DRUID
society

Paper to be presented at the

DRUID Society Conference 2014, CBS, Copenhagen, June 16-18

The wisdom of conformity and capabilities: How firm experience predicts the reception of typical and atypical products

Mirva Peltoniemi
Jyväskylä University School of Business and Economics
Strategy and Entrepreneurship
mirva.peltoniemi@jyu.fi

Henri Schildt
Aalto University School of Business
Management and International Business
henri.schildt@aalto.fi

Abstract
Is it better for firms in cultural industries to develop products typical or atypical to a given category? Prior research generally holds that deviating from the category norm results in more negative evaluations from the audience. We examine how firm capabilities can allow them to craft novel yet appealing products. Based on data on video games we find that deviance from expectations leads to inferior performance as expected, particularly for companies with focused experience in the product category. However, we find that deviance leads to better product reception for companies that have accumulated broad experience from diverse product categories. These findings highlight the contingency of categorical conformity on firm capabilities and the contingency of firm capabilities on product strategy.

Jelcodes:L10,Z10

The wisdom of conformity and capabilities: How firm experience predicts the reception of typical and atypical products

ABSTRACT

Is it better for firms in cultural industries to develop products typical or atypical to a given category? Prior research generally holds that deviating from the category norm results in more negative evaluations from the audience. We examine how firm capabilities can allow them to craft novel yet appealing products. Based on data on video games we find that deviance from expectations leads to inferior performance as expected, particularly for companies with focused experience in the product category. However, we find that deviance leads to better product reception for companies that have accumulated broad experience from diverse product categories. These findings highlight the contingency of categorical conformity on firm capabilities and the contingency of firm capabilities on product strategy.

INTRODUCTION

Research on categories finds that products deviating from existing categories tend to receive a negative response from the audience (Hsu, 2006; Zuckerman, 1999). Research on categorical conformity tends to focus predominantly on the "audience side" evaluations of products and other entities (Bowers, 2014). Thus it is not surprising that the literature has ignored possible heterogeneity in firm capabilities to create deviant yet appealing products. Many anecdotal examples show that creative individuals and teams can launch innovative products that are widely accepted despite their deviance from existing category norms. For example, the film Titanic launched in 1997 mixed characteristics of romantic, époque, action and catastrophe genres and yet ended up a huge success both at the box office and the Academy awards. In this case a capable team was able to reach success with a deviant product. We believe that a closer attention to firm capabilities could help elaborate the boundary conditions under which categorical deviance is detrimental. Therefore we ask: How do firm capabilities, conceived in terms of firm experience, influence the critical evaluation of typical and atypical products?

Prior research has examined product-category fit largely independently from the firms that actually produce the products. We suggest that the optimal positioning of products vis-à-vis existing categories is likely to be contingent on firm capabilities. Strategy literature broadly suggests firms develop capabilities through experience, leading to heterogeneous performance (see e.g. Anand & Khanna, 2000). Research on capabilities finds that experience from particular geographic areas improves future performance in them (Brouthers et al., 2008; Uhlenbruck, 2004), and experience from different technological and product-market domains has complementary benefits (Nerkar & Roberts, 2004). Our

paper is premised on the assumption that firm experience gained through past product launches in various product categories influences firm capabilities and therefore the firm's ability to successfully carry out diverse product strategies.

In order to unpack the relationships between firm capabilities and categorical conformity and deviance, we examine two specific types of experience: the depth of category-specific experience and the scope of experience across multiple categories. We hypothesize that categorical conformity is generally beneficial for companies, and particularly so the more experience they have in creating new products in a specific category. However, we argue that some firms are more capable in introducing attractive deviant products and reason such a capability to originate from a greater scope of experience obtained by launching products in a number of different categories.

To test our hypotheses, we use product data from the console games industry. Our dataset consists of 1428 video games, their critical evaluations, genre data and textual product synopses. We test our hypotheses using a mixed (multi-level) regression model that account for product-specific features and developer firm random effects. The analysis supports our main hypotheses. We find that in general the conformity to typical category content appears to improve the audience response to a product. Category-specific firm experience appears to be beneficial for products that are typical to their category. Our data also suggests that more diverse scope of experience is positively associated with audience evaluation. As expected, this positive relationship is particularly pronounced for atypical products.

Our findings help bridge the categories literature and the capabilities literature by revealing how benefits from capabilities and categorical conformity are contingent on one other. We

show that firms with a wide experience repertoire from multiple categories are better able to create successful products that deviate from what is typically expected in their category or categories. More broadly, our research shows that since products have difficult-to-observe heterogeneity that influences the effects of conformity, we ought also to attend to the characteristics of the entities that craft those products.

CATEGORY CONFORMITY AND FIRM CAPABILITIES

This section provides the necessary theoretical background for our hypotheses building by briefly introducing key findings related to product category conformity and the role of experience in the development of firm capabilities.

Audience response to category conformity and deviance

Categories research has established that a product deviating from the norms of its category tends to receive a less positive response from the audience. Underlying this finding are two theoretical explanations: audiences have difficulties in understanding atypical products and atypical products fail to meet audience expectations (Hsu, 2006; Zuckerman, 1999). Category deviance has been measured by multiple-category membership (e.g. Hsu, 2006; Negro & Leung, 2013; Ruef & Patterson, 2009; Vergne, 2012), disagreement over category membership (e.g. Hsu, 2006; Hsu et al., 2009), incoherence (Zuckerman, 2004) and atypicality of product or firm attributes in relation to its category (Durand et al., 2007; Jonsson et al., 2009; Smith, 2011). Such a variety of measurement methods has enabled

further scrutiny of how different kinds of deviance, and their co-existence, affect audience response.

Even though multiple-category membership (i.e. category straddling) is found to lead to less favorable evaluation, audience response improves with more common category combinations (Ruef & Patterson, 2009) and with more common category straddling (Rao et al., 2005). Being associated with several categories may also improve audience response when a firm is a member of a stigmatized category (Vergne, 2012). In this way multiple-category membership can neutralize connotations related to particular categories. However, recent research on wine producers shows that products by multiple-category firms receive more negative evaluations when product quality differences (blind wine testing scores) are controlled for (Negro & Leung, 2013). This means that being associated with several categories brings about ambiguity that cannot be offset by matching the quality of single-category actors. However, we do not know whether experience accumulated by operating in multiple categories has an effect on a firm's ability to target audience expectations.

Disagreement over which category a product or firm should be assigned to is associated with more negative evaluations. This has been shown by comparing the views of different audience members (Hsu, 2006) and the views of audience and producer (Hsu et al., 2009; Zuckerman, 1999). Coherence is related to audience agreement and is measured as the extent of overlap in the evaluated by the evaluators (Zuckerman, 2004). Coherent firms thus attract an audience that share the population of firms they are interested in. Incoherent firms are found more volatile (Zuckerman, 2004). Difficulties in assigning categories hence lead to more negative evaluations. Also this research stream has so far omitted category-related firm experience.

Some recent studies have measured category conformity as typicality of product attributes and found that deviance may also garner a positive response (Durand et al., 2007; Smith, 2011). This research has recorded "code-violating changes", i.e. deviation from the attributes of average category members (Durand et al., 2007), the extent of overlap in the description of a firm and its fellow category members (Smith, 2011) and similarity in firm size (Jonsson et al., 2009). Contrary to other measures of deviance, code-violating changes are rewarded as long as they are performed less frequently than competitors do (Durand et al., 2007), and firms with atypical descriptions benefit from amplification and buffering effects (Smith, 2011). Such firms receive a larger positive reaction for positive news and a smaller negative reaction for negative news from their audience. The findings by Durand et al. (2007) suggest deviance from category expectations may in some cases be beneficial. However, they do not shed light on the potential effects of experience from the violated category or experience from other categories on how audiences evaluate the product.

Much of existing research on product categories relates to cultural industries (see Peltoniemi, 2014 for a recent review), which are paradoxically divided into relatively widely recognized categories (often called "genres") yet demand constant innovation to please diverse and fickle consumer tastes (Aksoy & Robins, 1992; Mora, 2006; Power & Scott, 2004). Despite relatively stable genres, the producers operate in a product space with virtually unlimited aesthetic choices (Thompson et al., 2007). Many researchers have pointed out that audiences expect both novelty and familiarity from new cultural goods to be able to both understand and enjoy the products (Alvarez et al., 2005; Cillo et al., 2010). To generate familiarity, cultural producers use serials and superstars in addition to genres (Aksoy & Robins, 1992; Hesmondhalgh, 2002)

In order to understand how companies in cultural industries cope with audience demands, we turn our attention to firm capabilities. By attending to capabilities we aim at explaining when and why firms may be able to introduce atypical yet successful new products. In cultural industries it is so common for products to straddle multiple categories that it cannot be seen as deviance. Therefore we look into the typicality of products in relation to their categories.

Firm capabilities and category-related experience

Research on capabilities assumes that they are accumulated through experience (Helfat et al., 2007). The longer a firm has operated in a specific technological field, product-market domain or geographical area, the more capabilities it has acquired in relation to that particular mode of operation. In addition to tenure in one business, Eisenhardt and Martin (2000) highlight the importance of related but different experiences in developing capabilities. In the same vein, Bowman and Hurry (1993) state that the variety of past actions creates the firm's portfolio of options for the future. Different types of capabilities are accumulated through (1) the tenure of experience in a single business and (2) broad experience across several businesses. These can be called specialization and scope, respectively.

Benefits of specialization are widely documented (Huber, 1991; Levitt & March, 1988). Research on specialization has shown that experience from a particular geographic area improves future performance in it (Brouthers et al., 2008; Uhlenbruck, 2004), and that alliance experience improves the success rates of future alliances (Rothaermel & Deeds, 2006; Sampson, 2005). Therefore, gaining experience from a particular activity enables a

firm to perform better in it in the future. This translates into category-specific experience improving a firm's performance in the category in question in the future. There is preliminary empirical evidence of this effect relating to film genres (Shamsie et al., 2009).

Several studies have argued that a broad scope of experience improves a firm's ability to innovate, particularly in technology-based industries (Suzuki & Kodama, 2004; Miller, 2006; Hargadon & Sutton, 1997). Broader scope of expertise enables firms to combine knowledge across domains in order to solve problems (Hargadon & Sutton, 1997) and to absorb new knowledge more rapidly (Schildt et al., 2012). In general, since management theory tends to see problem solving as the matching of problems and solutions the ability of a firm to solve problems would be increased by the scope of experience (Cohen et al., 1972).

Vindicating these arguments, Nerkar and Roberts (2004) found that experience in distinct domains was not merely complementary, but that the value of one kind of experience may be increased by the presence of another kind. Similarly, internationalization capabilities develop through not just the quantity of experience, but also through the number of different countries a firm has operated in (Brouthers et al., 2008). Experience from entering new technological niches improves a firm's ability to do so in the future (King & Tucci, 2002). These findings suggest that gaining experience from diverse product categories has value for a firm, and makes it better equipped to compete in the future.

HYPOTHESES

To answer our research question, we outline three sets of hypotheses linking firm experience and category conformity to audience response, measured through critical reviews. Our Hypotheses 1-3 are not particularly novel, as they build heavily on existing findings in the literature. The novel contribution of the paper lies in Hypotheses 4 and 5, which examine the role of firm capabilities on the effects of categorical conformity and deviance.

Firm experience and critical response

The very basic tenet of organizational learning is that domain-specific experience improves a firm's performance in the domain in question (Argyris & Schön, 1978). This means that there are domain-specific capabilities that accumulate through tenure. We hypothesize that depth of past experience in a product category by the firm has a positive effect on the reception of its product.

H1: Depth of category-specific firm experience is positively related to critical appraisals of a new product.

Capabilities research also highlights the importance of the variety of experience. Related but different experiences enable firms to develop capabilities and open new avenues for future actions (Bowman & Hurry, 1993; Eisenhardt & Martin, 2000). We hypothesize that firm experience from multiple categories has a positive effect on the reception of its product.

H2: Diverse firm experience from multiple categories is positively related to critical appraisals of a new product.

Category conformity and critical response

Research holds that products conforming to established categories receive more positive evaluations. They offer the benefits of convenient sense-making by the audience and a good fit with category expectations (Glynn & Abzug, 2002; Zuckerman, 1999). Moreover, research on cultural industries shows that conforming to a category is perceived positively by gatekeepers (Mauws, 2000). Hence, we propose that products that are similar to past products within their categories tend to receive better critical reviews.

H3: Conformity to typical category content is positively related to critical appraisals for a new product.

Firm experience and the effects of category conformity and deviance

Firms gain domain-specific capabilities through specialization (cf. Brouthers et al., 2008; Rothaermel & Deeds, 2006). Firms with plenty of category-specific experience have learned to produce attributes that are expected from the audience consuming category products. This category-specific experience can be leveraged to the fullest when the product developed conforms to typical category content. Therefore, we hypothesize that the benefits from depth of experience in the category will be higher for products that are more typical category members.

H4: Conformity to typical category content increases the positive effects from greater depth of category-specific firm experience on critical appraisals.

Experience from multiple different domains builds a firm's dynamic capabilities and equips it with the skills required in complex environments that are in constant flux, both through an improved ability to combine knowledge across domains (Granstrand, 1998; Hargadon & Sutton, 1997; Suzuki & Kodama, 2004) and an improved ability to absorb and utilize new knowledge (Schildt et al., 2012; Schilling et al., 2003). In this way firms accumulate a larger repertoire of skills that enable them to combine elements across product categories. Such firms should be better equipped to create radical new innovations and introduce novel products. In order to accommodate innovations, however, firms may need to launch products that deviate from the expectations of any existing category. We thus hypothesize firm with greater scope of experience to benefit from deviating from established category content.

H5: Deviance from typical category content increases the positive effects from greater scope of firm experience from multiple categories on critical appraisals.

DATA AND METHOD

To study the critical evaluations of cultural products, we focused on the video game industry. Our data comes from the Metacritic database (http://www.metacritic.com/games). The Metacritic data is based on game titles. For each game title the following information is available: release date, name of developer firm, name of publisher firm, average score of reviews by critics and a short textual synopsis of the content of the game. We analyze the critical acclaim of games published between 2002 and 2011, although games dating back to

11

1996 are used to calculate various independent and control variables experience and past performance.

We sampled a set of games published for the major consoles of the fifth, sixth and seventh generation: Dreamcast, GameCube, Nintendo 64, PlayStation, PlayStation 2, PlayStation 3, Wii, Xbox and Xbox360. This selection was made for three reasons: (1) the market structure among console manufacturers and the technological environment for developers has remained similar during the chosen era, (2) games released for major consoles represent mainstream consumer products, and (3) the number of games published for these consoles and covered in the database is considerable.

To analyze game categories and content typical or atypical to them, we used genre data and the official textual synopses of the games. Genre data has been often used in category research to assign products into categories (cf. Hsu, 2006; Hsu et al., 2012a; Zuckerman & Kim, 2003). Various forms of content analysis have been used to assess the characteristics of cultural goods (e.g. Alexander, 1996; Cappetta et al., 2006). We chose to use a quantitative word based method with the goals of distinguishing words typical or atypial for different genres.

In addition to the Metacritic database, genre data was retrieved from the Mobygames database (mobygames.org) in order to increase the robustness of our analysis. The Metacritic categorization includes 36 genres and each game is assigned one genre. In the Mobygames categorization there are 8 main genres and 72 sub-genres, and games can be assigned more than one main and sub-genre. We decided to include all genre information relating to a game into the analysis. This means that each game can be assigned to several

categories. However, we cleaned small differences in terminology between the sources deciding that, for example, "role-playing" in Metacritic and "Role-Playing (RPG)" in Mobygames are the same genre. In consequence we eliminated the possibility of artificially creating additional genres due to different spellings.

The game synopses were cleaned through established content analysis methods. First, all punctuation marks were removed and all text was reduced to lower case. Second, so-called STOP-words were removed (see Salton, 1971). This list contains the most common words in the English language, such as "is", "and" and "the" (full list available here: ftp://ftp.cs.cornell.edu/pub/smart/english.stop). Third, the remaining words were stemmed, i.e. inflected words were reduced to their stem (see Porter, 1980). For example, plurals were turned into singulars and verbs were changed into present tense.

To increase the robustness of our findings, we controlled for the number of prior game releases of each developer firm. We dropped all the games that were produced by completely new entrants, as no information was available on the scope or depth of their prior experience. Including the new developers in models did not change the results or their significance, but increased variance. Moreover, we dropped all games that had less than 10 reviews at the Metacritic database because (1) such games were niche products that had received very limited attention and (2) those games were commonly reviewed by less professional outlets (i.e. hobbyist websites rather than professional magazines), wildly increasing the variance in ratings. Dropping games without substantial number of reviews reduced the sample by less than 10%. The highest number of different professional critic scores recorded for a single game was 107 (Heavy Rain on PlayStation 3, developed by Quantic Dream).

Dependent variable

We measured the critical reception of games with the average of scores given by professional critics for a game, a variable tracked by Metacritic (Average critic score). The database includes a broad range of evaluations from gaming-focused websites and trade magazines, recording scores in normalized scale from 0 to 100. For games published on multiple platforms, we used the highest average score among different versions (Metacritic tracks scores independently for each platform).

There are three reasons for using critical reviews rather than sales as the dependent variable. First, we are interested in audience evaluation and sales figures cannot tell us whether the consumers eventually thought highly of what they bought. Second, critics are a central audience in cultural industries and several studies have shown a connection between positive reviews and increased sales (Brewer et al., 2009; Eliashberg & Shugan, 1997; Gazley et al., 2011; Lampel & Shamsie, 2000; McKenzie, 2009). Hence critical reviews and sales data tend to not to contradict each other. Third, there is also evidence that user reviews and expert reviews for cultural goods tend to agree especially when the users are experienced in the product domain (Plucker et al., 2009). This means that professional critics and consumers tend to value similar product characteristics.

Independent variables

Developer depth of experience is measured as the number of products the developer(s) of the game has previously published in the genre(s) of the product. The variable is first calculated for each genre represented by the new product and then averaged. When the game was developed by multiple companies, we used the maximum score among

14

developers (i.e. captured the experience of the most experienced game developer). This was motivated by the assumption that the most experienced developer would likely have the greatest responsibility over the game design.

Developer scope of experience is calculated as the number of genres the developer(s) of the game in question had launched any products in during the preceding five years. When the game had multiple developers, we again used the most experienced.

Game typicality measures the percentage of words in a game synopsis that have appeared previously in products within the same game genre. The index is first calculated for each game genre the product belongs to and averaged over genres (i.e. the greatest typicality is reached when the game uses words that are present in all genres the game belongs to while the lowest typicality is reached when the game has no words that have previously appeared in any of the game's genres). When calculating the typicality of words, we only took into account words that had been used in at least some genre before in order not to create spurious novelty by invented names. Thus, the typicality score is unaffected by words that are completely new to video games; words that are considered new to a genre are borrowed from other genres.

Control variables

In order to control for effects of developer characteristics, we incorporate a number of control variables.

Developer release count captures the developer's general experience in the video game market as the number of games the developer of the game has introduced to the market

within the last seven years. When games listed multiple developers, we used the measure for the most experienced developer.

Game genre combination average rating controls for how good evaluations games representing the game's genres receive on average. The average of the critical ratings for games launched in each genre the game in question represents is calculated first. If the game represents more than one genre, the averages for each genre are again averaged. It is calculated for the past 7 years.

Category crowding controls for the differences in the release numbers in different genres, i.e. the size of the genre. It is calculated as the number of games released during the past three years in each genre assigned to the game, averaged across all genres represented by the game.

Release year is controlled through yearly dummy variables because the release activity within the game market varies across years due to the console cycle (new console launches are followed by a peak in game launches with some delay). The default year in the model is 2002, with dummies included for years 2003-2011.

During our research, we ran all models also with a control for past performance of the developer. While the variable was significant, it had no substantive effect on the findings. By including the firm-level random effects otherwise unobserved heterogeneity created by firms can be accounted for by the model (see the next section).

Analysis

Our dataset is nested as our observations are at the level of products many of which are created by the same developer. To take into account the varying capabilities of different developers, we use a multi-level model. The first level is product and the second level is the developer firm. A multi-level model allows us to draw more efficient inferences from the data by taking into account the interdependence of observations based on each company without underestimating standard error variance (Gelman & Hill, 2007, Ch. 1). We ran our regression models using Stata version 13. Since the firm-level characteristics in our data vary across products (as they are launched at different dates), our independent variables are all product-specific and we simply include a random developer-level effect in our model.

The application of multi-level model using the 'mixed' command at Stata revealed that the multi-level model had statistically significant differences with the pooled linear model, suggesting that the inclusion of random effects was warranted. We also examined our data using random-slope models, which allows us to see how the effect of game typicality varies across distinct groups (developer firms). This analysis revealed that relative typicality of games in their categories tended to have distinct effects for different firms when none of the other variables were considered (see Figure 1). The figure plots relationship for developers with at least four game releases in our data set, each represented with a distinct line. As our hypotheses do not motivate a random slopes model, we do not report these tests in greater detail.

======================

Figure 1 around here

======================

Descriptive statistics

The descriptive statistics for our sample are shown in Table 1. Our results are based on 1428 unique game titles developed by 335 distinct developers, with EA Canada the most widely represented in the dataset (49 titles). On average every developer in our data set has four products.

======================

Table 1 around here

======================

The correlation coefficients shown in Table 1 are relatively low with a few exceptions. As can be expected, the companies with a greater number of releases have higher scope (having launched games in a greater number of different genres) and greater genre experience (having launched more games in the genres represented by the new product). Because of the relatively large sample size, nearly all correlation coefficients are statistically significant.

RESULTS

Our findings are summarized in Table 2. Model 1 illustrates the effects of control variables with firm-level random effects, while Model 2 illustrates a regression model with basic effects of independent variables without interactions. As is commonly the case, we interpret all our findings using the full model (Model 3).

==================

Table 2 around here

==================

Our full model provides support for our Hypotheses 2, 3, 4 and 5. Curiously, our Hypothesis 1 is supported only in Model 2, which does not consider the interaction effects. Category-specific firm experience appears to modestly increase critical reviews for a product in the given category, but once typicality is taken into account, the effect is contingent on product characteristics. Supporting Hypothesis 2, we find that firm experience from multiple categories has a positive effect on critical appraisals. This signals that the diversity of experience makes firms better equipped to respond to audience expectations in a particular category.

We found support also for Hypothesis 3. Products conforming to their categories by content typical to them receive more positive critical appraisals. This is in line with the majority of categories research testing the effects of conformance to and deviance from a category. While in line with categories research, the finding seems to challenge the view common in the cultural industries literature that audiences expect a constant stream of novel and distinct goods to please diverse and fickle tastes (Aksoy & Robins, 1992; Mora, 2006;

19

Power & Scott, 2004). In our sample of video game critics, the audience had a strong need for familiarity (cf. Alvarez et al., 2005; Cillo et al., 2010; Peterson, 1997), perhaps stronger than the literature has so far assumed.

Our Hypotheses 4 and 5 considered interaction effects and the contingency of categorical conformity on firm capabilities. Supporting Hypothesis 4, we find that the depth of experience has greater positive effect when the product conforms to category expectations. This interaction graph, depicted in Figure 2, shows that when the product closely resembles the previous products in the category, companies enjoy a significant advantage from their depth of category-specific experience. Phrased in terms of product characteristics, we find that companies with greater depth of category-specific experience gain significant returns from categorical conformity.

Supporting Hypothesis 5, greater scope of experience from multiple categories appears to make companies significantly better in launching products that deviate from category norms. The interaction, depicted in Figure 3, suggests that while firm capabilities obtained through broad scope of past activities have insignificant advantage for products that conform to the category expectations, they become a significant advantage when the company focuses on products atypical in their categories. Indeed, our findings suggest that in contrast to the generic conclusion that categorical deviance is detrimental, deviance becomes an advantage as it allows the firm to exploit its capabilities for innovation in more innovative product offerings.

We ran a number of different tests to examine possible curvilinear relationships. While it could easily be hypothesized that a medium level of novelty might trump both extreme conformity and deviance from category norms, we did not find this to be the case with our data. However, since we only have a medium-sized sample (n=1428) and a dependent variable that is particularly hard to predict (reception of cultural products), we cannot fully rule out some mild curvilinearity of the effects. We can only conclude that the linear interaction model presented here provides the best fit with the data and supported our main hypotheses.

DISCUSSION

We build on our hypothesis development and results to theorize the roles of category conformity and firm experience in determining the audience response to cultural goods. Our findings offer firm capabilities as an overlooked dimension in the study categorical conformity and audience evaluations. We found that conformity increased the positive effect of category-specific firm experience on product reception, while deviance increased the positive effect of diverse firm experience from multiple categories. These findings can be interpreted from two perspectives, the contingency of category-related characteristics on firm's ability to launch conforming and deviant products or the contingency of firm capabilities on the product strategy that it pursues. We finish this section with a discussion of key limitations and avenues for future research offered by our theory development.

Firm capabilities and categorical conformity

Our findings vindicate the argument that there is more to audience reception and categories than the sole credibility or understandability of the goods being evaluated. Our findings strongly suggest that heterogeneity among game developers leads to subtle differences in the products that go beyond their fit with categories or their evident typicality. These subtle differences are the result of firm capabilities that have been accumulated during the development of previous products. Hence there is heterogeneity among firms in how capable they are at producing deviant goods. Firms with plenty of experience from a particular category benefit from sticking closely to established category content. Firms with diverse experience from multiple categories, on the other hand, benefit from deviance. Operating in multiple categories accumulates dynamic capabilities relating to the complex product space.

While earlier research has not focused on capabilities, some of its findings may be interpreted through the capabilities lens. Rao, Monin and Durand (2005) found that high-status chefs were able to bridge categories without penalty. In that study high status was defined through the possession of one or more Michelin stars. This particular signal of high status is also a strong measure of skill. Interpreted through the capabilities lens, Rao et al. (2005) found that highly capable professionals were able to create deviant offerings without a penalty. Similarly, in the study on typecasting by Zuckerman, Kim, Ukanwa and Rittmann (2003) it was shown that young actors are more likely to be hired if they have genre-specific experience, but veteran actors benefit from having experience from multiple genres. They reason this change to be due to veteran actors achieving recognition in the market and hence not being dependent on a clear genre-related identity in finding work. We

22

would like to point out the possibility that skill development may also play a role in this process. Multiple-genre experience takes time to accumulate and therefore veterans are better able to benefit from it than novices.

Out findings have direct implications on firm-level strategy, since both conformity versus deviance and specialization versus wide scope are managerial choices. Our findings indicate that audience response is influenced by how these strategies are combined (see Figure 4). By building category-specific capabilities through specialization firms are able to receive a reward for conforming products. On the other hand, such firms tend to receive a penalty in case they introduce deviant products. Due to their limited experience in various styles they are less likely to tap latent audience preferences. By building a broad portfolio of experiences from multiple categories firms may gain dynamic capabilities that allow them to create innovative content that deviates from category expectations. Such products tap latent demand for novelty. Whichever capabilities the firm builds, their value is contingent on a matching product strategy.

In the long term, specialization appears to be the riskier strategy once the role of genre effects are taken into account. The control variable Game genre combination average rating shows that products launched in genres that tend to receive higher ratings on average are more successful. The variable Category crowding in turn shows that the more products have been launched in the product's categories, the worse off the new product tends to be. These effects can create a dynamic where companies can benefit from proactively moving away from crowded product niches and into "fashionable" attractive niches currently valued by the critics. Companies that specialize cannot readily appropriate their capabilities outside their established niche, while firms with a wider scope can receive a reward for

both conforming and deviating products and expand their innovative abilities by maneuvering towards more fashionable and less crowded product niches. In this way, firms with a wider scope would seem to be better positioned to hedge their bets in the volatile market place for cultural goods.

==================

Figure 4 around here

==================

Limitations and future research

Each empirical study is limited by the chosen setting. To develop theory on the relationship of categories and firm capabilities, we examined the video games industry. We found that the appraisal of products in this setting is dauntingly difficult to predict or explain, as each product is unique in ways that are difficult to codify. Thus, attempts to quantitatively model the critical reception of goods lead to a disappointingly low predictive power. While this is likely to be the case in creative industries more generally, similar analyses on other industries would strengthen our conclusions. Moreover, games may have some specific characteristics that may limit our ability to extend implications to creative industries more generally. For example, while the critics in our case are professionals employed by the media, they lack a professional or educational background than book and film critics more commonly share.

To further add nuance to our analysis, future studies could examine the differences between categories as training grounds for firms. Are there differences between categories in how much experience in them contributes to a firm's learning? Categories may entail varying

levels of content diversity measured as the number of distinct words (cf. Hsu et al., 2012b).

This may have an effect on learning. Moreover, categories with lower and higher

boundaries (cf. Pontikes, 2012) may offer different experience effects.

Moreover, cultural economists have found that there are different types of audiences: some

expect familiarity and immediate enjoyment (Alderighi & Lorenzini, 2012; Lévy-Garboua

& Montmarquette, 1996) whereas others crave for novel experiences and want to

accumulate their cultural capital in order to be better able to enjoy diverse future products

(Becker & Murphy, 1988). Future research could try to tie different audience groups to firm

strategy in cultural industries.

CONCLUSIONS

Our study joins a growing literature on audience response to category conformity and

deviance. We found that broad scope of firm experience from multiple categories improves

critical reviews when the company launches products that deviate from category

expectations, while category-specific depth of experience has a positive effect only when

the firm conforms to category expectations. These findings highlight the contingency of

categorical conformity on firm capabilities and the contingency of firm capabilities on

product strategy.

Our findings imply two archetypal and opposite strategies for cultural industries: a firm can

either focus on novelty and create innovative capacity through broad scope of product

portfolio, or the firm may focus on incremental innovation and create category-specific

capabilities by narrow scope and greater depth of its product portfolio. We contribute to

categories literature by introducing two types of category-related firm experience, and

showing that they have an effect on a firm's ability to target audience expectations and hence on their response to conformity/deviance.

REFERENCES

Aksoy A, Robins K. 1992. Hollywood for the 21st-century - Global competition for critical mass in image markets. Cambridge Journal of Economics **16**(1): 1-22.

Alderighi M, Lorenzini E. 2012. Cultural goods, cultivation of taste, satisfaction and increasing marginal utility during vacations. Journal of Cultural Economics **36**(1): 1-26.

Alexander PJ. 1996. Entropy and popular culture: Product diversity in the popular music recording industry. American Sociological Review **61**(1): 171-174.

Alvarez JL, Mazza C, Pedersen JS, Svejenova S. 2005. Shielding Idiosyncrasy from Isomorphic Pressures: Towards Optimal Distinctiveness in European Filmmaking. Organization **12**(6): 863-888.

Anand BN, Khanna T. 2000. Do firms learn to create value? The case of alliances. Strategic Management Journal **21**(3): 295-315.

Argyris C, Schön DA. 1978. Organizational learning: A theory of action perspective. Addison-Wesley Reading, MA.

Becker GS, Murphy KM. 1988. A theory of rational addiction. The journal of political economy: 675-700.

Bowers A. 2014. Relative Comparison and Category Membership: The Case of Equity Analysts. Organization Science **0**(0): null.

Bowman EH, Hurry D. 1993. STRATEGY THROUGH THE OPTION LENS - AN INTEGRATED VIEW OF RESOURCE INVESTMENTS AND THE INCREMENTAL-CHOICE PROCESS. Academy of Management Review **18**(4): 760-782.

Brewer SM, Kelley JM, Jozefowicz JJ. 2009. A blueprint for success in the US film industry. Applied Economics **41**(5): 589-606.

Brouthers KD, Brouthers LE, Werner S. 2008. Resource-based advantages in an international context. Journal of Management **34**(2): 189-217.

Cappetta R, Cillo P, Ponti A. 2006. Convergent designs in fine fashion: An evolutionary model for stylistic innovation. Research Policy **35**(9): 1273-1290.

Cillo P, De Luca LM, Troilo G. 2010. Market information approaches, product innovativeness, and firm performance: An empirical study in the fashion industry. Research Policy **39**(9): 1242-1252.

Cohen MD, March JG, Olsen JP. 1972. A Garbage Can Model of Organizational Choice. Administrative Science Quarterly **17**(1): 1-25.

Durand R, Rao H, Monin P. 2007. Code and conduct in French cuisine: Impact of code changes on external evaluations. Strategic Management Journal **28**(5): 455-472.

Eisenhardt KM, Martin JA. Year. Dynamic capabilities: What are they?

Eliashberg J, Shugan SM. 1997. Film critics: Influencers or predictors? Journal of Marketing **61**(2): 68-78.

Gazley A, Clark G, Sinha A. 2011. Understanding preferences for motion pictures. Journal of Business Research **64**(8): 854-861.

Gelman A, Hill J. 2007. Data analysis using regression and multilevel/hierarchical models. Cambridge University Press.

Glynn MA, Abzug R. 2002. Institutionalizing identity: Symbolic isomorphism and organizational names. Academy of Management Journal **45**(1): 267-280.

Granstrand O. 1998. Towards a theory of the technology-based firm. Research policy **27**(5): 465-489.

Hargadon A, Sutton RI. 1997. Technology brokering and innovation in a product development firm. Administrative science quarterly: 716-749.

Helfat C, Finkelstein S, Mitchell W, Peteraf MA, Singh H, Teece DJ, Winter SG. 2007. Dynamic Capabilities: Understanding Strategic Change in Organizations. Blackwell: Oxford, U.K.

Hesmondhalgh D. 2002. The Cultural Industries. SAGE Publications.

Hsu G. 2006. Jacks of all trades and masters of none: Audiences' reactions to spanning genres in feature film production. Administrative Science Quarterly **51**(3): 420-450.

Hsu G, Kocak O, Hannan MT. 2009. Multiple Category Memberships in Markets: An Integrative Theory and Two Empirical Tests. American Sociological Review **74**(1): 150-169.

Hsu G, Negro G, Perretti F. 2012a. Hybrids in Hollywood: a study of the production and performance of genre-spanning films. Industrial and Corporate Change **21**(6): 1427-1450.

Hsu G, Roberts PW, Swaminathan A. 2012b. Evaluative Schemas and the Mediating Role of Critics. Organization Science **23**(1): 83-97.

Huber GP. 1991. Organizational Learning: The Contributing Processes and the Literatures. Organization Science **2**(1): 88-115.

Jonsson S, Greve HR, Fujiwara-Greve T. 2009. Undeserved Loss: The Spread of Legitimacy Loss to Innocent Organizations in Response to Reported Corporate Deviance. Administrative Science Quarterly **54**(2): 195-228.

King AA, Tucci CL. 2002. Incumbent entry into new market niches: The role of experience and managerial choice in the creation of dynamic capabilities. Management Science **48**(2): 171-186.

Lampel J, Shamsie J. 2000. Critical push: Strategies for creating momentum in the motion picture industry. Journal of Management **26**(2): 233-257.

Levitt B, March JG. 1988. Organizational Learning. Annual Review of Sociology **14**(1): 319-338.

Lévy-Garboua L, Montmarquette C. 1996. A microeconometric study of theatre demand. Journal of Cultural Economics **20**(1): 25-50.

Mauws MK. 2000. But Is It Art? The Journal of Applied Behavioral Science **36**(2): 229-244.

McKenzie J. 2009. Revealed word-of-mouth demand and adaptive supply: survival of motion pictures at the Australian box office. Journal of Cultural Economics **33**(4): 279-299.

Mora E. 2006. Collective production of creativity in the Italian fashion system. Poetics **34**(6): 334-353.

Negro G, Leung MD. 2013. "Actual" and Perceptual Effects of Category Spanning. Organization Science **24**(3): 684-696.

Nerkar A, Roberts PW. 2004. Technological and product-market experience and the success of new product introductions in the pharmaceutical industry. Strategic Management Journal **25**(8-9): 779-799.

Peltoniemi M. 2014. Cultural Industries: Product–Market Characteristics, Management Challenges and Industry Dynamics. International Journal of Management Reviews: n/a-n/a.

Peterson K. 1997. The distribution and dynamics of uncertainty in art galleries: A case study of new dealerships in the Parisian art market, 1985-1990. Poetics **25**(4): 241-263.

Plucker JA, Kaufman JC, Temple JS, Qian MH. 2009. Do Experts and Novices Evaluate Movies the Same Way? Psychology & Marketing **26**(5): 470-478.

Pontikes EG. 2012. Two Sides of the Same Coin: How Ambiguous Classification Affects Multiple Audiences' Evaluations. Administrative Science Quarterly **57**(1): 81-118.

Porter MF. 1980. An algorithm for suffix stripping. Program **14**(3): 130−137.

Power D, Scott AJ. 2004. Cultural Industries and the Production of Culture. Routledge.
31

Rao H, Monin P, Durand R. 2005. Border Crossing: Bricolage and the Erosion of Categorical Boundaries in French Gastronomy. American Sociological Review **70**(6): 968-991.

Rothaermel FT, Deeds DL. 2006. Alliance type, alliance experience and alliance management capability in high-technology ventures. Journal of Business Venturing **21**(4): 429-460.

Ruef M, Patterson K. 2009. Credit and Classification: The Impact of Industry Boundaries in Nineteenth-century America. Administrative Science Quarterly **54**(3): 486-520.

Salton G. 1971. The SMART Retrieval System - Experiments in Automatic Document Processing. Prentice-Hall, Inc.

Sampson RC. 2005. Experience effects and collaborative returns in R&D alliances. Strategic Management Journal **26**(11): 1009-1031.

Schildt H, Keil T, Maula M. 2012. The temporal effects of relative and firm-level absorptive capacity on interorganizational learning. Strategic Management Journal **33**(10): 1154-1173.

Schilling MA, Vidal P, Ployhart RE, Marangoni A. 2003. Learning by doing something else: Variation, relatedness, and the learning curve. Management Science **49**(1): 39-56.

Shamsie J, Martin X, Miller D. 2009. In with the old, in with the new: Capabilities, strategies, and performance among the Hollywood studios. Strategic Management Journal **30**(13): 1440-1452.

Smith EB. 2011. Identities as Lenses: How Organizational Identity Affects Audiences' Evaluation of Organizational Performance. Administrative Science Quarterly **56**(1): 61-94.

Suzuki J, Kodama F. 2004. Technological diversity of persistent innovators in Japan: Two case studies of large Japanese firms. Research Policy **33**(3): 531-549.

Thompson P, Jones M, Warhurst C. 2007. From conception to consumption: creativity and the missing managerial link. Journal of Organizational Behavior **28**(5): 625-640.

Uhlenbruck K. 2004. Developing acquired foreign subsidiaries: The experience of MNEs in transition economies. Journal of International Business Studies **35**(2): 109-123.

Vergne JP. 2012. Stigmatized categories and public disapproval of organizations: A mixed-methods study of the global arms industry, 1996-2007. Academy of Management Journal **55**(5): 1027-1052.

Zuckerman EW. 1999. The categorical imperative: Securities analysts and the illegitimacy discount. American Journal of Sociology **104**(5): 1398-1438.

Zuckerman EW. 2004. Structural incoherence and stock market activity. American Sociological Review **69**(3): 405-432.

Zuckerman EW, Kim TY. 2003. The critical trade-off: identity assignment and,box-office success in the feature film industry. Industrial and Corporate Change **12**(1): 27-67.

Zuckerman EW, Kim TY, Ukanwa K, Rittmann James v. 2003. Robust Identities or Nonentities? Typecasting in the Feature-Film Labor Market. American Journal of Sociology **108**(5): 1018-1073.

TABLE 1

DESCRIPTIVE STATISTICS

		Mean	Std. Dev.	Min	Max	1	2	3	4	5	6
1	Average critic score	70.93	13.66	0	98						
2	Developer depth of experience	6.59	10.02	0	72	0.12					
3	Developer scope of experience	8.10	7.44	0	44	0.04	0.67				
4	Game typicality	2.97	1.80	0.13	11.91	0.02	0.12	0.05			
5	Developer release count	13.86	18.14	1	103	0.09	0.77	0.73	0.05		
6	Game genre combination average rating	75.70	6.05	49	95	0.15	0.04	-0.03	-0.08	0.01	
7	Category crowding	578.31	357.10	6	2294	-0.02	0.20	0.11	0.29	-0.04	0.08

All correlation coefficients above .06 and below -.06 are significant at p<.05. n=1428.

TABLE 2

MULTI-LEVEL LINEAR REGRESSION MODEL EXPLAINING THE AVERAGE CRITIC SCORE RECEIVED BY A NEW PRODUCT

	Dependent variable: Average critic score			
	Model 1: Control variables	Model 2: Basic effects	Model 3: Interactions	Hypoth.
Fixed effects model (game)				
Developer depth of experience		0.19 **	0.01	**H1**
		(.06)	(.10)	
Developer scope of experience		0.07	0.29 *	**H2**
		(.09)	(.14)	
Game typicality		0.51 **	0.72 **	**H3**
		(.21)	(.30)	
Depth of experience * Game typicality			0.07 **	**H4**
			(.03)	
Scope of experience * Game typicality			-0.08 *	**H5**
			(.04)	
Developer release count	0.01	-0.10 +	-0.10 *	
	(.03)	(.05)	(.05)	
Game genre combination average rating	0.27 ***	0.29 ***	0.29 ***	
	(.06)	(.06)	(.06)	
Category crowding (x100)	-0.13	-0.39 **	-0.35 *	
	(.13)	(.14)	(.14)	
Year 2004	-0.52	-0.52	-0.61	
	(1.40)	(1.40)	(1.40)	
Year 2005	-1.53	-1.41	-1.57	
	(1.47)	(1.47)	(1.46)	
Year 2006	-3.98 *	-3.82 *	-4.10 **	
	(1.52)	(1.52)	(1.52)	
Year 2007	-3.96 *	-3.96 **	-4.19 **	
	(1.52)	(1.52)	(1.52)	
Year 2008	-3.06 *	-3.21 *	-3.66 *	
	(1.58)	(1.59)	(1.60)	
Year 2009	-1.17	-1.18	-1.62	
	(1.69)	(1.69)	(1.70)	
Year 2010	-0.66	0.38	-0.02	
	(1.85)	(1.86)	(1.86)	
Year 2011	1.59	2.82	3.01	
	(2.64)	(2.67)	(2.68)	
Constant	51.69 ***	49.21 ***	48.81	
	(4.73)	(4.84)	(4.88)	
Random effects model (developer)				
Variance of Constant	59.97	59.64	59.45	
	(8.81)	(8.77)	(8.77)	
Variance of Residual	128.61	126.83	126.22	
	(5.54)	(5.47)	(5.45)	

This is a mixed-effects regression model with random effects for developers, n=1428, groups=355 (group size 1-38, average size 4.0). Standard errors are in parentheses below unstandardized coefficients.
Significance levels are indicated in the typical manner: *** p<.001, ** p<.01, * p<.05, + p<.10.

FIGURE 1

FIRM-SPECIFIC RELATIONSHIPS OF GAME TYPICALITY AND AVERAGE
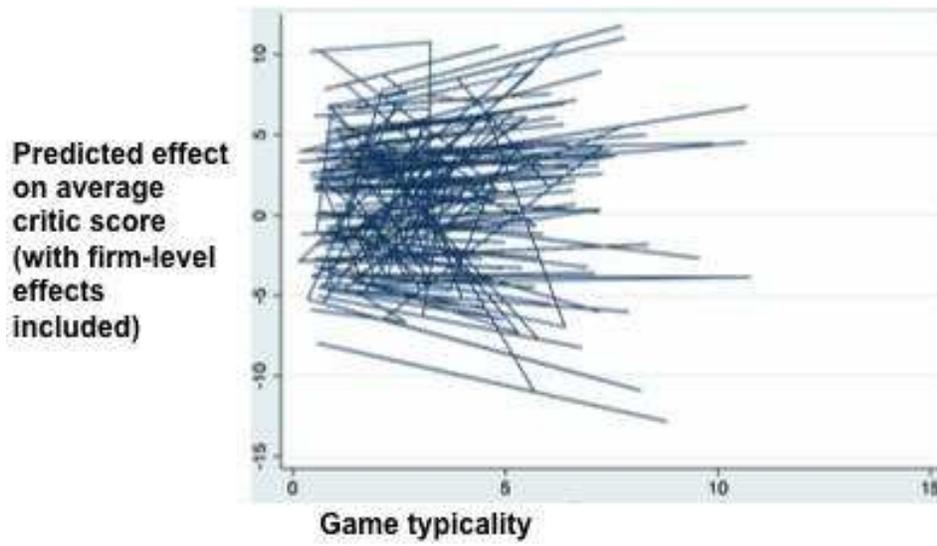
CRITIC SCORE

FIGURE 2

**INTERACTION OF DEPTH OF EXPERIENCE AND GAME TYPICALITY (X-AXIS)
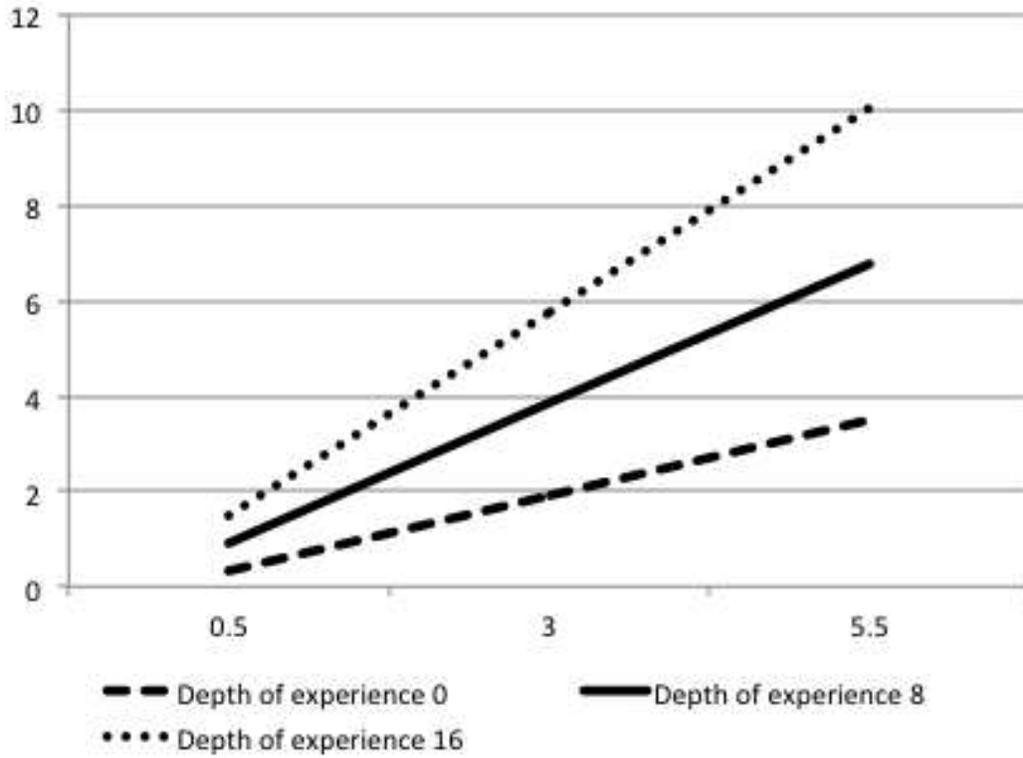ON AVERAGE CRITIC SCORE (Y-AXIS)**

FIGURE 3

**INTERACTION OF SCOPE OF EXPERIENCE AND GAME TYPICALITY (X-AXIS)
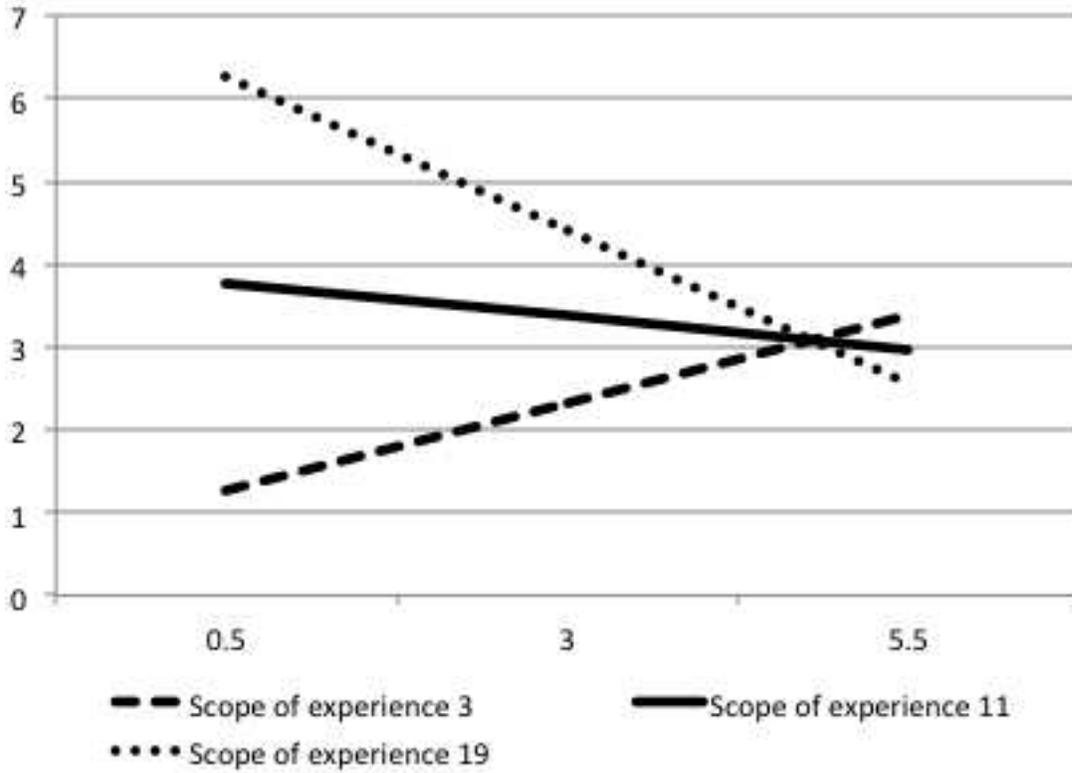ON CRITICAL APPRAISAL (Y-AXIS)**

FIGURE 4

PRODUCT RECEPTION AS A FUNCTION OF FIRM CAPABILITIES AND PRODUCT

STRATEGY

Capability development strategy

		Specialization	Wide scope
Product strategy	Conformity	Reward for capable familiarity	Reward for familiarity
	Deviance	Penalty for incapable innovation	Reward for capable innovation