Airlines	Flight socks	10	Cheerios		Frozen dinner
11		Running shoes	11		Breakfast cereal	Waffles
12		Suitcases	12			Lollipops
13	Automobile	Camera	13	Guess		Watches
14		Lawnmower	14		Clothing	USB
15		Motorboat	15			Ski gear
16	Breakfast cereal	Frozen dinner	16	Haagen-Dazs		Chocolate gift
17		Lollipops	17		Ice cream	Popcorn
18		Waffles	18			Ice tea
19	Digital camera	Breakfast cereal	19	Johnson&Johnson	Consumer packaged goods	Skin care lotion
20		Digital camcorder	20			Stuffed toy
21		Televisions	21			Instant noodles
22	Fast food	Frozen French fries for home cooking	22	McDonald's	Fast food	Frozen French fries for home cooking
23		Pralines	23			Roasted coffee
24		Roasted coffee	24			Pralines
25	Ice cream	Chocolate gift	25	Nike	Sports wear and goods	Comfort insoles
26		Ice tea	26			Treadmills
27		Popcorn	27			Camera
28	Luxury watches	Fountain pens	28	Nikon		Cereal
29		Handbags	29		Camera	Camcorder
30		Sunglasses	30			TV
31	Photocopy machines	Digital photo printers	31	Rolex		Sunglasses
32		Modems	32		Luxury watches	Fountain pens
33		Wristwatches	33			Handbags
34	Potato chips	Cheese crackers	34	Xerox	Photocopy machines	Modems
35		Cookies	35			Wristwatches
36		Ice cream	36			Digital photo printers

Appendix 2.D.

Technical Details of Data Analyses

D.1. Identification of Six Formative Measurement Items

Table 2.D.1 shows how we arrived at these six measurement items in table 2.3 by matching between references of the items/measures used by the various authors in this large literature and the final six measurement items in table 2.3. For example, Aaker and Keller (1990) utilize three measures of fit, as follows: "COMPLEMENT, indicates the extent to which consumers view two product classes as complements . . . SUBSTITUTE, is the extent to which consumers view two product classes as substitutes . . . TRANSFER reflects the perceived ability of any firm operating in the first product class to make a product in the second product class" (p. 30). Our item "usage occasion-based" essentially covers their item COMPLEMENT; our item "function-based" covers their item SUBSTITUTE; and our item "resource-based" covers their item TRANSFER. We proceeded in this fashion through the papers in the literature, matching their items with the six items in table 2.3.

Table 2. D. 1. Formative measures of congruity in brand extensions.

Research Paper	Type of Research	Type of Measure	Measure Used	Dimensionality
Aaker and Keller (1990). JM	Cor	F	(1) The extent to which the products were substitutes that they would select between in certain usage situations (SUBSTITUTE); (2) the extent to which the products were complements that they would be likely to use together in certain usage situations (COMPLEMENT); (3) Would the people, facilities, and skills used in developing, refining, and making the original product be helpful if the manufacturer were to make the product extension? (TRANSFER).	Usage occasion, resource, function
Park, Milberg, and Lawson (1991). JCR	Exp	R & F	The level of similarity on the feature similarity and concept consistency.	Feature, image
Keller and Aaker (1992). JMR	Exp	R & F	Overall similarity (1 = not at all similar, ..., 7 = very similar); how helpful the people, facilities, and skills used in developing, refining, and making the first product would be if the manufacturer were to make the second product (Aaker and Keller 1990).	Resource
Smith and Park (1992). JMR	Cor	F	How similar they believed the focal product was to each of the other products affiliated with the brand in terms of (1) the type of needs they satisfy, (2) the situation in which they are used, (3) skills required to manufacture them, and (4) their physical features.	Function, usage occasion, resource, feature
Sunde and Brodie (1993). IJRM	Cor	F	(1) TRANSFER, the usefulness of manufacturing skills and resources in the original product class for making the extension product; (2) SUBSTITUTE, substitutability of the original and extension products in use ; (3) COMPLEMENT, complementarily of the original and extension product classes in use (Aaker and Keller 1990)	Resource, usage occasion, function
Broniarczyk and Alba (1994). JMR	Exp	F, R	(1) 9-point scale used in the pretest that ranged from "not similar" to "very similar". (2) the relevance of the brand associations in each of the potential extension categories on a 9-point scale ranging from "not at all relevant" to "very relevant."	Image
Martin, Stewart, and Matta (2005). JAMS	Exp	F, R	(1) Feature-based perceived similarity measures: (a) Overall Perceived Similarity: "How similar\typical is Benetton leather shoes and Benetton clothing?"; (b) Manufacturing Similarity: "What is the ability of Benetton to manufacture and produce dress leather shoes,clothing?" (2) Usage Similarity:(a) "How similar are Reebok athletic shoes and Reebok dress leather shoes in terms of how\when they are used?" (b) "How likely are you to use Reebok athletic shoes and Reebok dress leather shoes together?" (c) "How appropriate is it to use Reebok athletic wear to exercise?" (3) Goal-derived categorization measures: Goodness-of-fit: (a) "How well does Benetton's leather shoes fit with the goal of wanting high quality, colorful clothing?" (b) "How consistent is Benetton's dress leather shoes with the goal of wanting high quality, colorful clothing?" (c) "How well does Benetton's dress leather shoes exemplify the goal of wanting high quality, colorful clothing?"	Feature, usage occasion, function
Echambadi, Arroniz, Reinartz, and Lee (2006). IJRM	Cor	F	Three fit variables: (1) transferability, (2) complementarity and (3) substitutability, (Aaker and Keller 1990)	Function, usage occasion, resource
Kalamas, Cleveland, Laroche, and Laufer (2006). JSM	Exp	R, F	7-point scaled items: fit, substitute, complement, and transfer (Aaker and Keller 1990)	Function, usage occasion, resource
Pina, Martinez Salinas, de Chernatony, and Drury (2006). EJM	Cor	F	(1) The company's perceived capacity to offer the extension (transfer fit), (2) the possibility of jointly using the extension and the current services (complement fit) . (Aaker and Keller 1990)	Resource, usage occasion
Völckner and Sattler (2006). JM	Cor	F, R	Perceived fit: (1) the overall similarity of the brand extension to the parent brand, (2) Would the people, facilities, and skills used in making the original product be helpful if the manufacturer were to make the extension product? (Aaker and Keller 1990), and (3) the relevance of the brand-specific associations in the extension product category (Broniarczyk and Alba 1994).	Resource, image
Völckner and Sattler (2007) IJRM	Cor	F, R	(1) Global similarity: global similarity between the parent brand and the extension product (Boush and Loken 1991); (2) Brand concept consistency: [Brand name] and [extension product] have similar images (Bhat and Reddy 2001); (3) Relevance of the extended associations for the extension product: extent that parent brand-specific associations are relevant in the extension category (Broniarczyk & Alba 1994); (4) Symbolic value of the parent brand (i.e., image orientation of extended information): respondents indicate the extent that associations such as exclusivity or a particular brand image can be transferred to the extension product (Reddy et al. 1994); (5) Linkage of the utility of the parent brand to product attributes of the original product category: [Brand name] is closely tied to the attributes of the original product category (Rangaswamy, Burke, and Oliva 1993)	Function, image, attribute

Table 2.D.1. Formative measures of congruity in brand extensions. (continued)

Research Paper	Type of Research	Type of Measure	Measure Used	Dimensionality
Ahluwalia (2008). JMR	Exp	R, F	Pretest: the level of similarity between each proposed extension product (1) the product categories, (2) attributes, (3) usage situations, and (4) target market associated with the parent brand of the extension; Main study: the perceived fit of the new product with the brand ("similar/dissimilar," "inconsistent/consistent"),	Attribute, usage occasion, target market
Oakley, Duhachek, Balachander, and Sriram (2008). JCR	Exp	F	How similar is the extension to the parent brand product (1) in terms of consumer needs being satisfied and (2) in terms of skills required for production.	Function, resource
Voelckner, Sattler, and Kaufmann (2008). ML	Cor	F, R	Five 7-point scales anchored by high/low global similarity, high/low similarity of the extension to products the brand currently makes, very logical/not at all logical for the company, high/low perceived ability of the company to make a product in the extension product class, and appropriate/inappropriate for the parent brand.	Resource
Martínez Salinas and Pina Pérez (2009). JBR	Cor	F, R	(1) Category fit :(a)The extension is similar to the brand's products, (b) The firm's resources are helpful to make the product extension. (2) Image fit: (a) The product extension fits with the brand image, (b) Launching the extension is logical for the company, (c) Launching the extension is appropriate for the company	Image
Sichtmann, Schoefer, Blut, and Kemp (2017). EJM	Cor	F, R	(1) How does the picture you have of [brand name] fit [extension product]? (2) In your opinion, how does the [extension product] fit with the other products and services that are offered by [brand name]? (3) Would the people, facilities and skills of [brand name] used to deliver the original service be helpful if the service provider were to offer the following products and services? (Völckner and Sattler 2006)	Resource, image

Note. Cor: correlational research; Exp: experimental researc; F: formative; R: reflective.

D.2. Cluster Analysis

Figure 2.D.1 shows the dendogram of the cluster analysis, which is based on average linkage (complete linkage was nearly identical).

Figure 2. D. 1. Cluster analysis.

D3. Factor Analysis

Table 2.D.2 below shows the detailed results of the factor analysis (using Varimax rotation method) in study 1, including variance explained, reliability, and factor loadings.

Table 2. D. 2. Exploratory factor analysis.

Latent and Observed Variable	% of Variance	Reliability	Factor Loading
Market-based congruity	39.811	.839	
Target-market-based			.911
Image-based			.841
Usage-occasion-based			.779
Engineering-based congruity	35.999	.783	
Feature-based			.940
Function-based			.775
Resource-base			.701

Note. Extraction method: Principal Component Analysis. Rotation method: Varimax. Reliability test: Cronbach's alpha

D4. Model Estimation Results of Figure 2.5

Table 2.D.3 below shows the detailed estimates of the results of the model summarized in figure 2.5.

Table 2. D. 3. Model results (study 2-a & 2-b).

Latent and Observed Variable	Study2-a			Study 2-b		
	Loadings/Weights[a]	AVE	CR	Loadings/Weights[a]	AVE	CR
First order construct						
Market-based congruity		.827	.934		.871	.953
Target-market-based	.903***			.923***		
Image-based	.914***			.935***		
Usage-occasion-based	.910***			.942***		
Engineering-based congruity		.834	.938		.847	.943
Feature-based	.925***			.930***		
Function-based	.891***			.924***		
Resource-base	.924***			.908***		
Congruity (reflective)		.912	.977		.930	.981
Natural extension	.962***			.973***		
Fit	.953***			.964***		
Similar	.941***			.953***		
Congruent	.963***			.967***		
Second order construct						
Congruity (formative)						
Market-based congruity	.527***			.534***		
Engineering-based congruity	.526***			.511***		

Note. AVE: Average Variance Extracted; CR: Composite Reliability; a: standardized weight; ***: $p<0.001$

Table 2.D.4 below shows the detailed estimates of important correlations for several variables for the model summarized in figure 2.5.

Table 2. D. 4. Inter-construct correlation (study 2-a & 2-b).

Interconstruct Correlation	Market-based Congruity		Engineering-based Congruity		Congruity (reflective)		Congruity (formative)	
	Study 2-a	Study 2-b	Study 2-a	Study 2-b	Study 2-a	Study 2-b	Study 2-a	Study 2-b
Market-based Congruity	1.000	1.000						
Engineering-based Congruity	.802	.833	1.000	1.000				
Congruity (reflective)	.871	.895	.835	.955	1.000	1.000		
Congruity (formative)	.949	.959	.949	.852	.899	.913	1.000	1.000

D5. Nomological Analysis (Studies 2-a and 2-b)

In the brand extension literature, abundant empirical research consistently shows that the congruity between the parent brand and the extension product has a significant and positive effect on consumers' overall attitude toward the brand extension. Therefore, we develop three structural models to help evaluate the nomological validity of the formative measure of congruity. Model (1) includes the reflective measure of congruity only and examines the link between reflective-measured congruity and consumers' overall attitude toward brand extensions (*reflective-only model*). Model (2) adds the formative measure of congruity to model (1) as an antecedent of the reflective measure of congruity (*full model*). Model (3) replaces the reflective measure of congruity in model (1) with the formative measure (*formative-only model*). The goal is to show that the formative congruity measure behaves similarly to the reflective measure as one compares models (1) and (2), models (1) and (3), and models (2) and (3). The factor loadings and weights for all three models are presented in table 2.D.5; and graphical representations and estimated results of the three models are presented in figures 2.D.2, 2.D.3, and 2.D.4, respectively.

Table 2. D. 5. Factor loadings and weights of measurement models.

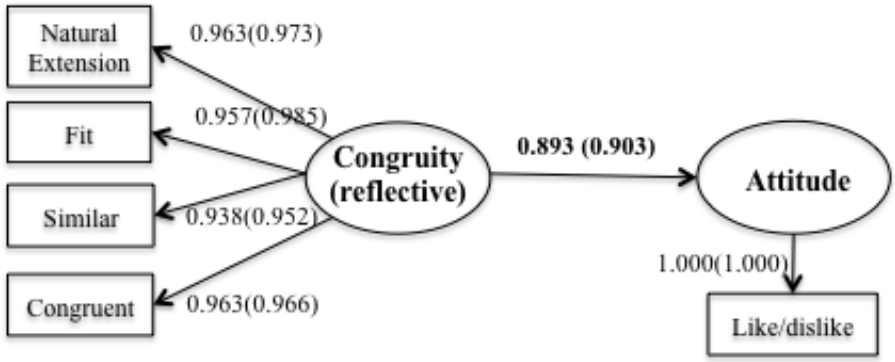
Latent and Observed Variable	Study2-a			Study 2-b		
	Reflective	Full Model	Formative	Reflective	Full Model	Formative
First order construct						
Market-based congruity						
Target-market-based		.903***	.903***		.923***	.923***
Image-based		.914***	.914***		.935***	.935***
Usage-occasion-based		.910***	.910***		.942***	.942***
Engineering-based congruity						
Feature-based		.925***	.925***		.930***	.930***
Function-based		.891***	.891***		.924***	.924***
Resource-base		.924***	.924***		.908***	.908***
Congruity (reflective)						
Natural extension	.963***	.962***		.973***	.973***	
Fit	.957***	.953***		.965***	.964***	
Similar	.938***	.941***		.952***	.953***	
Congruent	.963***	.963***		.966***	.967***	
Second order construct						
Congruity (formative)						
Market-based congruity		.527***	.531***		.534***	.536***
Engineering-based congruity		.526***	.522***		.511***	.508***

Note. ***: $p<0.001$

Reflective-only and full models. By comparing these two models, we note several things. First, the measurement models of the reflective only model and the full model fit very well (see table 2.D.5). Second, coefficients of corresponding structural paths are similar between both models (figure 2.D.2 and 2.D.3). Third, both structural models (figure 2.D.2 and 2.D.3) have similarly good model fit (study 2-a: $GoF_{Reflective} = .872$, $GoF_{Full} = .866$; study 2-b: $GoF_{Reflective} = .887$, $GoF_{Full} = .803$). Fourth, in both models, almost the same amount of variance in the key dependent variable (i.e., attitude toward the brand extension) is explained (study 2-a: $R^2_{Relective} = .798$, $R^2_{Full} = .797$; study 2-b: $R^2_{Relective} = .816$, $R^2_{Full} = .816$). Fifth, consistent with the redundancy analysis, we find that the formative measure of congruity converges well to the reflective measure. Sixth, most importantly, although the reflective measure of congruity can alone explain the dependent variable well, the full model with the formative measure of

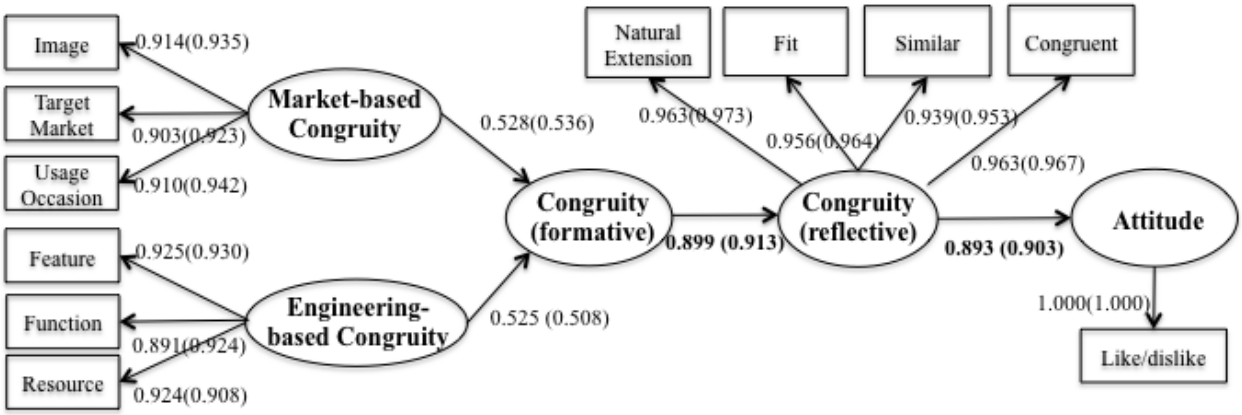
congruity distinguishes predictions between the individual effects of market-based and engineering congruity.

Figure 2. D. 2. Effect of congruity on brand extension attitude (Reflective only – model 1).



Note: first number is estimate from study 2-a (number in parenthesis is estimate from study 2-b).

Figure 2. D. 3. Effect of congruity on brand extension attitude (Full model – model 2).

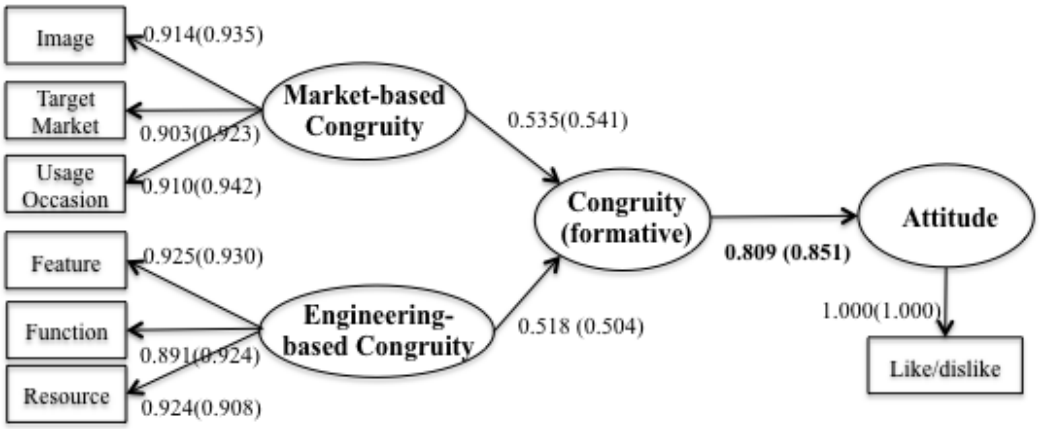


Note: first number is estimate from study 2-a (number in parenthesis is estimate from study 2-b).

Reflective-only and formative-only Models. Now, by comparing these two models, we note similar points: the measurement models fit well (see table 2.D.5) with similar corresponding coefficients and weights across study 2-a and 2-b; both models have satisfactory model fit (study 2-a: $GoF_{Reflective} = .872$, $GoF_{Formative} = .839$; study 2-b: $GoF_{Reflective} = .887$, $GoF_{Formative} = .869$); and similar amount of variance of the overall brand extension attitude explained by the formative and the reflective congruity constructs (study 2-a: $R^2_{Relective} = .798$, $R^2_{Formative} = .654$; study 2-b: $R^2_{Relective} = .816$, $R^2_{Formative} = .723$). These results show that the formative measure of congruity

works similarly as the reflective measure to affect the key dependent variable; and the formative measure has the additional advantage of affording predictions from changes in engineering-based versus market-based congruity.

Figure 2. D. 4. Effect of congruity on brand extension attitude (Formative only – model 3).



Note: first number is estimate from study 2-a (number in parenthesis is estimate from study 2-b).

Full and formative-only models. The comparison between the formative-only (figure 2.D.4) and the full model (figure 2.D.3) leads to some new findings. Although the formative measure of congruity can predict the dependent variable as well as the reflective measure, the full model including both formative and reflective measures of congruity works best in terms of predictive power (study 2-a: R^2_{Full}= .798, $R^2_{\text{Formative}}$= .654; study 2-b: R^2_{Full}= .816, $R^2_{\text{Formative}}$= .723).

In addition, we also found that, in the full model, the reflective measure of congruity is a full mediator of the effect of the formative measure of congruity on the brand extension attitude. Specifically, the path from the formative measure of congruity to attitude becomes insignificant (path coefficient = .033, *t* value= .937, $p > .05$), after the inclusion of the reflective measure.

This finding is consistent with the argument in the brand extension literature that in consumers' evaluation process, they will first evaluate different aspects or dimensions of congruity (i.e., formative measure of congruity), and then form an overall perception of

congruity (i.e., reflective measure of congruity). Furthermore, during the process of forming the overall perception of congruity, many other factors (e.g., moderators such as contextual factors and individual differences) also play a role. Therefore, the formative measure of congruity (i.e., market-based and engineering-based congruity) cannot explain all the variance in the reflective measure, and the reflective measure has a slightly better predictive power than the formative measure.

Discussion. Overall, across the two empirical studies (2-a and 2-b), several consistent findings are identified. First, the validity of the formative measurement model of congruity is supported via redundancy analysis and nomological analysis. Both analysis results show that the formative measure of congruity not only covers similar consumers' perception of congruity, but also have similar predictive power for the overall attitude toward brand extensions, compared to the reflective measure of congruity. Second, the formative measure of congruity can go beyond the reflective measure to show that its two formative dimensions (market-based and engineering-based congruity) have an equally important role in forming consumers' overall brand extension attitude. Third, the reflective measure of congruity serves as a key mediator of the effect of formative measure of congruity on consumers' overall brand extension attitude. Lastly, after comparing the results of study 2-a and study 2-b, whose key difference is that study 2-a uses fictional parent brands and study 2-b uses real parent brands in the stimuli, we can see that the results of study 2-b are slightly better than those of study 2-a. Our conjecture is that a brand extension idea with a real parent brand may be easier for consumers to assess.

Essay 3

What Makes Products Weird?

Antecedent to and Underlying Mechanism of Weirdness

Abstract

The word *weird* is widely used in our daily lives, but little is known about the concept of weirdness, or about what makes things weird, probably because of the negative connotations of weirdness. This paper challenges this view, not only by identifying the key antecedent to weirdness and the underlying mechanism in the context of product designs, but also by showing the marketing potential of weirdness. Via a series of experiments, this paper demonstrates that a key antecedent to weirdness is extreme incongruity between a product's design and its category schema (study 1), because of a failed sense-making process (study 2 & 3). Furthermore, providing information to facilitate the sense-making process can significantly decrease the perception of weirdness (study 4). Moreover, although extremely incongruent products are weird, and consumers like them less and are less willing to buy them, those same consumers are more willing to share information about these products than about regular ones (study 5). This paper also proposes a definition of weirdness and discusses the theoretical contributions, limitations, and practical implications of the research findings.

Keywords: Weirdness; Sense Making; Extreme Incongruity; Scale; Word-of-Mouth.

Introduction

North Americans might use the word *weird* to describe eating roasted sheep's head, common in Norway, or serving KFC (Kentucky Fried Chicken) as Christmas food, as done in Japan (Hopkins, 2017). Another example of weirdness is the injury suffered by Major League Baseball player Wade Boggs, who had to miss a week's worth of games after straining his back while trying to pull on his cowboy boots (Shaw, 2014). The Guinness World Records include many categories that could be called weird; for example, the current record for the longest kiss is 58 hours, 35 minutes, and 58 seconds, and the record for being covered by the largest number of bees is 1.1 million (Jacob, 2015). In politics, Donald Trump, the president of the United States, took a water break during his national security speech, which was regarded as weird and has gone viral in social media (Baragona, 2017).

The word *weird* is also often used by marketers and consumers to describe particular products, services, or experiences. For example, one may find and purchase many weird products on Amazon, such as a life-sized Bigfoot statue, live cockroaches, Nicolas Cage pillowcases, and a pocket-sized suture pad (Dickerson, 2016). A Prada store was built in the middle of nowhere, on an empty stretch of U.S. 90 outside of the tiny town of Valentine, some 150 miles from El Paso, and it has become a weird but popular tourist attraction (Dermody, 2016).

Given its popular usage in people's daily life and online, what is weirdness? According to the Oxford dictionary, *weird* has the meanings of "connected to fate," "something unearthly," and "strange and bizarre." However, there have been relatively few studies of weirdness in academia; some of these focus on bizarreness in psychology (McDaniel & Einstein, 1986; McDaniel, Einstein, & Lackey, 1989; McDaniel, Dornburg, & Guynn, 2005), weird and outrageous acts in entertainment (Sergius Koku, 1995), and bizarre experiences in marketing

(Latimer, 2015; Latimer & Raghubir, 2012). Yet, no clear definition of weirdness has so far been developed, and the methods of creating and manipulating weirdness vary among different fields of study and among different individual studies.

One may argue that the limited research on weirdness is probably due to the negativity associated with this construct. If consumers' responses toward weird products or services are negative, why bother to study weirdness? However, bizarre imagery, as compared to more usual imagery, has been shown to be able to increase memory recall (McDaniel & Einstein, 1986), weird and outrageous acts in entertainment industry have greater marketing and promotion value (Sergius Koku, 1995), and bizarre experiences are retrospectively perceived as more pleasant than mundane experiences (Latimer, 2015). Therefore, the construct of weirdness holds marketing potential, and is definitely worthy of further study. However, the more important question to answer in such a study is *what makes things weird*.

This research takes the first step to answer this question by examining what makes products weird in the context of product designs, specifically aiming to understand the key antecedent of weirdness and the related underlying mechanism. Based on the relevant literature in psychology, marketing, and product designs, we posit that extreme incongruity, which refers to extreme differences between a product design and its product category schema, is a key antecedent to the perception of weirdness. More specifically, when a product has certain features that are extremely atypical for its own product category, this product will be perceived as weird. This is because an extremely incongruent product design violates consumers' existing expectations for this kind of product, and consumers will then engage in a sense-making process in order to restore their own sense or state of the world (Weick, 1995). However, at the same time, the extremity of such a violation may be too hard for consumers to make sense of or

resolve. Therefore, the failure of the sense-making process leads to the perception of weirdness. Furthermore, if consumers can receive extra useful information to facilitate sense making, their perception of weirdness should be decreased.

This paper, via a series of experiments, uses various methods to manipulate congruity across many product categories, with different research contexts and samples, to demonstrate that extreme incongruity, compared to congruity and moderate incongruity, consistently leads to the perception of weirdness. Through direct measurement and indirect manipulation of relevant factors, we further demonstrate that the underlying mediator is a failed sense-making process. Finally, we extend this research framework by investigating some marketing outcome variables of extremely incongruent products and weirdness. Specifically, although extremely incongruent products are weird and produce less positive attitude and lower purchase intention than more conventional products, consumers' sharing intentions toward weird products are significantly higher than toward more regular ones.

Conceptual Framework

Weirdness

According to the Oxford dictionary, the word *weird* has a Germanic origin: the old English word *wyrd*, meaning *destiny*. In late Middle English, *weird*, as an adjective, meant *having the power to control destiny*. The word was most commonly associated with the Weird Sisters, which originally referred to the Fates, and later to the three witches in Shakespeare's play *Macbeth*. Because of Shakespeare's *Macbeth*, by the early nineteenth century *weird* took on the meaning of *unearthly*.

Google Ngram Viewer, which is an online search engine that charts frequencies of any set of comma-delimited search strings using a yearly count of n-grams found in sources printed

between 1500 and 2008, shows that since the beginning of the nineteenth century, the usage of *weird* started to increase significantly in the 1950s with a decline in the 1960s, and that since the 1970s, *weird* has been used more frequently.

The word *weird* has three common meanings, according to the Oxford dictionary: (1) suggesting something supernatural and unearthly; (2) connected to fate; and (3) very strange or bizarre (informal). The first two meanings can be traced back to the word's origins in Old and Middle English. The last meaning, *strange and bizarre,* is the most commonly used meaning in modern times. The word *weird* can be used to describe almost anything, such as objects, products, companies, brands, artworks, people, events, or places. At the same time, probably due to the increasing usage of *weird*, even in academia, research attention has been paid to this concept, as demonstrated by related constructs in various fields, including psychology, entertainment, and marketing.

The bizarreness effect and the distinctiveness process in psychology. In psychology, the most relevant research stream for the current research paper is the bizarreness effect or bizarre imagery effect, which can be traced back to the ancient Greeks, suggesting that bizarre images can improve memory (Lorayne & Lucas, 1974). A large number of studies in psychology since the 1980s have explored the mnemonic benefits of bizarre materials in order to enhance learning and reduce forgetting. The most famous examination of this effect is McDaniel and Einstein's 1986 study. Via a series of experiments, they demonstrated that bizarre imagery can increase memory recall, but only when bizarre imagery is presented together with common imagery (in a within-list/within-subject design), not when it is presented alone (in a between-list/between-subject design).

Since then, various efforts have been made to understand the reasons behind these effects. These studies have taken two main arguments: one involving information encoding, and the other involving information retrieval. Specifically, at the information encoding stage, bizarre items, presented with common items at the acquisition stage, induces more elaborative encoding that is more distinctive and thus more memorable than common items alone (Hunt & McDaniel, 1993; McDaniel, et al., 2005; Mulligan, 2000; Waddill & McDaniel, 1998). In the retrieval context, the bizarre items also have some distinctive features, which have advantages of being more easily recalled (Knoedler, Hellwig, & Neath, 1999; McDaniel et al., 2000).

However, these studies did not develop a clear definition of bizarreness, and instead, they focus more on the operationalization of bizarreness. Bizarre materials are "constructed by presenting common items in bizarre relations to one another, either through pictorial or sentential materials" (McDaniel & Einstein, 1986; McDaniel, et al., 2005). For example, the standard operationalization of bizarre stimuli is to vary the relations among the noun triplets in a sentence. For example, a bizarre context for the triplet items DOG, BICYCLE, and STREET is "The DOG rode the BICYCLE down the STREET," while the more common context is "The DOG chased the BICYCLE down the STREET" (Einstein et al., 1989).

The distinctiveness effect/process (Einstein & McDaniel, 1986; McDaniel, DeLosh, & Merritt, 2000) is highly related to the bizarreness effect. This process suggests that stimuli that are distinctive from other elements in one's focal context, such as other common stimuli presented together in a memorization task, can facilitate memory encoding and retrieval. The distinctiveness process, from a slightly different perspective, points out that one thing that makes an object bizarre is its distinctiveness from other elements in the focal context (i.e., other common stimuli presented together in the memory task).

Weird and outrageous acts in entertainment. In the entertainment field, artists' propensity to use bizarre or outrageous elements is an important part of the process. Sergius Koku (1995) defines the bizarre or the outrageous act in the entertainment industry as a "conduct – an act, an artistic expression or statement – that is out of the norms of the culture of a people" (p.19), and further found that using weird acts as a marketing promotion tool is "a double-edged sword: while it can attract some consumers, it also can turn off others" (p.29). However, in this research, weird and outrageous acts are operationalized as "creative activities."

This brings our attention to another related construct, creativity, which is usually defined as "a phenomenon whereby something new and somehow valuable is formed" (Mumford, 2003). In this sense, weird acts by artists can be understood as actions that diverge from social norms, while at the same time carrying marketing and production value. However, although weird things may be perceived as creative in certain situations or contexts, not all weird things are creative. One commonality between weirdness and creativity we can observe here is something new or divergent from the social norms. However, what factor differentiates weirdness from creativity is still an open question, and we will further discuss and examine the difference between weirdness and creativity in a later section of this paper.

Distinctive and bizarre experiences in marketing. Studies of bizarre and distinctive experiences in marketing mainly build upon and apply theories and findings from psychology into marketing. For example, in 2012, Latimer and Raghubir, using prototype theories in the categorization literature and the distinctiveness effect in psychology, examined the ability of retrieved exemplars and the presence of a distinctive peak to predict participants' overall retrospective evaluations of an experience immediately after and a day after that experience had occurred. Specifically, via field survey data, Latimer and Raghubir found that a distinctive peak

of an experience decreased overall evaluations immediately after the experience, supporting the prototype effect, but increased the overall evaluations after a delay, supporting the distinctiveness effect. In 2015, Latimer built on the mnemonic advantage of bizarre stimuli to investigate retrospective enjoyment of bizarre experiences as compared to mundane experiences. A series of experiments show that bizarre experiences are perceived as more pleasant in retrospect than they are during the initial experience. Although interesting, this research lacks a clear definition and theorization of the focal construct of bizarreness, leading to very different operationalization of bizarreness (i.e., paintings and sculptures pretested to be bizarre), which makes it hard to draw useful reference about weirdness in general.

To summarize, after reviewing relevant constructs and studies across various fields, one can notice that no clear definition of weirdness or bizarreness has been developed, and that the manipulation methods of weirdness or bizarreness vary greatly across different fields. This not only shows the research gap in the literature, but also calls for more research to clearly define and understand the weirdness construct.

At the same time, some useful themes can be identified. First, one commonality among bizarreness, distinctiveness, and creativity is that something in the focal stimuli (e.g., words, actions, or artistic works/actions) is different and distinctive from the other elements in the relevant environment/context (e.g., a word in a sentence, or a certain action performed by an entertainer). Second, bizarre and distinctive stimuli, as compared to mundane or regular stimuli, have memory advantages and can generate value in relevant contexts.

Extreme Incongruity in Product Designs

The focus of this paper is to understand what makes products weird. In the product design literature, one construct that is highly related to the common theme identified in other research

fields related to weirdness, is congruity, which refers to the similarity, typicality, or consistency between a product design and its product category schema in consumers' minds.

In the product design literature, the focal question relating to congruity is how congruity influences consumers' responses, primarily consumers' attitudes, acceptance, and purchase intentions. There are two streams of research on this question, with different theoretical grounds and predictions of outcomes.

The first of these research streams is the congruity hypothesis, which predicts a positive linear relationship between congruity and consumers' responses. In other words, a higher level of congruity leads to more positive responses. The consideration of congruity or fit in the marketing literature, in general, has its antecedents in the social psychology literature, going back to Gestalt psychology, which suggests that people prefer to perceive their environment in simple and coherent ways (Kohler, 1929). Various cognitive consistency theories have arisen from this perspective, including the strain toward symmetry (Newcomb, 1953), congruency theory (Osgood & Tannen-baum, 1955), and the affective-cognitive consistency model (Rosenberg, 1956). The presence of congruity can be associated with (a) mental comfort (Festinger, 1957), (b) a balance state (Heider, 1958), (c) ease of processing or categorizing different objects (Fiske & Pavelchak, 1986), and (d) affect or image transference via an established memory link (Anderson, 1983; Shimp, 1981; and, in marketing, Aaker & Keller, 1990; Smith, 2004).

The other stream of research on this topic involves the schema congruity effect, also known as Mandler's effect, whose central prediction is that moderate incongruity leads to more positive affective responses than congruity or extreme incongruity. This research stream relies on the connection between arousal and incongruity, and there may be a tradeoff between arousal arising from incongruity and aroused acceptance of congruity. Mandler's schema congruity

theory (1982), a classic in this field, suggests that although congruity can result in positive values, affective intensity is marginal. When slight incongruity occurs, the existing schema can incorporate or accommodate the new information without any major structural changes, and relatively better values can be expected. When incongruity is so severe that it cannot be easily assimilated, deeper structural changes to the existing schema are needed to accommodate the new information, and outcomes become uncertain, often negative, depending on the broader evaluation of the context. As a result, this theory proposes that moderate incongruity can generate more favorable responses than either congruity or severe incongruity, and congruity can lead to more favorable outcomes than incongruity. Much research has found empirical evidence supporting the inverted U-shaped relationship between congruity and consumers' responses (e.g., Meyers-Levy & Tybout, 1989; Noseworthy, Muro, & Murray, 2014; Peracchio & Tybout, 1996).

The key difference between the congruity hypothesis and the schema congruity effect is whether congruity or moderate incongruity would lead to optimal affective responses by consumers. One point of agreement between these two research streams is that, compared to congruity and moderate incongruity, extreme incongruity leads to the worst affective responses from consumers. Because of this, limited research attention is devoted to extreme incongruity.

Since the focus of this paper is to understand the concept of weirdness, we argue that among various levels of congruity, extreme incongruity, which means the focal product is extremely different from the general product category schema, has a greater tendency to lead to the perception of weirdness than congruity or moderate incongruity. Previous research in psychology and marketing shows that in order to be weird or bizarre, something has to be distinctively different from other elements in the relevant context. In the context of product designs, only extreme incongruent product designs are distinctively different from other products

in its own product category, and congruent or moderately incongruent product designs are not. Therefore, the first hypothesis of this research paper is as follows:

H1: Extremely incongruent product designs will be perceived as weirder than congruent and moderately incongruent designs.

Sense Making as the Underlying Mechanism

When one's current state of the world is different from his/her expected state of the world, or when there is no obvious way to engage the world, people tend to engage in explicit effort at sense making (Weick, Sutcliffe, & Obstfeld, 2005). Sense making refers to "meaning creation based on current and prior interpretations of thoughts generated from three sources: external stimuli, focused retrieval from internal memory, and seemingly random foci in working memory" (Woodside, 2001, p.416).

In the context of product designs, when consumers see a product design that is extremely different from their expectations, which arise from the product category schema invoked in their minds, consumers also tend to engage in a sense-making process, which is a conscious activity (Craig-Lees, 2001). Sense making is a broad concept, which can cover various cognitive activities. For example, in the product design literature, sense making can be simple assimilation, or accommodation (Piaget & Cook, 1952), depending on the severity of expectation violations. More specifically, for slightly or moderately incongruent product designs, consumers can easily assimilate these new products into their existing product category schema. For extremely incongruent product designs, consumers have to change their existing product category schema to accommodate these products. Given the extremity of the incongruity involved, most of the time, the outcome of this sense-making process would be unsuccessful. But if the sense-making process is successful, consumers can accommodate extreme incongruity (and it will be perceived

as less problematic and weird). Thus, we argue that a failed sense-making process mediates the effect of extreme incongruity on the perception of weirdness.

There is another plausible explanation for the effect of extreme incongruity on the perception of weirdness, which emphasizes on the role of negative affect. According to Mandler's theory (1982), unsuccessful accommodation of extreme incongruent products into an existing product category schema usually leads to some negative affect, such as anxiety, disgust, or helplessness, with a very high intensity level, and this negative affect could be the reason creating the perception of weirdness (we refer this possible explanation as the affect account).

Between the two possible underlying mechanisms, we argue that weirdness perception is more an outcome of people's cognitive sense-making process, and may be independent of peoples' affective responses. We empirically tested these two explanations (see study 3). To summarize, the second hypothesis of this paper is as follows.

H2: Sense making will mediate the effect of extreme incongruity on weirdness.

Moderators of Sense Making

In product designs, researchers have devoted attention and effort to identify factors that facilitate sense-making and eventually increase consumers' responses. For example, Jhang, Grant and Campbell (2012) posit that cognitive flexibility is a key factor in the facilitation of sense making. They demonstrate that three different manipulations of cognitive flexibility, including positive affect, a future launch, and a cognitive flexibility prime, successfully increase consumers' evaluations of extremely incongruent products. Another way to help people make sense of extreme incongruity is through feature-based association by incorporating an enabler, as proposed by Noseworthy, Murray and Muro (2017).

Since the sense-making process is proposed as an important part of the underlying mechanism, the current research explores factors related to some key characteristics of the sense-making process. In 1995, Weick posited seven properties of sense making: it is (a) grounded in identity construction, (b) retrospective, (c) enactive of sensible environments, (d) social, (e) ongoing, (f) focused on and by extracted cues, and (g) driven by plausibility rather than accuracy. Among the seven properties, we select three of them for further investigation, because we believe these three are central to the sense-making process, and factors related to some of these properties (e.g., elaboration, social norms) have been demonstrated in previous literature of congruity to be key factors in resolving incongruity (mostly moderate incongruity) (e.g., Campbell & Goodstein, 2001; Maoz & Tybout, 2002; Walchli, 2007).

Facilitating information. According to Weick (1995), one of the seven properties of sense making is "focus on and by extracted cues: paying close attention to ways people notice, extract cues, and embellish that which they extract." In other words, people extract cues from the environment to help them decide "what information is relevant and what explanations are acceptable" for their sense-making process (Brown, Stacey, & Nandhakumar, 2007). Therefore, the most direct way to facilitate sense making is to provide useful information in a relevant context so that people can extract this information and better make sense of the unexpected.

Induced elaboration. Sense making is "retrospective: people can know what they are doing only after they have done it". One key issue of this property is that sense making is "an attention process, but it is the attention in the past" (Weick, 1995). When people reflect on, and try to make sense of certain things, attention at the moment of retrospection, and interruptions to that attention, are important to the process of sense making (Gephart, 1993). Therefore, inducing

people to elaborate as a way to direct their attention to sense making should facilitate the process and change the weirdness perception.

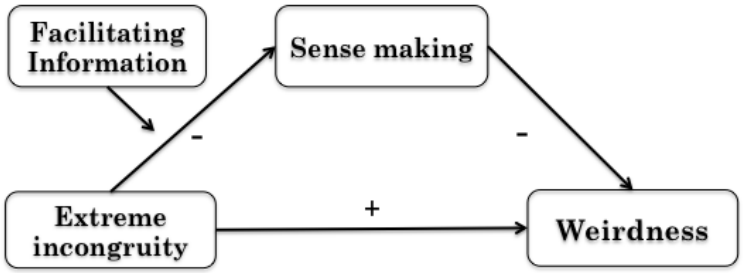
Social norms. Sense making is a social process, "because an organization of people (of any type) is a network of inter-subjectively shared meanings— a collective mind—that are sustained through the development and use of common language and everyday social interaction" (Weick, 1995). The presence of others, whether actual, imagined, or implied, affects one's sense-making process. Therefore, the social norm approach, which shows typical behavior (e.g., high or low product attitudes/acceptance) performed by others, should also influence the sense-making process and the perception of weirdness.

Although the three factors discussed above should moderate the sense-making process and could further change the weirdness perception, whether all of them can be significant moderators is open to question. When faced with extremely incongruent products, consumers' sense-making ability may be limited by the amount and kind of information they can use. According to Weick (1995), people have three sources of information: external stimuli, focused retrieval from internal memory, and seemingly random foci in working memory. For extremely incongruent product designs with which the majority of consumers are unfamiliar, useful information for sense making that can be retrieved from internal and working memory should be very limited. Therefore, compared to induced elaboration and social norm information, providing useful product-related information as external stimuli for the sense-making process may be the only one way to significantly facilitate sense making and further change people's perceptions of weirdness. Therefore, we predict that facilitating information may be the only significant moderator of the relationship between extreme incongruity and weirdness among the three moderators discussed above. As a result, the third hypothesis of this paper is as follows:

H3: Facilitating information is a significant moderator of the effect of extreme incongruity on weirdness.

To summarize, we posit that in the context of product designs, extremely incongruent products will be perceived as much weirder than congruent and moderately incongruent products (H1), because consumers cannot make sense of extreme incongruity (H2). Furthermore, providing consumers with information to assist in the sense-making process could significantly decrease the perception of weirdness for extremely incongruent product designs (H3). The overall theoretical framework is presented in figure 3.1.

Figure 3. 1. Theoretical framework.



Overview of Studies

Five studies, including eight experiments, have examined the effect of extreme incongruity on the perception of weirdness across various product categories and research contexts. Study 1 investigates whether extreme incongruity, compared to congruity and moderate incongruity, leads to the perception of weirdness. Study 2 examines whether the proposed mediator, the sense-making process, is the underlying mechanism of the effect of extreme incongruity on weirdness. Study 3-1 narrows sense making down to product usefulness as the underlying mechanism, excludes one possible alternative explanation: the affect account, and develops a more comprehensive reflective measurement scale of weirdness. Study 3-2 further

validates the measurement scale of weirdness via confirmatory factor analysis. Study 4 tested three moderators of the relationship between extreme incongruity and sense making, including facilitating information in study 4-1, induced elaboration in study 4-2, and social norms in study 4-3. Finally, study 5 looks into marketing-related outcomes of weirdness, including consumers' overall attitude, purchase intention and share intention toward weird product designs.

Study 1: Extreme Incongruity to Weirdness

The purpose of study 1 is to demonstrate that extremely incongruent designs will be perceived as significantly weirder than congruent and moderately incongruent product designs.

Design. A 3-condition (congruity vs. moderate incongruity vs. extreme incongruity) between-subject experiment design was used for this study. The experimental stimuli used in this study were developed in the following stages. (1) The Product category, MP3 players, was selected, because general consumers are familiar with and have relatively neutral or positive attitude toward this product category. A pretest with 50 Mturk participants shows that their overall attitude toward MP3 players is 5.35 on a seven-point scale (1: extremely dislike, 7: extremely like), and that their familiarity level with this product category is 5.2 on a seven-point scale (1: extremely unfamiliar, 7: extremely familiar). (2) The manipulation stimuli for congruity (five MP3 players ranging from congruent to extremely incongruent designs) were first selected by the author of this research from various online resources, and then pretested with another 50 Mturk participants with a four-item congruity Likert scale (including "it is very likely that a MP3 player comes likes this", "this MP3 player matches my expectations for MP3 players in general", "this MP3 player is a typical MP3 player", and "this MP3 player is congruent with a traditional MP3 player". 1: extremely disagree, 7: extremely agree). (3) Based on the pretest results, three MP3 player designs were selected. The congruent design was a black rectangular-shaped MP3

player; the moderately incongruent design was a black wireless headset MP3 player; and the extremely incongruent design was a golden MP3 player shaped like a grenade (see appendix 3.A).

Procedure and Measures. 151 participants (Male: 48.9%) were recruited on Mturk, with the cover story that the purpose of the study was to understand how consumers evaluate new product designs. After a consent form and instructions, participants were randomly assigned into one of three experimental conditions to evaluate a MP3 player. The dependent variable was measured by a one-item seven-point Liker scale ("this MP3 player is weird." 1: strongly disagree; 7: strongly agree). After that, perceived overall congruity was measured by the four-item seven-point Liker scale used in the pretest. At the end, some demographics were measured.

Results. For the manipulation check of congruity, an index of congruity was calculated by taking the mean of the four measurement items of congruity (Cronbach's Alpha= .942). The rectangular MP3 player was perceived as more congruent (M=5.68, SD=1.33) than the wireless one (M=3.29, SD=1.35), which was more congruent than the grenade MP3 player (M=2.37, SD=1.18, $F(2, 148)=88.77$, $p<.001$). Planned contrast tests further showed that the rectangular MP3 player was significantly more congruent than the wireless design ($t(148)=9.51$, $p<.001$), and that the wireless MP3 player was significantly more congruent than the grenade design ($t(148)=3.52$, $p=.001$).

For the dependent variable of the perception of weirdness, the extremely incongruent MP3 player (M=5.91, SD=1.56) was perceived as much weirder than the moderately incongruent one (M=4.58, SD=1.54) and the congruent one (M=2.40, SD=1.60, $F(2,148)=64.16$, $p<.01$). Planned contrast tests showed that the extremely incongruent design was perceived as significantly weirder than the moderately incongruent design ($t(148)=4.21$, $p<.001$), and the

moderately incongruent one was perceived as significantly weirder than the congruent one ($t(148)=7.13$, $p<.001$).

Discussion. In study 1, we successfully demonstrated that extreme incongruity is a key antecedent of the perception of weirdness in the context of product designs, supporting H1. The product category used here is the MP3 player. The next study uses a different product category to test the underlying mechanism of sense making.

Study 2: Sense Making as Mediator

The purposes of study 2 are two-fold. First, study 2 aims to replicate the results of study 1 by examining the effect in a different product category: speakers. Second, study 2 tests H2, that the sense-making process is the underlying mechanism of the effect of extreme incongruity on perceived weirdness.

Design. A 3-condition (congruity vs. moderate incongruity vs. extreme incongruity) between-subject experiment design was used. The congruent design was a pair of black rectangular speakers, the moderately incongruent design was a pair of blue round shaped speakers, and the extremely incongruent design was a pair of white speakers shaped like shoes (see appendix 3.A). The stimuli used in this study were develop in the same way as study 1.

Procedure and Measures. 107 Mturk participants (Male: 49.5%) were recruited. The cover story and procedure of this study are the same as those of study 1. Weirdness and the manipulation check of congruity were measured in the same ways as in study 1. The mediator of sense making is measured by a single-item seven-point Likert scale of "this pair of speakers makes sense to me" (1: strongly disagree; 7: strongly agree), adapted from research by Noseworthy, Muro and Murray (2014). At the end, some demographics were measured.

Results. For the manipulation check of congruity, the rectangular speakers (M=4.55, SD=1.35) were perceived as more congruent than the round ones (M=3.64, SD=1.27) and the shoe-shaped ones (M=2.54, SD=1.34, $F(2, 104)=22.48$, $p<.001$). Planned contrast tests showed that the rectangular speakers were significantly more congruent than the round design ($t(104)=2.79$, $p=.006$), and the round speakers were significantly more congruent than the shoe-shaped one ($t(104)=3.31$, $p=.001$).

A one-way ANOVA indicated that the extremely incongruent product (M=5.73, SD=1.43) was perceived as weirder than the congruent one (M=4.21, SD=1.94) and the moderately incongruent one (M=3.24, SD=2.17, $F(2,104) = 18.07$, $p<.001$). Planned contrast tests showed that the perceived weirdness of the extremely incongruent product was significantly higher than that of the moderately incongruent one ($t(104)=3.28$, $p=.001$), which was significantly higher than that of the congruent one ($t(104)=2.17$, $p=.032$), supporting H1.

To test the mediating role of sense making, we conducted a mediator analysis (model 4, Hayes, 2013) for one part of the data (only the moderate incongruity and the extreme incongruity conditions, as a more conservative test), for which congruity was the independent variable, perceived weirdness was the dependent variable, and sense making was the mediator. First, sense making was predicted by congruity ($t(63)=4.54$, $p<.001$). Second, with sense making in the model, congruity no longer predicted weirdness ($t(62)=4.54$, $p=.44$), but sense making ($t(62)=-5.67$, $p<.001$) did predict weirdness. Third, the mediator analyses with 5,000 bootstrap samples showed that the indirect effect of congruity on weirdness through sense making was significant (b=-1.19; 95% CI: [-.54, -1.97]). Furthermore, the total effect of congruity on weirdness was originally significant (b=-.1.52, $p<.001$). After controlling the indirect effect of sense making, the direct effect of congruity on weirdness was insignificant (b=-.33, $p=.44$). In summary, sense

making was the full mediator of the relationship between congruity and perceived weirdness, supporting H2.

Discussion. Study 2 successfully replicated the results of study 1 in a different product category: speakers. It provided direct evidence of the underlying mechanism of sense making. However, several questions were still left to be addressed. First, one may question, since speakers and MP3 players belongs to the same product category of electronics, whether the effect will still hold in other categories. Second, sense making is a very broad construct, so one may ask what the consumer actually makes sense of in the context of product designs and evaluation. Third, no alternative explanations for the effect have been considered or tested. The next study addresses these issues.

Study 3: Alternative Explanation and Discriminant Validity

The purposes of study 3 include the following: (1) Study 3 tries to replicate the results of studies 1 and 2 by examining the effect in different product categories outside of the electronic product category. (2) This study also aims to further understand, as a result of consumers' sense-making process, which factors play key roles in explaining the effect of congruity on perceptions of weirdness. In the context of product designs and evaluation, we predict that product usefulness may be the most relevant outcome of consumers' sense-making process and the key factor influencing consumers' perception of weirdness. (3) Another main purpose of study 3 is to rule out one alternative explanation. As discussed earlier, the affect theory is highly relevant for this project. According to previous literature on Mandler's theory (1982), any incongruity will automatically change people's physiological arousal level, when compared to congruity. Correspondingly, one's arousal level could further change the intensity level of one's affect, the various emotions one experiences at the moment. Which specific emotion(s) will be evoked

further depends on one's skill level, according to Csikszentmihalyi's flow model (1997). Theoretically speaking, extreme incongruity may cause negative emotions such as anxiety, frustration or upset, because the incongruity between a new product and people's general product category schema is severe, which means the challenge level is high, and ordinary consumers' ability to make sense of the new product design is relatively low. Therefore, these induced negative emotions may lead to the perception of weirdness. In study 3, consumers' arousal levels and affect were measured to test this account. (4) Study 3 also tries to develop a more comprehensive reflective measurement scale of weirdness to capture and cover the theoretical domain of the weirdness construct. Moreover, in the psychology and marketing literature, many constructs have been identified as sharing a common antecedent with weirdness (i.e., incongruity), such as creativity and coolness. Therefore, study 3 also aims to establish its discriminant validity versus these related constructs.

Study 3-1: Alternative Explanations

Design. A 3-condition (congruity vs. moderate incongruity vs. extreme incongruity) between-subject experiment design was used. The product category used in this study was packaged butter, a very common and daily used good in North America. The congruent design was an individual butter packet (butter in a square plastic box with a cover that can be easily peeled off); the moderately incongruent design was an individual butter packet with a wooden knife-shaped cover (referred to as a butter knife); and the extremely incongruent design was a stick of butter packaged in a container shaped like a glue stick (referred to as a butter glue stick) (see appendix 3.A).

Procedure and Measures. 153 undergraduate students were recruited in a North American University to participate in a research bundle, which included multiple separate studies,

in exchange for one research credit. The cover story and the procedure of this study were the same as those of study 1. In order to better capture the theoretical domain of the concept of weirdness, four reflective measurement items were developed, including "weird", "bizarre", "strange" and "odd", and were used in this study to measure weirdness. A short version of the manipulation check measure of congruity was used, including a two-item seven-point Likert scale of "it is consistent with typical butter in my mind", and "it is a typical kind of butter". The mediator of sense making was measured in the same way as that of study 2. Product usefulness was measured by a single-item seven-point bipolar scale (useless: 1; useful: 7). The alternative mediator of arousal was measured by a self-reported four-item seven-point bipolar scale (tense-relax, stimulated-bored, unlively-lively, dull-bright). Affect was measured by the PANAS scale, based on that of Watson, Clark, and Tellegen (1988). After that, participants proceeded to other studies in the research bundle.

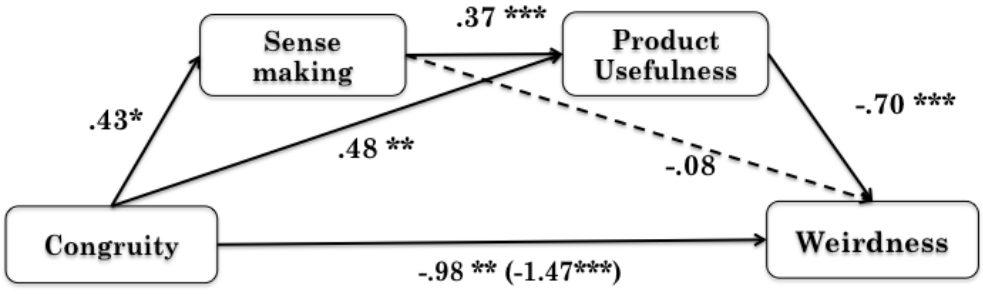
Results. For the manipulation check of congruity, the individual butter packet (M=4.72, SD=1.37) was perceived as more congruent than the butter knife (M=4.33, SD=1.35) and the butter glue stick (M=3.16, SD=1.77, $F(2, 150)=15.00$, $p<.001$). Planned contrast tests showed that the perceived congruity of the individual butter packet was higher than that of the butter knife, but not significantly different ($t(150)=1.29$, $p=.199$), and the perceived congruity of the butter knife is significantly higher that of the butter glue stick ($t(150)=3.97$, $p<.001$). Although the operationalization of the congruity was not successful given the individual butter packet is not significantly different from the butter knife in terms of the perceived congruity, the extreme incongruity was showed to be significantly less congruent than the other two designs.

The four-item measurement scale of weirdness was analyzed according to exploratory factors. The analysis results showed that the four items explained 86.84% of the total variance,

and were loaded on a single factor with a very high internal consistency (Cronbach's Alpha = 0.948). As a result, we took the average of these four measurement items to form an index of weirdness. A one-way ANOVA showed that the extremely incongruent product (M=4.50, SD=1.77) was perceived as weirder than the congruent one (M=3.04, SD=1.37) and the moderately incongruent one (M=2.73, SD=1.35, $F(2,150) = 20.21$, $p<.001$). Planned contrast tests showed that the perceived weirdness of the extremely incongruent product was significantly higher than that of the moderately incongruent one ($t(150)=4.91$, $p<.001$), which is not significantly different from that of the congruent one ($t(150)=1.05$, $p=.295$), consistent with the manipulation checks.

As discussed earlier, we propose that product usefulness is the most important outcome of consumers' sense-making process and the key factor influencing the perception of weirdness. Therefore, we conducted a serial mediator analysis (model 6, Hayes, 2013) for a part of our data (the extreme incongruity and moderate incongruity conditions), for which congruity is the independent variable, perceived weirdness is the dependent variable, and sense making and product usefulness are the mediators (see Figure 3.2).

Figure 3. 2. Serial Mediation Analysis of Study 3-1.



First, sense making was predicted by congruity ($t(100)= 1.80$, $p=.04$). Second, product usefulness was predicted by both congruity ($t(99)= 3.19$, $p=.002$) and sense making ($t(99)= 7.41$, $p<.001$). Third, with both sense making and product usefulness in the model, weirdness was

predicted by product usefulness ($t(98)= -3.78$, $p<.001$) and congruity ($t(98)= -3.32$, $p=.0013$), but not sense making ($t(98)= -.69$, $p=.59$).

The bootstrapping results showed that the only significant path was the indirect effect of congruity on weirdness through usefulness ($b=-.340$; 95% CI: [-.772, -.08]). Furthermore, the indirect effects of congruity on weirdness through sense making ($b=-.03$, 95% CI: [-.31, .07]) and through sense making and usefulness ($b=-.11$, 95% CI: [-.37, .01]) were insignificant. In summary, these results provide some preliminary evidence for our prediction that product usefulness could be the key outcome of the sense-making process, and further influence the weirdness perception.

Alternative Explanation. To test the possible mediating effect of arousal, an independent-sample *t* test was conducted to examine the mean difference between the arousal levels of the moderately incongruent and extremely incongruent product designs. As a result, the arousal levels between the conditions of moderate incongruity ($M=4.29$, $SD=.63$) and extreme incongruity ($M=4.33$, $SD=.66$) were not significantly different from each other ($t(100)=-.27$, $p=.789$). Furthermore, the mediation analysis (model 4, Hayes, 2013) also showed that arousal was not a significant mediator of the effect of congruity on weirdness. More specifically, neither the effect of congruity on arousal ($b=-.03$, $p=.79$), nor the effect of arousal on weirdness ($b=-.20$, $p=.42$) is significant. The total effect of congruity on weirdness is significant ($b=-1.47$, 95% CI: [-2.08, -.85]), but the indirect effect of congruity on weirdness through arousal is insignificant ($b=.01$, 95% CI: [-.05, .14]).

As for the possible mediating effect of affect, every measurement item on the PANAS scale was first analyzed by an independent-sample *t* test to detect any significant difference between the conditions of moderate incongruity and extreme incongruity. Two emotions, “upset”

($M_{\text{Moderate incongruity}}$=1.55, SD=1.29; $M_{\text{Extreme incongruity}}$=2.16, SD=1.63, t(100)=2.09, p=.04) and "disgusted" ($M_{\text{Moderate incongruity}}$=1.27, SD=.90; $M_{\text{Extreme incongruity}}$=1.75, SD=1.38, t(100)=2.04, p=.045), were found to be significantly different. However, the mediation analysis (model 4, Hayes, 2013) showed that neither upset (b=-.13, 95% CI: [-.40, .0002]) nor disgust (b=-.05, 95% CI: [-.25, .05]) were significant mediators of the effect of extreme incongruity on weirdness.

Study 3-2: Discriminant Validity of Weirdness

Study 3-2 tries to (1) further establish the external validity of the effect by changing the product category of congruity manipulation, and (2) more importantly, to examine the discriminant validity of the weirdness scale versus related psychological constructs including creativity and coolness. Creativity is defined as the process by which something new and somehow valuable is formed (Mumford, 2003). In the context of product designs and evaluation, a product is perceived as creative because it is novel and useful. Coolness is "a subjective and dynamic, socially constructed positive trait attributed to cultural objects (people, brands, products, trends, etc.) inferred to be appropriately autonomous" (Warren & Campbell, 2014, p.544). In the context of product designs and evaluation, cool products usually diverge from the norm, which can be interpreted as the product category schema. One can notice that something novel in the creativity literature and something diverging from the norm in the coolness literature share the common trait of incongruity with weirdness. In study 3-2, weirdness, creativity and coolness were measured and analyzed using confirmatory factor analysis to test their discriminant validity.

Design. A 2-condition (congruity vs. extreme incongruity) between-subject experiment design was used. The product category used was tea infusers. The congruent design was a

stainless steel steeper filter, and the extremely congruent design was a silicon tea infuser shaped like a small person (see appendix 3.A).

Procedure & Measures. 87 undergraduate students were recruited in a North American University in exchange for research credits. The cover story and the procedure of this study were the same as those of study 1. Weirdness and the manipulation check of congruity were measured in the same ways as study 3-1. Product creativity was measured by a three-item seven-point Liker scale ("this product is creative," "this product is smart," and "this product is innovative"). Coolness was measured by a two-item seven-point Liker scale ("this product is cool" and "this product is hip") adopted from Warren and Campbell's 2014 study.

Results. For the manipulation check of congruity, the stainless steel steeper filter (M=5.52, SD=1.37) was perceived as more congruent than the tea infuser shaped like a small person (M=2.93, SD=1.37, $t(81)$=8.66, p<.001).

In terms of perceived weirdness (the mean of the four measurement items), the extremely incongruent product (M=4.25, SD=1.62) was perceived as weirder than the congruent one (M=1.87, SD=1.87, $t(81)$ = -7.82, p<.001).

The four measurement items of weirdness, the three items of creativity, and the two items of coolness were analyzed by confirmatory factor analysis using AMOS. First, the analysis demonstrated a good model fit (root mean square error of approximation [RMSEA]=.119; confirmatory fit index [CFI]=.953; incremental fit index [IFI]=.954; relative fit index [RFI]=.850). Second, the weirdness (Cronbach's Alpha=.965), creativity (Cronbach's Alpha=.769), and coolness (Cronbach's Alpha=.866) scales showed good internal reliability. Third, inter-correlations between the weirdness construct and the creativity construct (r =-.064, n.s.) and between the weirdness and the coolness construct (r =.098, n.s.) were insignificant.

However, the correlation between creativity and coolness was significantly high (r=.75, p<.001). Finally, the discriminant validity of weirdness relative to creativity was supported by the comparisons between average variance extracted (AVEs) ($AVE_{Weirdness}$=. .87, $AVE_{Creativity}$= .56) and the squared correlation (.004), and so was the discriminant validity of weirdness to coolness ($AVE_{Weirdness}$=. .87, $AVE_{Creativity}$= .79, squared r=.01). Although creativity and coolness were correlated with each other, their average AVEs (.68) was still greater than their squared correlation (.56), which supports their discriminant validity.

Furthermore, the perceived creativity of the extremely incongruent design (M=5.37, SD=.82) was significantly higher than that of the congruent (M=4.77, SD=1.30, t(81)=2.25, p<.001), but the perceived coolness between the extremely incongruent (M=4.54, SD=1.26) and congruent designs (M=4.15, SD=1.33) were not significantly different from each other (t(81)=1.34, p=.184).

This study demonstrated the discriminant validity of weirdness versus creativity and coolness, but this study also show some interesting findings in terms of how congruity affects weirdness, creativity and coolness. This should be further investigated by future research.

Discussion. In summary, study 3-1 successfully replicated the results of studies 1 and 2 in a different product category (i.e., butter), and excluded arousal and negative affect as possible alternative underlying mechanism of the effect of extreme incongruity on the perception of weirdness. Moreover, a more comprehensive reflective measurement scale of weirdness was developed to better capture the weirdness construct. In study 3-2, the discriminant validity of the weirdness measurement scale versus creativity and coolness was further established. Up to now, the basic effect of extreme incongruity on weirdness and the underlying mechanism of sense

making have been established and replicated. The next study examines relevant moderators of the effect.

Study 4: Moderators of the Sense-Making Process

The main objective of Study 4 is to further examine the underlying mechanism of sense making by manipulating relevant moderators, including facilitating information for sense making (study 4-1), induced elaboration (study 4-2), and social norms (study 4-3). Furthermore, different operationalization of congruity was developed and used to replicate the previous results in different product categories.

Study 4-1: Moderating Effect of Facilitating Information

Design. A 2 (congruity vs. extreme incongruity) by 2 (facilitating information vs. control) between-subject experiment design was used. The product category used in this study was hats. The congruent design was a black knitted beanie, and the extremely incongruent design was a black knitted beanie with an attached part shaped like a beard (referred to as the beard hat) (see appendix 3.A). Facilitating information was manipulated by varying the product information provided. In the facilitating information condition, the product information provided was useful and positive for sense making (e.g. "The beard is detachable and foldable. You can fold away the beard easily for fancy occasions!"); and in the control condition, the product information provided was meaningless or less helpful yet positive information about the product (e.g. "Looks good for all kinds of occasions!"). For both conditions, four pieces of information were provided to control the total amount of information provided.

Procedure and Measures. 181 participants (Male: 56.9%) were recruited from Mturk. The cover story was slightly different from previous studies. In this study, participants were asked to imagine that they were browsing randomly online and saw the products shown on the

following page. They were then randomly assigned into the four experiment conditions. After reviewing the product information, participants answered various questions, including the perceived weirdness, a sense-making measure, manipulation check measures for congruity and facilitating information, and finally some demographic information. Most of the measures were the same as those in study 3, and the manipulation check measure for facilitating information included a seven-point question of "to what extent do you think the product information is useful for your evaluation?" (1: extremely useless, 7: extremely useful). In order to control the valence of the product information, a seven-point question of "how positive or negative is the product information?" (1: extremely negative, 7: extremely positive) was also asked.

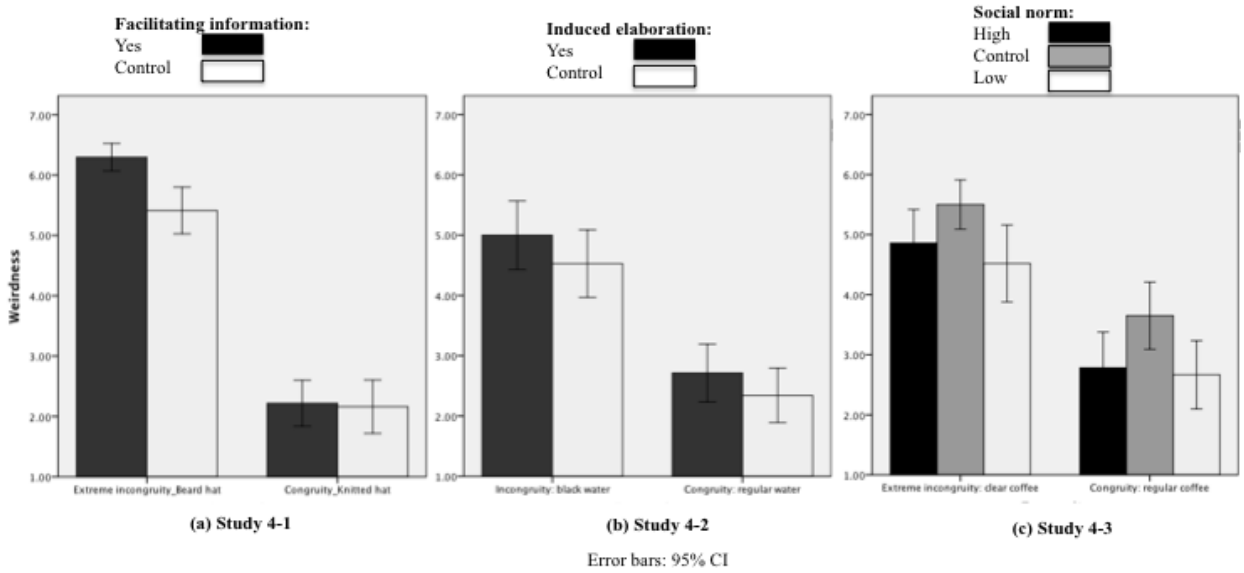
Results. For the manipulation check of congruity, the beanie (M=5.20, SD=1.26) was perceived as more congruent than the beard hat (M=2.27, SD= 1.36, t(179)=14.97, p<.001).

For the manipulation of facilitating information, participants in the facilitating information condition (M=5.62, SD=.99) perceived the product information as significantly more useful than those in the control condition did (M=5.03, SD= 1.18, t(179)=3.87, p<.001). In terms of information valence, the product information was perceived as equally positive between the facilitating information condition (M=5.84, SD=.87) and the control condition (M=5.70, SD= .89, t(179)=1.13, p=.26).

A two-way ANOVA with congruity and facilitating information as two independent factors demonstrated a main effect of congruity (F(1, 177)=602.32, p<.001), a main effect of facilitating information (F(1, 177)=9.92, p=.01), and more importantly, a significant interaction effect (F(1, 177)=7.65, p=.025). Consistent with our prediction (see figure 3.2-a), for the extremely incongruent product, the perceived weirdness was significantly lower in the facilitating information condition (M=5.41, SD=1.36) than that in the control condition (M=6.30,

SD=.76, *F*(1,177)=12.30, *p*=.001). However, for the congruent product, the perceived weirdness did not differ between the facilitating information condition (*M*=2.16, *SD*=1.37) and the control condition (*M*=2.22, *SD*=1.30, *F*(1,177)=.05, *p*=.83), supporting H3.

Figure 3. 3. Results of study 4.



We further tested the underlying mediator of sense making using a moderated mediation analysis (model 7, Hayes, 2013). Consistent with our prediction, the moderated mediation was significant (index =.68, 95% CI: [.10, 1.40]). Furthermore, for the extremely incongruent product, the indirect effect of facilitating information via sense making was significant (*b*=-.70; 95% CI: [-1.33, -.22]), but for the congruent product, the indirect effect of facilitating information via sense making was insignificant (*b*=-.02; 95% CI: [-.32, .32]).

Study 4-2: Moderating Effect of Induced Elaboration

Design. A 2 (congruity vs. extreme incongruity) by 2 (Elaboration: induced vs. control) between-subject experiment design was used. The product category used for this study was bottled mineral water. The congruent design was a bottle of clear mineral water with a cylinder-shaped transparent plastic bottle, and the extremely incongruent design was a bottle of black

mineral water (referred to as the black water) (see appendix 3.A). A pretest (109 Mturk participants) of this operationalization of congruity was conducted, in which the regular bottled mineral water (M=5.32, SD=1.36) was perceived as more congruent than the black mineral water (M=2.77, SD=1.56, t(107)=9.07, p<.001).

Induced elaboration was manipulated by varying the filler task before the dependent variable measurement. In the induced elaboration condition, participants were asked to write down three situations in which they may want to purchase or try this product, and briefly explain why. In the control condition, participants were asked to write down three tourism destinations that they wanted to visit, and briefly explain why.

Procedure and Measures. 139 undergraduate students were recruited in a North American University in exchange for one research credit. At the beginning of the lab sessions, participants were informed that they would be participating in multiple separate small studies. Our focal study was a product evaluation task. After the study introduction, participants were randomly assigned into one of the four experiment conditions. First, they were asked to review a consumer product. Second, they were asked to participate in an elaboration task. Third, they answered various questions regarding perceptions of weirdness, sense-making measures, and demographic information. All of the measures were the same as those in study 4-1.

Results. A two-way ANOVA with congruity and induced elaboration as two independent factors showed only a significant main effect of congruity (F(1, 135)=77.35, p<.001). The interaction effect between congruity and elaboration was not significant (F(1, 135)=.04, p=.84) (see figure 3.2-b). The pair-wise comparisons demonstrated that, for the extremely incongruent product, induced elaboration could directionally decrease the perceived weirdness, but not significantly ($M_{\text{Elaboration}}$=4.53, SD= 1.63; M_{Control}=5.00, SD= 1.63; F(1, 135)=1.71, p=.19). For

the congruent product, induced elaboration did not significantly change the perceived weirdness ($M_{Elaboration}$=2.34, SD= 1.31; $M_{Control}$=2.71, SD= 1.39; $F(1, 135)$=1.08, p=.30).

Study 4-3: Moderating Effect of Social Norms

Design. A 2 (congruity vs. extreme incongruity) by 3 (social norm: high vs. low vs. control) between-subject experiment design was used. The product category used in this study was bottled coffee. The congruent design was a cylindrical transparent glass bottle of milky coffee, and the extremely incongruent design was identical except that the coffee was clear and transparent (referred to as clear coffee) (see appendix 3.A).

Social norm was manipulated by varying the market research information provided with the focal product. In the high social norm condition, participants were informed that according to our previous marketing research, 85% of consumers liked this product. In the low social norm condition, participants were informed that 85% of consumers disliked the product. In the control condition, no information about the marketing research was provided.

Procedure and Measures. 147 undergraduate students were recruited in a North American University to participate in a research bundle in exchange for one research credit. Our focal study was a product evaluation task, and the cover story was to help a coffee manufacturer evaluate one of their new product designs. After the study introduction, participants were randomly assigned into one of the six experiment conditions. First, they were asked to review a new product design. Second, except for those in the control condition, participants were provided with marketing research information about the new product design, in which the social norm manipulation was embedded. Third, they answered questions about their perceptions of weirdness, sense-making measures, and demographic information. All of the measures were the same as those in study 4-2.

Results. For the manipulation check of congruity, the milky coffee (M=4.97, SD=1.25) was perceived as more congruent than the clear coffee (M=2.63, SD=1.22, t(145)=11.49, p<.001).

A two-way ANOVA with congruity and social norm as two independent variables showed only a significant main effect of congruity (F(1, 141)=75.49, p<.001; $M_{\text{Congruity}}$= 2.04, SD=1.42; $M_{\text{Extreme Incongruity}}$=4.95, SD=1.36) and a significant main effect of social norms (F(2, 141)=7.16, p<.01; M_{Control}= 3.84, SD= 1.72; $M_{\text{High Norm}}$=3.61, SD= 1.72; $M_{\text{Low Norm}}$= 4.56, SD=1.50). However, the interaction effect between congruity and social norm was not significant (F(2, 135)=.12, p=.89).

Discussion. In summary, the three studies of study 4 replicated the basic effect of congruity on weirdness using three new manipulation stimuli of congruity in three product categories. These studies examined three moderators, facilitating information, induced elaboration, and social norms, to further explore the underlying mechanism of sense making. Among these three moderators, only facilitating information was a significant moderator, supporting H3. These results can be interpreted as follows. First, extreme incongruity is a key antecedent and a strong and robust manipulation for weirdness. Second, the underlying mechanism of a failed sense-making process is difficult to reverse due to the severity and extremity of incongruity. Induced elaboration cannot help people resolve extreme incongruity, probably because their ability to resolve is bounded, and their motivation to resolve and their devoted efforts are not enough to fully make sense of extreme incongruity. The social norm information may change consumers' attitudes toward or acceptance of an extremely incongruent product design, but it does not directly help with the sense-making process and the perception of weirdness. Overall, only the information that is essential to the sense-making process can moderate the effect of extreme incongruity on perceptions of weirdness.

Study 5: So What? Behavioral Outcomes of Extremely Incongruent Products

The previous four studies have consistently shown that extreme incongruity is a key antecedent of perceived weirdness, and that a failed sense-making process, particularly the limited usefulness of a focal object, is an important mediator in the effect of extreme incongruity on the perception of weirdness. Given the common negative connotations of weirdness, one may question the marketing implications of extremely incongruent products and weirdness. How do consumers respond toward these extremely incongruent product designs, with respects to some marketing-related outcome variables (e.g., attitude, behavioral intention)? If these influential responses toward these products are all negative, why bother to study extremely incongruent product designs or weirdness? Therefore, the main objective of study 5 is to examine some consequential behavioral outcomes and to demonstrate the marketing potential of extremely incongruent product designs.

The first set of behavioral outcomes we choose to examine is consumers' attitude and purchase intention. In marketing, consumers' product attitude and purchase intention have been regarded as a key index for the success of product designs. Previous literature on congruity theory and product designs consistently finds that compared with congruent and moderately incongruent product designs, extremely incongruent product designs usually lead to negative attitude and limited purchase intention (e.g., Noseworthy, Muro, & Murray, 2014; Peracchio & Tybout, 1996). As discussed and shown in previous studies in this paper, extremely incongruent products are hard to be made sense of by consumers (study 2), and the failed sense making directly leads to low perceived product usefulness (study 3). Therefore, compared with congruent product designs, extremely incongruent product designs should lead to lower attitude and purchase intention, because of the limited usefulness.

H4: Extremely incongruent product designs, compared with congruent ones, will lead to lower attitude.

The second behavioral outcome we choose to examine is consumers' Word-of-Mouth (WOM) intention, another important marketing-related outcome variable. Current literature on consumers' WOM behavior shows that consumers are very conscious and strategic about their sharing behavior (Chen & Kirmani, 2015; Fleming, et al., 1990), and that there are many different motivations behind consumers' sharing behavior (Berger, 2014). At the same time, extremely incongruent or weird products have many characteristics that could motivate consumers to share these products with others. First, one of the motivations for WOM behavior is to fill conversational space (Berger, 2014). Previous literature on bizarreness has shown that weird or bizarre objects or experiences are distinctive from other elements in the relevant context, and the distinctiveness can facilitate memory recall. Therefore, compared to regular product designs, extremely incongruent products are more likely to be recalled, and then more likely to be mentioned in daily conservations. Second, another motivation behind WOM behavior is to facilitate sense making (Rimé, 2009). Since extremely incongruent and weird products are hard to be made sense of, consumers may be motivated to talk with other in order to seek for facilitating information for their own sense-making process. Therefore, consumers are more likely to share extremely incongruent products. Third, another motivation behind WOM behavior is the self-enhancement motive. More specifically, previous research finds that people like to share entertaining things (e.g., interesting, surprising, novel things) in order to make themselves look good to others (Berger, 2014; Berger & Schwartz, 2011). As shown by the research in the entertainment area, weird and bizarre things or acts could have certain entertaining value, so

consumers are more likely to share the extremely incongruent and weird products with others. Based on the three arguments above, we propose the following hypothesis.

H5: Extremely incongruent product designs, compared with congruent ones, will lead to higher share intention.

Furthermore, one may notice and criticize operationalization of extreme incongruity in previous studies. In some studies, we manipulated more than one attribute of the focal product. For example, we changed both the color and shape for the MP3 player in study 1 and the speakers in study 2, which may inflate the effect size of extreme incongruity on weirdness. In some studies, our operationalization may cause mis-categorization. For example, the clear coffee in study 4-3 can be categorized as coffee or water, which may also lead to the perception of weirdness. In order to address this concern, study 5 used a different operationalization strategy, in which only one product attribute is changed, and the mis-categorization problem is avoided.

Design. A 2-condition (weirdness vs. control) between-subject experiment design was used. The weirdness was manipulated by changing one attribute in our focal product, and the product category used in this study is another fast consumption product: toilet paper. The regular product in the control condition was a roll of white toilet paper, and the weird product was a roll of black toilet paper (see Appendix 3.A).

Procedure and Measures. 116 participants were recruited in a North American University to participate in a research bundle in exchange for one research credit. The cover story and the basic procedure were the same as study 4-1. The dependent variables included (1) overall attitude, measured by a seven-point question of "overall, how do you like/dislike this product?" (1: extremely dislike; 7: extremely like), (2) purchase intention, measured by a seven-point question of "would you be interested to buy this product?" (1: not interested at all; 7:

extremely interested), and (3) share intention, measured by two seven-point questions of "will you share information about this product on social networks?" and "will you mention this product during daily conversations with other people such as your family and friends?" (1: definitely will not; 7: definitely will). After the dependent variables, participants proceeded to answer the manipulation check measures of weirdness and congruity.

Results. For the manipulation check of congruity, the black toilet paper (M=2.88, SD=1.42) was perceived as less congruent than the white one (M=5.68, SD= 1.24, $t(114)$=11.29, p<.001).

First, as for the perceived weirdness, the black toilet paper (M=5.20, SD=1.27) was perceived as weirder than the white one (M=2.54, SD= 1.33, t(114)=-11.00, p<.001).

Second, consumers' overall attitude toward the weird black toilet paper (M=3.97, SD=1.77) was significantly lower than that of the white one (M=4.86, SD= 1.09, t(114)=3.24, p=.002). Similarity, consumers' purchase intention toward the weird black toilet paper (M=3.83, SD=1.98) was significantly lower than that of the white one (M=4.71, SD= 1.22, t(114)=2.87, p=.005).

Third, the two measurement items of share intention were averaged to form an index of share intention. The analysis results showed that consumers' share intention toward the weird black toilet paper (M=4.18, SD=1.64) was significantly higher than that of the white one (M=2.02, SD= 1.01, t(114)=-8.44, p<.001).

Discussion. This study, using a cleaner operationalization of extreme incongruity and weirdness than previous studies, demonstrates that although weird products produce lower overall attitudes and purchase intentions, they also result in higher share intentions, which has important implications for marketing.

The finding about the relationship between extreme incongruity and WOM behavior is new to the congruity and the product design literature. Although very interesting, examining the underlying reason behind the WOM behavior goes beyond the scope of this paper, so we did not empirically examine or test the possible mechanism (e.g., the high accessibility in memory recall, the motivation for sense making, and the self-enhancement motive). However, it should be noted that this finding definitely opens a new and important research question, and deserves future research.

General Discussion

This paper, via reviewing relevant literature and a series of experiments, presents evidence that in the context of product designs, the key factor causing the weirdness perception is extreme incongruity, in the sense that the focal product design is extremely incongruent or different from its general product category schema. More specifically, it is demonstrated that compared to congruent and moderately incongruent product designs, extremely incongruent designs lead to the perception of weirdness. Furthermore, the underlying mechanism identified is a failed sense-making process. This project also demonstrates the marketing potential of extremely incongruent products and weirdness: although extremely incongruently products are weird and produce less positive attitude and lower purchase intention, consumers are more willing to share information about such products than more conventional products.

This paper contributes to the marketing literature by providing a definition of weirdness in the context of product designs for future research: *weirdness is something having the quality of being bizarre, odd or strange, caused by extreme incongruity between the focal object and people's expectations (that originate from the product category schema), due to a failed sense-making process.* Although this definition was based on the results of this paper in the context of

product designs, it could also provide a useful reference for future studies on weirdness in other contexts. We further propose a general definition of weirdness beyond the product designs: *weirdness is something (e.g., products, persons, events, companies, places, et al.) having the quality of being bizarre, odd or strange, caused by extreme incongruity between the focal object and people's expectations, due to a failed sense-making process.* This general definition not only brings various research streams across different research areas (e.g., psychology, entertainment, and marketing), but also provides a useful theoretical foundation for future research on this topic.

This project also contributes to the congruity/incongruity literature by identifying extreme incongruity as the key antecedent of weirdness. In marketing and product designs, the research focus on congruity is typically on the influence of congruity on consumers' affective responses toward products (e.g., attitudes, or purchase intentions). And this focus has led to debate over whether congruity or moderate incongruity produces better affective outcomes, which has further led to a lack of attention to extreme incongruity due to its associated negative affective responses. This paper looks beyond the existing literature of congruity and identifies an important link between extreme incongruity and weirdness. In the product design literature, extreme incongruity is usually defined as a significant difference or inconsistency between a focal product and its general product category schema, and it can also be defined as the extreme violation of consumers' general expectations usually formed by the product category schema. Extreme violations of people's expectations make up a common theme in various related constructs across different disciplines. For example, bizarre materials in psychology are operationalized by "presenting common items in bizarre relations to each other" (McDaniel & Einstein, 1986), which violates people's expectations for these common items and their relations between each other. In the entertainment field, weird and outrageous acts are defined as conduct

"out of the norms of the culture of a people" (Sergius Koku, 1995), and cultural norms are important factors in the formation of expectations. Furthermore, extreme incongruity is also closely related to the expectancy violation theory (Burgoon, 1976), which is a well-established and widely examined theory in social science. Therefore, we expect that the extreme violation of people's expectations could a common antecedent for weirdness in many other application areas, such as social judgment and interpersonal communications

This paper may also shed some light on the underlying mechanism between congruity and consumers' responses in marketing. As noted above, the current literature on congruity has focused on its effect on consumers' affective responses. According to the congruity hypothesis and the schema congruity effect, extreme incongruity, whether due to mental discomfort (Festinger, 1957), an unbalanced state (Heider, 1958), processing difficulty (Fiske & Pavelchak, 1986), or a difficult accommodation process (Mandler, 1982), usually creates high levels of arousal and negative affect, and further leads to negative responses by consumers. Unlike much of the existing literature, this paper excludes arousal and affect as alternative explanations, and demonstrates that a failed sense-making process is the underlying mediator of the effect of extreme incongruity on the perception of weirdness. As Craig-Lees (2001) and other theorists of sense making have pointed out, sense making is a conscious process involving more than simply physiological responses. We acknowledge the existence of physiological responses (such as arousal) and affect toward extreme incongruity, but our empirical evidence shows that neither arousal level nor negative affect mediate the effect of extreme incongruity on weirdness.

This paper also contributes to our understanding of sense making by demonstrating that when faced with extreme incongruity, certain properties of sense making play more essential roles than others. This paper discusses and examines three moderators related to the properties of

sense making (Weick, 1995): useful information as external stimuli, induced elaboration to ensure attention toward sense making, and social norms related to the social characteristics of sense making. Among these three moderators, only useful facilitating information is identified as a significant moderator. This may suggest that people's general ability to make sense of extreme incongruity may be significantly bounded by a lack of essential information related to the focal object. Devoting more attention and cognitive effort to the sense-making process may not be enough to resolve incongruity. Social norm information, knowing that many other consumers dislike or like the product, also fails to facilitate the sense-making process or change the perception of weirdness.

This paper challenges the common perception that weirdness is negative by further demonstrating that weirdness has positive or constructive marketing potential. Specifically, we found that, compared to more conventional products, weird products produce lower attitude, which is consistent with existing findings in the congruity literature. However, consumers' sharing intentions toward weird products are significantly higher than those of regular products, a relatively new insight in marketing literature. Word-of-mouth (WOM) behavior has become a very important outcome variable for marketing nowadays, as evidenced by the large number of research papers on this topic (see the review paper by Berger, 2014). Consumers, online and in-person, talk about various products they have bought and used, experiences they have had, brands they like or dislike, or places they have visited. Such WOM discussions influence brand awareness, brand attitude, purchase decisions and behaviors, and, most importantly, product sales (Chevalier & Mayzlin, 2006; Chintagunta, Gopinath, & Venkataraman, 2010; Godes & Mayzlin, 2009; Trusov, Bucklin, & Pauwels, 2009; Van den Bulte & Wuyts, 2009). Given the

importance of the WOM behavior in marketing, the marketing potential of extremely incongruent or weird products is worthy of future study.

Limitations and Directions for Future Research

Since this paper is among the first research efforts to understanding the weirdness construct, there are many topics that can be addressed in future research.

This paper uses many different manipulation stimuli to test and validate the effect of extreme incongruity on weirdness. Existing literature on congruity discusses different types of congruity that lead to different outcomes, such as conceptual congruity, which focuses on positioning and brand associations, and perceptual congruity, which focuses on certain product attributes (e.g. Kuo & Rice, 2015). Future research could build on this literature to examine whether conceptual incongruity or extreme perceptual incongruity would lead to different perceptions of weirdness.

Study 3 argued that product usefulness is the most important outcome of consumers' sense-making processes as well as the key factor influencing consumers' perception of weirdness, and empirically demonstrated its significant mediating role between extreme incongruity and weirdness. However, serial mediation (extreme incongruity → sense making → product usefulness → weirdness) is not significant. The psychology literature on loss and trauma suggests (e.g., Davis, Nolen-Hoeksema, & Larson, 1998) that when coping with major loss or trauma, people usually engage in two different ways to find meaning: sense making ("meaning-as-comprehensibility") and benefit finding ("meaning-as-significance), and they play independent roles in the process. Future research could further investigate the relationship between and the different roles of sense making and product usefulness in driving the weirdness perception. Furthermore, usefulness is also a very abstract construct, so future research could

explore which aspects of usefulness (e.g., product quality, functionality, or convenience of usage) exert the most influence.

Study 3-2 empirically established discriminant validity of weirdness versus creativity and coolness, but the theoretical differences among these three constructs are still worthy of future investigation. As for the difference between weirdness and creativity, study 3-2 demonstrated that compared to congruent product designs, extremely incongruent designs were weirder and more creative at the same time. Our conjecture is that incongruity is a common antecedent of weirdness and creativity, and the distinguishing factor may be the outcome of the sense-making process, more specifically product usefulness. Weird products are hard to be made sense of, which leads to low perceived usefulness, while consumers have to see certain usefulness from a new product design to generate a creativity perception. As for the difference between weirdness and coolness, study 3-2 demonstrated that compared to congruent product designs, extremely incongruent designs were weirder, but equally cool. According to Warren and Campbell (2014), one key antecedent to coolness, apart from divergence from the norm, is appropriate autonomy. Our conjecture is that in a product evaluation task, people could not exert their control or influence on the product design to display their autonomy, which leads to no difference between their coolness perception of congruent and extremely incongruent designs.

Study 4-2 failed to find a moderating effect of induced elaboration as a proxy for the attention devoted to the sense-making process. It is possible that without induced elaboration, consumers are already motivated and actively engaged in sense making, and this could be why we did not observe a moderating effect. Other moderators, such as the cognitive load, may be more effective manipulators of attention devoted to sense making. Instead of directing more attention to sense making, distracting people's attention from sense making may change the

perception of weirdness, and this could also permit consideration of whether sense making is a conscious or unconscious process.

This paper discusses and examines moderators related to three specific properties of sense making, and future research could focus on other properties than the ones explored here. For example, one property of sense making is "ground in identity construction" (Weick, 1995). Related to this, all the product categories we examined in this paper are low-involvement ones. Perhaps the perception of weirdness could be significantly different for high-involvement, or identify-relevant products. Sense making is also an "on-going" process (Weick, 1995), and does not stop right after the evaluation of the product. Future research could adopt a longitudinal research paradigm to examine how sense making and the perception of weirdness change after a delay or over a period of time.

Finally, the link between extreme incongruity and Worth-of-Mouth behavior is definitely worthy of future research. Future research needs not only to validate this effect, but also to explore its underlying reasons. As discussed earlier in study 5, people may be more willing to share information about extremely incongruent products to make sense of them for emotion regulation purpose, or to present a better self-image to others due to the entertaining value being provided, or just to fill the conversation gap due to the easy accessibility. We are currently conducting another research project to examine this question.

Implications for Marketing Professionals

The findings of this paper offer some practical implications for companies and marketing professionals. For example, weird product designs do not need to be quickly killed in the product development process. Our research could help companies understand why certain products are perceived as weird. Furthermore, strategically designed marketing and promotion messages

could help decrease the perception of weirdness by emphasizing key information designed to facilitate the sense-making process.

More importantly, companies and marketing managers can realize the marketing value of weird product designs. With the rapid development and usage of the Internet in modern society, weird products could generate WOM and become viral online, a virtually free and effective marketing strategy to build product awareness. For example, Renova, a medium-sized European family business, launched a new product, black toilet paper, which not only gained worldwide attention and inspired online and in-person discussion, but also greatly expanded the business to 57 countries (Bart, et al., 2010). At the same time, this strategy should be applied with caution, because of its potential risks. There is no guarantee that this strategy could lead to a success, because other factors, such as the size of the company or existing brand image and reputation, could potentially come into play; weird products may have negative spillover effects on the parent brand (Lei, Dawar, & Lemmink, 2008). Furthermore, the company's target market could also play a role, because some consumers are more open to new designs (e.g., early adopter for innovations), but some are not (e.g., conservative laggards) (Rodgers, 2010). Indeed, segmentation based on variations of the product that are more or less novel, incongruent, and weird could be a highly effective marketing strategy for some categories of merchandise, such as for tattoos, bizarre movies, and particular music and visual arts.

Appendix 3.A.

Experimental Stimuli Used in Essay 3

Table 3. A. 1. Experimental stimuli used in study 1.

Experiment Condition	Stimuli	Source
Congruity		Source: https://mans.io/item/creative/zen-style-m300
Moderate Incongruity		Source: http://www.tnwelettronica.it/index.php?id_product=2784&controller=product
Extreme Incongruity		Source: http://www.dx.com/p/unique-bomb-style-rechargeable-mp3-player-speaker-w-sd-usb-fm-remote-controller-bronze-81267#.Wk1AZGQ-emE

Table 3. A. 2. Experimental stimuli used in study 2.

Experiment Condition	Stimuli	Source
Congruity		Source: http://revwatchco.com/shop/poweredspeakers/a5-plus-powered-speakers/
Moderate Incongruity		Source: https://www.pinterest.ca/pin/459507968216956802/?lp=true
Extreme Incongruity		Source: https://www.pinterest.ca/pin/151785449911006540/

Table 3. A. 3. Experimental stimuli used in study 3-1.

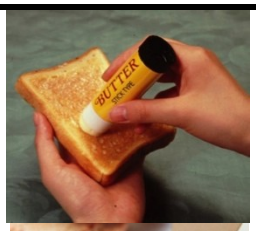
Experiment Condition	Stimuli	Source
Extreme Incongruity		Source: https://weirdamericanseatpeanutbutter.wordpress.com/2013/05/23/the-butter-stick-type/
Moderate Incongruity		Source: http://nano-graph.com/graphic/archives/category/product-design
Congruity		Source: https://www.costcobusinessdelivery.com/Lightly-Salted-Butter-Cups%2c-5g-cup%2c-720-ct.product.11611955.html

Table 3. A. 4. Experimental stimuli used in study 3-2.

Experiment Condition	Stimuli	Source
Congruity		Source: https://www.aliexpress.com/item/Stainless-Steel-Mesh-Tea-Infuser-Reusable-Tea-Strainer-Loose-Tea-Leaf-Spice-Filter/32736114985.html
Extreme Incongruity		Source: https://www.shoppersshop.com/mr-tea-infuser-perches-atop-your-tea-cup-70520139

Table 3. A. 5. Experimental stimuli used in study 4-1.

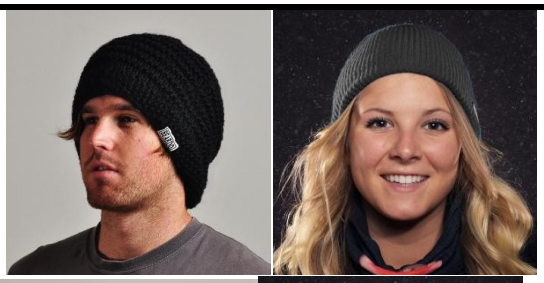
Experiment Condition	Stimuli	Source
Congruity		Source: https://www.amazon.com/Beardo-Original-Detachable-Beard-Hat/dp/B00FSAZLO2
Extreme Incongruity		Source: https://www.beardo.co.uk/products/beardo-black-attached-pink-beard

Table 3. A. 6. Experimental stimuli used in study 4-2.

Experiment Condition	Stimuli	Source
Extreme Incongruity		Revised. Source: https://laughingsquid.com/blk-a-black-mineral-water/
Congruity		Revised. Source: https://bukiajimobi.com/2016/02/02/review-blk-water/

Table 3. A. 7. Experimental stimuli used in study 4-3.

Experiment Condition	Stimuli	Source
Extreme Incongruity		Revised. Source: http://www.bigapple-shop.fr/612-clear-coffee-.html
Congruity		Revised. Source: https://www.starbucks.com/menu/drinks/bottled-drinks/bottled-frappuccino-vanilla

Table 3. A. 8. Experimental stimuli used in study 5.

Experiment Condition	Stimuli	Source
Extreme Incongruity		Source: https://www.pinterest.ca/explore/black-toilet-paper/
Congruity		Source: https://linustechtips.com/main/topic/18688-show-off-your-latest-purchase/?page=25

References

Aaker, D. A., & Keller, K. L. (1990). Consumer Evaluations of Brand Extensions. *The Journal of Marketing*, *54*(1), 27-41.

Ahluwalia, R. (2008). How far can a brand stretch? Understanding the role of self-construal. *Journal of Marketing Research*, *45*(3), 337-350.

Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, *22*(3), 261-295.

Bambauer‐Sachse, S., Hüttl, V., & Gierl, H. (2011). Can advertising elements improve consumer evaluations of brand extensions with a moderate or low fit?. *Psychology & Marketing*, *28*(2), 205-218.

Baragona, J. (2017, December 18). Twitter Has Field Day Over Trump's Weird Water Break During Speech: 'America Thirst'. Retrieved from: https://www.mediaite.com/online/twitter-has-field-day-over-trumps-weird-water-break-during-speech-america-thirst/

Barone, M. J. (2005). The interactive effects of mood and involvement on brand extension evaluations. *Journal of Consumer Psychology*, *15*(3), 263-270.

Barone, M. J., Miniard, P. W., & Romeo, J. B. (2000). The influence of positive mood on brand extension evaluations. *Journal of Consumer Research*, *26*(4), 386-400.

Bart, Y., Chandon, P., Sweldens, S., & Seabra de Sousa, R. (2010). Renova Toilet Paper: Avant-Garde Marketing In A Commoditized Category. *INSEAD*.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of general psychology, 5*(4), 323-370.

Berger, J. (2014). Word of mouth and interpersonal communication: A review and directions for future research. *Journal of Consumer Psychology, 24*(4), 586-607.

Berger, J., & Schwartz, E. M. (2011). What drives immediate and ongoing word of mouth?. *Journal of Marketing Research, 48*(5), 869-880.

Bierley, C., McSweeney, F. K., & Vannieuwkerk, R. (1985). Classical conditioning of preferences for stimuli. *Journal of Consumer Research*, *12*(3), 316-323.

Bijmolt, T. H. A., Wedel, M., Pieters, R. G. M., & DeSarbo, W. S. (1998). Judgments of brand similarity. *International Journal of Research in Marketing,* 15(3), 249-268.

Bottomley, P. A., & Holden, S. J. (2001). Do we really know how consumers evaluate brand extensions? Empirical generalizations based on secondary analysis of eight studies. *Journal of Marketing Research*, *38*(4), 494-500.

Boush, D. M., & Loken, B. (1991). A process-tracing study of brand extension evaluation. *Journal of Marketing Research, 28*(1), 16-28.

Boush, S. S., Loken, B., Gencturk, E., Crockett, S., Kennedy, E., Minshall, B., Misurell, D., Rochford, L., & Strobel, J. (1987). Affect Generalization To Similar And Dissimilar Brand Extensions. *Psychology and Marketing,* 4(3), 225–237.

Broniarczyk, S. M., & Alba, J. W. (1994). The importance of the brand in brand extension. *Journal of Marketing Research, 31*(2), 214-228.

Brown, A. D., Stacey, P., & Nandhakumar, J. (2008). Making sense of sense making narratives. *Human relations, 61*(8), 1035-1062.

Buil, I., de Chernatony, L., & Hem, L. E. (2009). Brand extension strategies: perceived fit, brand type, and culture influences. *European Journal of Marketing*, *43*(11/12), 1300-1324.

Burgoon, J. K., & Jones, S. B. (1976). Toward a theory of personal space expectations and their violations. *Human Communication Research, 2*(2), 131-146.

Campbell, M. C., & Goodstein, R. C. (2001). The moderating effect of perceived risk on consumers' evaluations of product incongruity: Preference for the norm. *Journal of Consumer Research*, *28*(3), 439-449.

Chen, Y. J., & Kirmani, A. (2015). Posting strategically: The consumer as an online media planner. *Journal of Consumer Psychology, 25*(4), 609-621.

Chernev, A., Böckenholt, U., & Goodman, J. (2015). Choice overload: A conceptual review and meta-analysis. *Journal of Consumer Psychology*, *25*(2), 333-358.

Chevalier, J. A., & Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research, 43*(3), 345-354.

Chin, W. W. (2010). How to write up and report PLS analyses. *Handbook of partial least squares*, 655-690.

Chintagunta, P. K., Gopinath, S., & Venkataraman, S. (2010). The effects of online user reviews on movie box-office performance: Accounting for sequential rollout and aggregation across local markets. *Marketing Science, 29*(5), 944–957.

Cochran, W. G. (1954). The combination of estimates from different experiments. *Biometrics*, *10*(1), 101-129.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155-159.

Cohen, J. B., & Basu, K. (1987). Alternative models of categorization: Toward a contingent processing framework. *Journal of Consumer Research*, *13*(4), 455-472.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, *82*(6), 407-428.

Coltman, T., Devinney, T. M., Midgley, D. F., & Venaik, S. (2008). Formative versus reflective measurement models: Two applications of formative measurement. *Journal of Business Research, 61*(12), 1250-1262.

Cooper, H., Hedges, L. V., & Valentine, J. C. (Eds.). (2009). *The handbook of research synthesis and meta-analysis*. Russell Sage Foundation.

Craig-Lees, M. (2001). Sense making: Trojan horse? Pandora's box?. *Psychology & Marketing, 18*(5), 513-526.

Csikszentmihalyi, M. (1997). *Finding flow: The psychology of engagement with everyday life*. Basic Books.

Churchill Jr, G. A. (1979). A Paradigm For Developing Better Measures Of Marketing Constructs. *Journal of marketing research*, 16(1), 64-73.

Czellar, S. (2003). Consumer attitude toward brand extensions: an integrative model and research propositions. *International Journal of Research in Marketing*, *20*(1), 97-115.

Davis, C. G., Nolen-Hoeksema, S., & Larson, J. (1998). Making sense of loss and benefiting from the experience: two construals of meaning. *Journal of personality and social psychology, 75*(2), 561-574.

Dermody, K. C. (2016, December 19). 22 of the World's Most Bizarre Tourist Attractions. Retrieved from: https://www.tripstodiscover.com/22-of-the-worlds-most-bizarre-tourist-attractions/

Diamantopoulos, A., Smith, G., & Grime, I. (2005). The impact of brand extensions on brand personality: experimental evidence. *European Journal of Marketing*, *39*(1/2), 129-149.

Diamantopoulos, A., & Winklhofer, H. M. (2001). Index construction with formative indicators: An alternative to scale development. *Journal of marketing research, 38*(2), 269-277.

Dickerson, B. (2016, December 30). 22 Irresistibly Weird Products To Buy On Amazon. Retrieved from: https://www.buzzfeed.com/betsydickerson/amazingly-weird-products-from-amazon?utm_term=.siY6Z7JKx#.pkABNEZmn

Dimitriu, R., Dimitriu, R., Warlop, L., Warlop, L., Samuelsen, B. M., & Samuelsen, B. M. (2017). Brand extension similarity can backfire when you look for something specific. *European Journal of Marketing*, *51*(5/6), 850-868.

Echambadi, R., Arroniz, I., Reinartz, W., & Lee, J. (2006). Empirical generalizations from brand extension research: How sure are we?. *International Journal of Research in Marketing, 23*(3), 253-261.

Einstein, G. O., McDaniel, M. A., & Lackey, S. (1989). Bizarre imagery, interference, and distinctiveness. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*(1), 137-146.

Fedorikhin, A., Park, C. W., & Thomson, M. (2008). Beyond fit and attitude: The effect of emotional attachment on consumer responses to brand extensions. *Journal of Consumer Psychology*, *18*(4), 281-291.

Festinger, L. (1957). *A Theory of Cognitive Dissonance*. Stanford University Press.

Fiske, S. T. (1982). Schema-triggered affect: Applications to social perception. In *Affect and cognition: 17th Annual Carnegie Mellon symposium on cognition* (pp. 55-78). Hillsdale: Lawrence Erlbaum.

Fiske, S. T., & Pavelchak, M. A. (1986). Category-based versus piecemeal-based affective responses: Developments in schema-triggered affect. In R. M. Sorrentino & E. T. Higgins (Eds.), *The Handbook Of Motivation And Cognition* (pp.167−203). New York: Guilford.

Fleming, J. H., Darley, J. M., Hilton, J. L., & Kojetin, B. A. (1990). Multiple audience problem: A strategic communication perspective on social perception. *Journal of personality and social psychology, 58*(4), 593-609.

Gephart, R. P. (1993). The textual approach: Risk and blame in disaster sense making. *Academy of Management Journal, 36*(6), 1465-1514.

Gierl, H., & Huettl, V. (2011). A closer look at similarity: The effects of perceived similarity and conjunctive cues on brand extension evaluation. *International Journal of Research in Marketing*, *28*(2), 120-133.

Godes, D., & Mayzlin, D. (2009). Firm-created word-of-mouth communication: Evidence from a field test. *Marketing Science, 28(*4), 721–739.

Hagtvedt, H., & Patrick, V. M. (2008). Art and the brand: The role of visual art in enhancing brand extendibility. *Journal of Consumer Psychology, 18*(3), 212-222.

Hayes, A. F. (2013). *Introduction to Mediation, Moderation, and Conditional Process Analysis. A Regression-Based Approach*. New York: Guilford.

Heider, F. (1958). *The psychology of interpersonal relations*. NY: John Wiley.

Higgins, J., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta‐analysis. *Statistics in Medicine*, *21*(11), 1539-1558.

Hopkins, B. (2017, December 17). 5 weird Christmas foods from around the world. Retrieved from: http://metro.co.uk/2017/12/17/5-weird-christmas-foods-from-around-the-world-7138415/?ito=cbshare

Huang, Y., Jia, Y., & Wyer, R. S. (2017). The effects of physical distance from a brand extension on the impact of brand-extension fit. *Psychology & Marketing*, *34*(1), 59-69.

Hunt, R. R., & McDaniel, M. A. (1993). The enigma of organization and distinctiveness. *Journal of Memory & Language, 32*(4), 421-445.

Jacob, S. (2015). 27 of the strangest world records ever set. Retrieved from: http://www.businessinsider.com/the-most-bizarre-world-records-ever-set-2015-10

Jhang, J. H., Grant, S. J., & Campbell, M. C. (2012). Get it? Got it. Good! Enhancing new product acceptance by facilitating resolution of extreme incongruity. *Journal of Marketing Research, 49*(2), 247-259.

Kalamas, M., Cleveland, M., Laroche, M., & Laufer, R. (2006). The critical role of congruency in prototypical brand extensions. *Journal of Strategic Marketing*, *14*(3), 193-210.

Kanouse, D. E., & Hanson, L. R.(1987). Negativity in evaluations. In E. E. Jones, Kanouse, D. E., H. H. Kelley, R. E. Nisbett, S. Valins, B. Weiner (Eds.)*, Attribution: Perceiving the causes of behavior* (pp. 47-62). NJ: General Learning Press.

Kapoor, H., & Heslop, L. A. (2009). Brand positivity and competitive effects on the evaluation of brand extensions. *International Journal of Research in Marketing*, *26*(3), 228-237.

Kaul, T. M., & Naire, A. (2012). FMCG Brand Extensions Five Times More Successful Than New Product Launches In India. Retrieved from http://www.nielsen.com/in/en/press-room/2012/brand-stretching0.html.

Keller, K. L. (1993). Conceptualizing, measuring, and managing customer-based brand equity. *The Journal of Marketing*, *57*(1), 1-22.

Keller, K. L. (1998). *Strategic brand management: Building, measuring, and managing brand equity*. New Jersey.

Keller, K. L., & Aaker, D. A. (1992). The effects of sequential introduction of brand extensions. *Journal of Marketing Research, 3*(3), 35-50.

Keller, K. L., & Lehmann, D. R. (2006). Brands and branding: Research findings and future priorities. *Marketing Science*, *25*(6), 740-759.

Kim, H., & John, D. R. (2008). Consumer response to brand extensions: Construal level as a moderator of the importance of perceived fit. *Journal of Consumer Psychology*, *18*(2), 116-126.

Klara, R. (2013). The Best (and Worst) Brand Extensions. Retrieved from: http://www.adweek.com/brand-marketing/best-and-worst-brand-extensions-146966/

Knoedler, A. J., Hellwig, K. A., & Neath, I. (1999). The shift from recency to primacy with increasing delay. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 25*(2), 474-487.

Kohler, W. (1929). *Gestalt Psychology*, New York: Liveright.

Kuo, A., & Rice, D. H. (2015). The impact of perceptual congruence on the effectiveness of cause-related marketing campaigns. *Journal of Consumer Psychology, 25*(1), 78-88.

Lane, V. R. (2000). The impact of ad repetition and ad content on consumer perceptions of incongruent extensions. *Journal of Marketing*, *64*(2), 80-91.

Latimer, R. (2015). The Peculiarly Persistent Pleasantness of Bizarre Experiences. *ACR North American Advances*.

Latimer, R., & Raghubir, P. (2012). The Best and the Bizarre: Prototype and Exemplar-Based Retrospective Evaluations of Experiences. *ACR North American Advances*.

Lefkoff-Hagius, R., & Mason, C. H. (1993). Characteristic, beneficial, and image attributes in consumer judgments of similarity and preference. *Journal of Consumer Research, 20*(1), 100-110.

Lei, J., Dawar, N., & Lemmink, J. (2008). Negative spillover in brand portfolios: exploring the antecedents of asymmetric effects. *Journal of Marketing, 72*(3), 111-123.

Loken, B., Barsalou, L. W., & Joiner, C. (2008). Categorization Theory And Research In Consumer Psychology. In C. P. Haugtvedt (Ed.), *Handbook of Consumer Psychology* (pp. 133-65), Hoboken: Lawrence Erlbaum Associates.

Loken, B., & John, D. R. (1993). Diluting brand beliefs: when do brand extensions have a negative impact?. *The Journal of Marketing*, 57(3), 71-84.

Liu, L., Li, C., & Zhu, D. (2012). A new approach to testing nomological validity and its application to a second-order measurement model of trust. *Journal of the Association for Information Systems, 13*(12), 950-975.

Liu, X., & Hu, J. (2012). Adolescent evaluations of brand extensions: The influence of reference group. *Psychology & Marketing*, *29*(2), 98-106.

Lorayne, H., & Lucas, J. (1974). *The memory book*. New York Stein & Day.

Mandler, G. (1982). The Structure of Value: Accounting For Taste. In M. S. Clark and S. T. Fiske (Eds.), *Affect and Cognition: The Seventeenth Annual Carnegie Symposium on Cognition* (pp. 3-36). Hillsdale: Erlbaum.

Maoz, E., & Tybout, A. M. (2002). The moderating role of involvement and differentiation in the evaluation of brand extensions. *Journal of Consumer Psychology*, *12*(2), 119-131.

Martin, I. M., Stewart, D. W., & Matta, S. (2005). Branding strategies, marketing communication, and perceived brand meaning: The transfer of purposive, goal–oriented brand meaning to brand extensions. *Journal of the Academy of Marketing Science*, *33*(3), 275-294.

Martinez, E., & Pina, J. M. (2003). The negative impact of brand extensions on parent brand image. *Journal of Product & Brand Management*, *12*(7), 432-448.

Martinez, E., Polo, Y., & De Chernatony, L. (2008). Effect of brand extension strategies on brand image: A comparative study of the UK and Spanish markets. *International Marketing Review, 25*(1), 107-137.

Mathur, P., Jain, S. P., & Maheswaran, D. (2012). Consumers' implicit theories about personality influence their brand personality judgments. *Journal of Consumer Psychology*, *22*(4), 545-557.

McCarthy, M. S., Heath, T. B., & Milberg, S. J. (2001). New brands versus brand extensions, attitudes versus choice: Experimental evidence for theory and practice. *Marketing Letters*, *12*(1), 75-90.

McDaniel, M. A., & Einstein, G. O. (1986). Bizarre imagery as an effective memory aid: The importance of distinctiveness. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12*(1), 54-65.

McDaniel, M. A., DeLosh, E. L., & Merritt, P. S. (2000). Order information and retrieval distinctiveness: Recall of common versus bizarre material. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(4), 1045-1056.

McDaniel, M. A., Dornburg, C. C., & Guynn, M. J. (2005). Disentangling encoding versus retrieval explanations of the bizarreness effect: Implications for distinctiveness. *Memory & Cognition, 33*(2), 270-279.

McGraw, A. P., Warren, C., Williams, L. E., & Leonard, B. (2012). Too close for comfort, or too far to care? Finding humor in distant tragedies and close mishaps. *Psychological Science, 23*(10), 1215-1223.

McSweeney, F. K., & Bierley, C. (1984). Recent Developments in Classical Conditioning. *Journal of Consumer Research,* 11(September), 619-631.

Meyers-Levy, J., & Tybout, A. M. (1989). Schema congruity as a basis for product evaluation. *Journal of Consumer Research*, *16*(1), 39-54.

Milberg, S. J., Sinn, F., & Goodstein, R. C. (2010). Consumer reactions to brand extensions in a competitive context: does fit still matter?. *Journal of Consumer Research*, *37*(3), 543-553.

Milberg, S. J., Goodstein, R. C., Sinn, F., Cuneo, A., & Epstein, L. D. (2013). Call back the jury: Reinvestigating the effects of fit and parent brand quality in determining brand extension success. *Journal of Marketing Management*, *29*(3-4), 374-390.

Monga, A. B., & John, D. R. (2006). Cultural differences in brand extension evaluation: The influence of analytic versus holistic thinking. *Journal of Consumer Research*, *33*(4), 529-536.

Morrin, M. (1999). The impact of brand extensions on parent brand memory structures and retrieval processes. *Journal of Marketing Research*, *36*(4), 517-525.

Mulligan, N. W. (2000). Perceptual interference at encoding enhances item-specific encoding and disrupts relational encoding: Evidence from multiple recall tests. *Memory & Cognition, 28*(4), 539-546.

Mumford, M. D. (2003). Where have we been, where are we going? Taking stock in creativity research. *Creativity Research Journal, 15*(2-3), 107-120.

Na, J., & Kitayama, S. (2011). Spontaneous trait inference is culture-specific: Behavioral and neural evidence. *Psychological Science*, *22*(8), 1025-1032.

Nan, X. (2006). Affective cues and brand‐extension evaluation: Exploring the influence of attitude toward the parent brand and attitude toward the extension ad. *Psychology & Marketing*, *23*(7), 597-616.

Newcomb, T. M. (1953). An Approach to the Study of Communicative Acts. *Psychological Review, 60*(6), 393-404.

Noseworthy, T. J., Muro, F. D., & Murray, K. B. (2014). The Role of Arousal in Schema-Based Product Evaluations. *Journal of Consumer Research, 41*(4), 1108-1126.

Noseworthy, T. J., Murray, K. B., & Di Muro, F. (2017). When Two Wrongs Make a Right: Using Conjunctive Enablers to Enhance Evaluations for Extremely Incongruent New Products. *Journal of Consumer Research*, 1379-1396.

Oakley, J. L., Duhachek, A., Balachander, S., & Sriram, S. (2007). Order of entry and the moderating role of comparison brands in brand extension evaluation. *Journal of Consumer Research*, *34*(5), 706-712.

Osgood, C. E., & Tannenbaum, P. H. (1955). The principle of congruity in the prediction of attitude change. *Psychological review*, *62*(1), 42-55.

Park, C. W., Milberg, S., & Lawson, R. (1991). Evaluation of brand extensions: the role of product feature similarity and brand concept consistency. *Journal of Consumer Research*, *18*(2), 185-193.

Peracchio, L. A., & Tybout, A. M. (1996). The moderating role of prior knowledge in schema-based product evaluation. *Journal of Consumer Research*, *23*(3), 177-192.

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology*, *19*, 123-205.

Piaget, J., & Cook, M. T. (1952). *The origins of intelligence in children*. New York, NY: International University Press.

Pina, J. M., Martinez, E., De Chernatony, L., & Drury, S. (2006). The effect of service brand extensions on corporate image: an empirical model. *European Journal of Marketing, 40*(1/2), 174-197.

Rimé, B. (2009). Emotion elicits the social sharing of emotion: Theory and empirical review. *Emotion Review, 1*, 60–85.

Rogers, E. M. (2010). *Diffusion of innovations*. Simon and Schuster.

Rosenberg, M. J. (1956). Cognitive structure and attitudinal affect. *The Journal of Abnormal and Social Psychology*, *53*(3), 367-372.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, *86*(3), 638-641.

Rossiter, J. R. (2005). Reminder: a horse is a horse. *International Journal of Research in Marketing, 22*(1), 23-25.

Scheibehenne, B., Greifeneder, R., & Todd, P. M. (2010). Can there ever be too many options? A meta-analytic review of choice overload. *Journal of Consumer Research*, *37*(3), 409-425.

Sergius Koku, P. (1995). On bizarreness in the entertainment industry. Journal of *Services Marketing, 9*(4), 19-30.

Shaw, B. (2014, March 30). The 24 Most Bizarre Injuries in Baseball History. Retrieved from: http://mentalfloss.com/article/29000/24-most-bizarre-injuries-baseball-history

Shen, Y. C., Bei, L. T., & Chu, C. H. (2011). Consumer evaluations of brand extension: The roles of case‐based reminding on brand‐to‐brand similarity. *Psychology & Marketing*, *28*(1), 91-113.

Shimp, T. A. (1981). Attitude toward the ad as a mediator of consumer brand choice. *Journal of Advertising*, *10*(2), 9-48.

Shine, B. C., Park, J., & Wyer Jr, R. S. (2007). Brand synergy effects in multiple brand extensions. *Journal of Marketing Research*, *44*(4), 663-670.

Sichtmann, C., Sichtmann, C., Schoefer, K., Schoefer, K., Blut, M., Blut, M., ... & Kemp, C. J. (2017). Extending service brands into products versus services: Multilevel analyses of key success drivers. *European Journal of Marketing, 51*(1), 200-218.

Smith, D. C., & Park, C. W. (1992). The effects of brand extensions on market share and advertising efficiency. *Journal of Marketing Research*, *29*(3), 296-313.

Smith, G. (2004). Brand image transfer through sponsorship: A consumer learning perspective. *Journal of Marketing Management*, *20*(3-4), 457-474.

Sood, S., & Keller, K. L. (2012). The effects of brand name structure on brand extension evaluations and parent brand dilution. *Journal of Marketing Research*, *49*(3), 373-382.

Steel, P. (2006). MetaExcel (Version 10.0). [Computer software]. http://www.ucalgary.ca/steel/procrastinus/meta/meta.html.

Sujan, M. (1985). Consumer knowledge: Effects on evaluation strategies mediating consumer judgments. *Journal of Consumer Research, 31*(1), 31-46.

Sunde, L., & Brodie, R. J. (1993). Consumer evaluations of brand extensions: Further empirical results. *International Journal of Research in Marketing*, *10*(1), 47-53.

Tauber, E. M. (1981). Brand franchise extension: new product benefits from existing brand names. *Business Horizons*, *24*(2), 36-41.

Torelli, C. J., & Ahluwalia, R. (2011). Extending culturally symbolic brands: A blessing or a curse?. *Journal of Consumer Research, 38*(5), 933-947.

Trusov, M., Bucklin, R. E., & Pauwels, K. (2009). Effects of word-of-mouth versus traditional marketing: Findings from an Internet social networking site. *Journal of Marketing, 73*(5), 90–102.

Tversky, A. (1977). Features of similarity. *Psychological review*, *84*(4), 327-352.

Van den Bulte, C., & Wuyts, S. (2009). Leveraging customer networks. In Jerry Yoram Wind, & Paul Kleindorfer (Eds.), *The network challenge: Strategy, profit and risk in an interlinked world* (pp. 243–258). Upper Saddle River, NJ: Wharton School Publishing.

Völckner, F., & Sattler, H. (2006). Drivers Of Brand Extension Success. *Journal of Marketing, 70*(2), 18-34.

Völckner, F., & Sattler, H. (2007). Empirical generalizability of consumer evaluations of brand extensions. *International Journal of Research in Marketing, 24(*2), 149-162.

Völckner, F., Sattler, H., & Kaufmann, G. (2008). Image feedback effects of brand extensions: Evidence from a longitudinal field study. *Marketing Letters, 19*(2), 109-124.

Waddill, P. J., & McDaniel, M. A. (1998). Distinctiveness effects in recall: Differential processing or privileged retrieval? *Memory & Cognition, 26*(1), 108-120.

Walchli, S. B. (2007). The effects of between‐partner congruity on consumer evaluation of co‐branded products. *Psychology & Marketing, 24*(11), 947-973.

Warren, C., & Campbell, M. C. (2014). What makes things cool? How autonomy influences perceived coolness. *Journal of Consumer Research, 41*(2), 543-563.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal Of Personality And Social Psychology, 54*(6), 1063-70.

Weick, K. (1995). *Sense making in organizations*. Thousand Oaks, CA: Sage.

Weick, K. E., Sutcliffe, K. M., & Obstfeld, D. (2005). Organizing and the process of sense making. *Organization Science, 16*(4), 409-421.

Woodside, A. G. (2001). Sense making in marketing organizations and consumer psychology: theory and practice. *Psychology & Marketing, 18*(5), 415-421.

Wright, P. (1975). Consumer choice strategies: Simplifying vs. optimizing. *Journal of Marketing Research, 11*(February), 60-67.

Yeo, J., & Park, J. (2006). Effects of parent-extension similarity and self regulatory focus on evaluations of brand extensions. *Journal of Consumer Psychology*, *16*(3), 272-282.

Yeung, C. W., & Wyer Jr, R. S. (2005). Does loving a brand mean loving its products? The role of brand-elicited affect in brand extension evaluations. *Journal of Marketing Research*, *42*(4), 495-506.

Yorkston, E. A., Nunes, J. C., & Matta, S. (2010). The malleable brand: The role of implicit theories in evaluating brand extensions. *Journal of Marketing*, *74*(1), 80-93.

Zhang, S., & Sood, S. (2002). "Deep" and "surface" cues: Brand extension evaluations by children and adults. *Journal of Consumer Research,* 29(1), 129-141.