Review of: "An integrative model of new product evaluation: A systematic investigation of perceived novelty and product evaluation in the movie industry"

James Cutting[1]

1 Cornell University

Potential competing interests: The author(s) declared that no potential competing interests exist.

Luan and Kim (2022) present a compelling analysis of the curvilinear relationship between objective and subjective measures of a movie's novelty. Their proxy for an objective measure was coarse, but straightforward and effective. It was determined by the genre combination of a given movie compared against the frequency of that combination in the past. That is, the more frequently an "action, adventure, comedy" movie appeared in the past, for example, the less novel a new "action, adventure, comedy" movie would be scored. Their proxy for a subjective measure was a modified content analysis of reviews of each movie, scored for the proportion of words in the review that seemed to reflect novelty—"fresh," "original," "different," etc.

Strikingly, Luan and Kim found that a moderate amount of objective novelty garnered the highest measure of subjective novelty. That is, not only would a new "comedy, romance" (a romcom) likely not be regarded as novel, but a highly unusual genre combination, like "family, horror, romance" would not be regarded as novel either. This seems odd, but of course such a movie would likely never be made and, if it were, it would not likely do well at the box office.

Moreover, Luan and Kim found a subtle interaction in their data as well. Movies by more-celebrated directors had peak subjective novelty scores that were objectively less novel than those of less-celebrated directors. This means that celebrated filmmakers could make movies that were objectively more novel but weren't perceived as being more novel. Thus, Steven Spielberg and Christopher Nolan could make almost any kind of movie and it would be regarded as less novel than if it were an indie made by an unheralded filmmaker.

One might have qualms with Luan and Kim's assessment of objective novelty. Spielberg, for example, made*The Color Purple* (1985), a "drama" (the most common genre category of all time, see Cutting, 2022) that was regarded as a "turning point in his career, as it was a departure from the summer blockbusters for which he had become known" (Wikipedia). It would score low on Luan and Kim's objective novelty. But it was highly novel on at least two other grounds: the genre flip from Spielberg's usual forte, and its dealing with rape and incest in the US Black south. Nonetheless, anecdotes about single movies do not carry the weight of data, and I believe Luan and Kim's data. However, let me recast what they have found in a somewhat different terms.

A theoretical purpose of Luan and Kim's article is to integrate two disparate threads of prior results. One thread is that increased novelty of a product increases positive evaluation. They quote Spielberg (p. 1): "The public has an appetite for … anything that is as far away from reality as possible." The other thread is the reverse. They quote a renown filmmaker of animations, Brad Bird: "Familiarity is all the rage. And if you are doing something that doesn't have its rhythms preset … everybody's a bit uncomfortable." Luan and Kim then back up these quotes with literatures that support each view.

Such a contrast in literatures, and even a synthesis of them, have been found elsewhere. Put these trends side by side and, together, they make an inverted U-shaped function. This is an idea towards which I am quite partial (Cutting 2019, 2020). A bit of novelty is often thought to be a good thing, but too much novelty is not.

The dependent measure in the Luan and Kim study is the relative occurrence in movie reviews of terms like "unusual," "new," "unique," and so forth. These words usually have positive value. The use of such words generally means that the reviewer has liked the movie. The ordinate of Luan and Kim's graphs is "product *evaluation*" (my emphasis). Not only might these ordinates be taken as a measure of *evaluated* novelty, but they can also be taken as a measure of affective response. Historically, this goes under the umbrella term *hedonic value*, where positive values indicate things that you like and negative values indicate things that you don't. The explanation for all this necessitates some history (see also Cutting, 2020). What follows is a supportive backdrop of Luan and Kim's results from psychology, a discipline much older than, but nonetheless related to, marketing.

The general idea can be traced back at least to Eleanor Mure's (1831) story of the three bears, which the English poet Robert Southey popularized. Only later was the character Goldilocks inserted (Steele, 1918), replacing an old woman. The end result was: one bowl of porridge is the right temperature (not too hot, not too cold), one chair is the right height (not too high, not too low), and one bed feels just right (not too hard, not too soft). Later in the 19th century this idea of an optimal middle ground found its way into psychology through the German psychologist Wilhelm Wundt (1832-1920). It was next filtered through the ideas of the British-Canadian psychologist and aesthetician Daniel Berlyne (1924-1976). It then became one of the most enduring ideas in psychology. Moreover, it has since spread to information science (McCormack and d'Inverno 2012), economics (Kaimann, Stroh-Maraun, & Cox 2018), consumer research (Anand & Holbrook 1986) and now, I would claim, to movies (Luan & Kim 2022; see also Berliner 2017).

The phenomenon also has several variants. One is the *discrepancy hypothesis* (Blijlevens et al 2012; Haber 1958; McCall & McGhee 1977), which suggests that people have a preference for stimuli that differ slightly, but not too much or too little, from a norm (see also Mather, 2013). Another is the "law" of Yerkes and Dodson (1908), where people are said to perform better at a moderate level of arousal rather than lower or higher. And a third is an *opponent-process theory* (Solomon & Corbit 1974; Solomon 1980) for which reward and aversion compete.

For Luan and Kim (2022) the relationship between reviewer evaluation and movie novelty is captured by what is typically called the Wundt curve, or the inverted-U. This curve plots hedonic value—how much you like something—against

another variable. Across the post-Berlyne literatures on aesthetics, arousal, and preference, the axes of this type of plot are given in many different ways. A version of it, from Berlyne (1970), is shown in the middle panel of Figure 1, below. The left panel shows Wundt's figure, which has a much narrower band of variation for which positive hedonic value is high. The German abscissa label translates as "intensity of sensation" and the ordinate labels roughly as "pleasure" and "displeasure." For Berlyne (1960, 1971) arousal potential is on the horizontal axis and hedonic value, in his case how much one likes a movie, on the vertical axis. Notice that, as part of the abscissa of the central panel, novelty and complexity play a role.

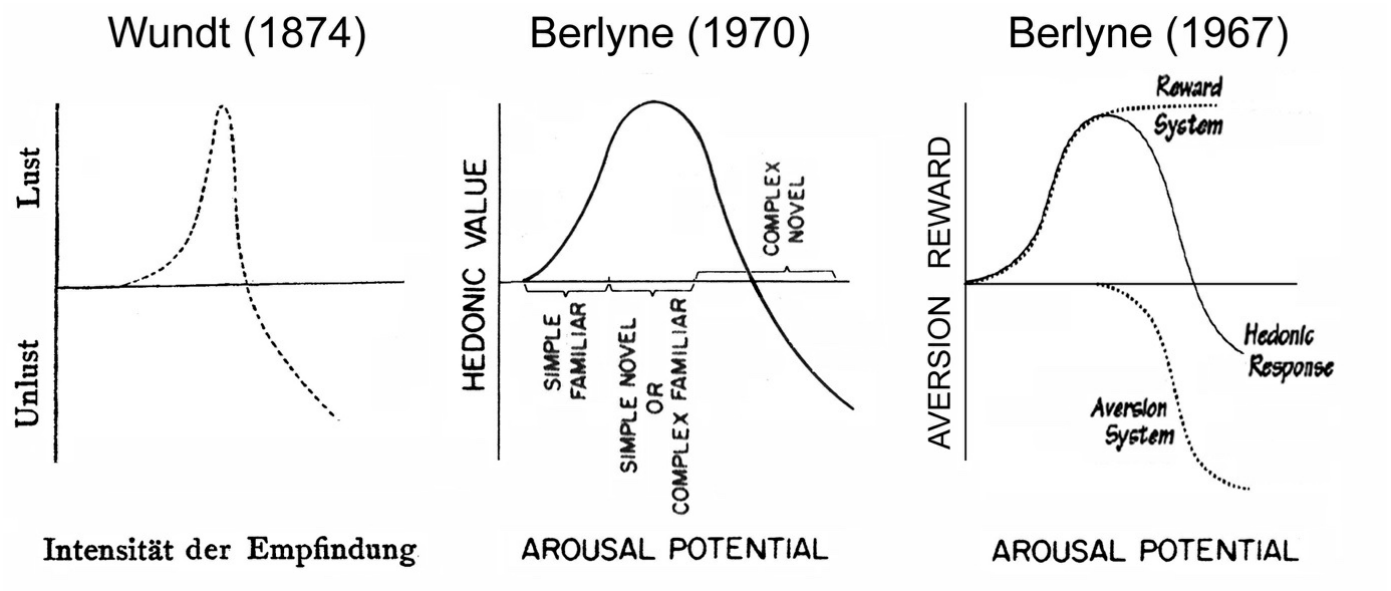

Figure 1: Three representations of the Wundt curve. The left panel is simplified from Wundt (1874, p. 432, Figure 97), the central panel modified from Berlyne (1970, Figure 4), and the right panel reworked from Berlyne (1967, p. 88, Figure 2). The whole figure is from Cutting (2020).

The notion of arousal potential, rather than simply arousal, is important. It acknowledges that some stimuli (e.g. movies) can affect people in different ways and at different times. According to Berlyne (1971, pp. 68-69), arousal is dependent on three properties. The first are *psychophysical*. As an example, consider temperature. Generally, like Goldilocks, we don't like soup that is too cold or too hot. In between is best. The second are *ecological*, those properties around us every day that are personally meaningful, and include even how we feel at a given time. You might not want to watch *Sophie's Choice* (1982) at your birthday party. Cultures, and the individuals within them, vary widely in their embrace of such properties at particular times.

Third, and most important, are *collative* properties. Collative means to collect and compare, gathering information from diverse sources, and then being able to rank order things based on all that information. Collative properties are evaluative. Examples include novelty, complexity, conflict, uncertainty, suprisingness, and ambiguity (Berlyne 1960, p.44; 1971, p. 69). Chmiel and Schubert (2017, p. 887) noted that "While Berlyne proposed that all three types of variables contribute towards aesthetic preference, his legacy is the discovery of collative variables and the idea that these are the 'most significant' determinants of preference."

Berlyne's intent was that collative properties—particularly novelty and complexity— could be "usefully discussed in the language of information theory" (Berlyne 1971, pp. 69-70). These could be pitted against hedonic value (observer preferences) to yield the Wundt curve. He started simplistically, measuring familiarity as the number of times that listeners heard a novel melody. And he measured complexity in terms of the number of sides of a polygon (Berlyne 1970). Unfortunately, his own data didn't really reveal the pattern of the Wundt curve.

Remarkably, one would be hard pressed to find an idea more persuasive in psychology that for forty years flunked corroboration so thoroughly. For example, across more than a dozen experiments, Walker (1981) found no evidence for a Wundt curve. In addition, Marin et al (2016, p. 2) cited a dozen other papers, often with multiple experiments, that also found no supportive results. Instead of an inverted U-shaped function, hedonic judgments typically increased with complexity or decreased with familiarity—an upward and downward set of results not unlike those that Luan and Kim (2022) report for novelty.

Why such empirical nonsupport? Consider three possibilities. First, the focus on arousal seems clearly misplaced. We now know that equal values of arousal (for example, the autonomic measures of heart rate, blood pressure, and skin conductance) can be associated with both positive and negative responses. Thus, arousal confuses the issue. Second, hedonic tone—the ability to feel pleasure—is itself multidimensional and dependent on both the stimulus set and the individual (Cupchik 1986; Martin et al 2016). Third, a single underlying process is insufficient. As Berlyne (1971, pp. 86-95) explained, to generate a U-shaped function one generally needs two processes and a staggered relation between them. One process generates the upward part of the function and a second generates the downward part. Moreover, these become relevant in different collative ranges, one in the lower range and one in the upper. The two processes are a reward system and an aversion system. The neurophysiology of these systems is well documented in contemporary theory and data (Hu 2016).

Consider the graph in the right panel of Figure 1, modified from Berlyne (1967). The reward system builds in its response to a collative property, like novelty or complexity. It forms an ogival curve, asymptotic at both ends, one at indifference on the left and the other at some maximum value on the right. With increases in the collative variable the aversion system waits until some threshold is exceeded and then begins to be activated. Its response then follows a function that is also an ogive and doubly asymptotic, but with a reversed polarity and a bit greater in amplitude. Add the two together and you have a Wundt curve. This underlying, two-process account can be found in other, related domains (Solomon & Corbit 1974).

Gratifyingly, and despite the checkered history, recent research has found results endorsing the Wundt curve. Making good on one of Berlyne's wishes, Kidd et al (2012) found that, using carefully calibrated information-theoretic criteria, infants preferred visual sequences of moderate complexity over those that were both simpler and more complex. And Gravino et al (2019) analyzed big data from recommender systems, those online algorithms that recommend books,

songs, and other items based on what the internet surfer has purchased. Among popular songs, they found that an individual listener was more likely to listen to a song she's heard, say, ten times before than one she has listened to fifty times or only twice. Such results suggests that both familiarity and novelty drive preferences, as suggested in the central panel of Figure 1, with a Goldilocks zone in between,

Moreover, as Berlyne originally suggested, individual differences can matter. Güçlütürk et al (2016) generated complex visual patterns for aesthetic judgments by viewers. Importantly, they first normalized results within individuals and then made comparisons across them. This procedure yielded an inverted U-shaped pattern for individual preferences, which was absent when no individual differences were considered. Similarly, Sousa et al (2019) modeled preferences to popular music along three collative variables—novelty, complexity, and uncertainty—and, once they were combined, found separate and different Wundt curves for individual listeners. These sets of results may explain why many earlier findings were inconsistent with Berlyne's theory.

Finally, in the domain of movies and somewhat similar to the work of Luan and Kim (2022), Sreenivasan (2013) looked at the key-words viewers had used to tag all movies on the Internet Movie Database (IMDb.com) between 1930 and 2010. He then measured the general uniqueness of the terms used by viewers for each particular movie compared to those terms used for all movies released in prior years (dating from 1910). In general, the movies that were most successful were ones that had an intermediate number of unique terms. Movies with fewer unique terms and those with more unique terms did not do as well at the box office. This supports the idea that Hollywood movies continually search out new but related themes as they continue mostly to match genre expectations of viewers. It may also help explain why remakes, sequels, prequels, and spinoffs are so popular—and always have been so (Cutting, 2021, p.2).

In summary, Luan and Kim's (2022) appraisal of perceived novelty in popular cinema can be made to fit comfortably within existing psychological theory. Its backbone is Daniel Berlyne's theory of hedonic value. Here, I have reviewed, reworked, and updated that theory—one that had a pervasive but rocky existence in the latter third of the 20th century—and find consistent and relevant empirical support for it in the data of Luan and Kim (2022). I am happy to see that the theory continues to extend into the domain of popular cinema.

It really is the case that we share a lot with Goldilocks as we forage for pleasure in the world around us.

References

Anand, P. & Holbrook, M. P. (1986). Chasing the Wundt curve: An adventure in consumer esthetics. *Advances in Consumer Research* (R. J. Lutz, Ed.), *13*, 655-657. Provo, UT: Association for Consumer Research.

Berliner, T. (2017). *Hollywood aesthetic: Pleasure in American cinema.* New York: Oxford University Press.

Berlyne, D. E. (1960). *Conflict, arousal and curiosity* . New York: McGraw-Hill.

Berlyne, D. E. (1967). Arousal and reinforcement." *Nebraska Symposium on Motivation, 15,* 1-110. doi: 10.1007/978-1-4613-3195-7_2

Berlyne, D. E. (1970). Novelty, complexity, and hedonic value. *Attention, Perception and Psychophysics, 8* (5), 279-278. doi: 10.3758/BF03212593

Berlyne, D.E. (1971). *Aesthetics and psychobiology*. New York: Appleton-Century-Crofts.

Blijlevens, J., Carbon, C.-C., Mugge, R. & Schoormans, J. P. L. (2012). Aesthetic appraisal of product designs: Independent effects of typicality and arousal. *British Journal of Psychology, 103* (1), 44-57. doi: 10.1111/j.2044-8295.2011.02038.x

Chmiel, A., & Schubert, E. (2017). Back to the inverted-U for music preference: A review of the literature. *Psychology of Music, 45*(6), 886-909. doi: 10.1177/0305735617697507

Cupchik, G. C. (1986). A decade after Berlyne: New directions in experimental aesthetics. *Poetics, 15*(4-6), 345-369. doi: 10.1016/0304-422X(86)90003-3

Cutting, J. E. (2019). Simplicity, complexity, and narration in popular movies. In M. Grishakova and M. Poulaki (Eds.). *Narrative complexity: Cognition, embodiment, evolution* (pp. 200-222). Lincoln, NE: University of Nebraska Press.

Cutting, J. E. (2020). Goldilocks aesthetics. *Projections, 14*(2), 66-74. doi: 10.3167/proj.2020.140206

Cutting, J. E. (2021). *Movies on our minds: The evolution of cinematic engagement.* New York: Oxford University Press.

Cutting, J. E. (2022). Evolution of the depiction of telephone calls in popular cinema. *Projections, 16*(2), in press.

Gravino, P., Monechi, B. & Loreto, V. (2019). Towards novelty-driven recommender systems. *Comptes Rendu Physique, 20*(4), 371-379. doi: 10.1016/j.crhy.2019.05.014

Güçlütürk, Y., Jacobs, R. H. A. H., & van Lier, R. (March 2016). Liking versus complexity: Decomposing the inverted-U curve. *Frontiers of Human Neuroscience, 10*, Article 112, 1-11. doi: 10.3389/fnhum.2016.00112

Haber, R. N. (1958). Discrepancy from adaptation level as a source of affect. *Journal of Experimental Psychology, 56* (4), 370-375. doi: 10.1037/h0041761

Hu, H. (2016). Reward and aversion. *Annual Review of Neuroscience, 39*, 297-324. doi: 10.1146/annurev-neuro-070815-014106

Kaimann, D., Stroh-Maraun, N., & Cox., J (2018). Variety in the video game industry: An empirical study of the Wundt curve. *Management Decision Economics, 39* (3), 354-362. doi: 10.1002/mde.2909

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The Goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple or too complex. *PloS One, 7* (5), Article e36399, 1-8. doi: 10.1371/journal.pone.0036399

Luan, Y. & Kim, Y. J. (11 March 2022). An integrative model of new product evaluation: A systematic investigation of perceive novelty and product evaluation in the movie industry. *PLoS One.* doi: 10.1371/journal.pone.0265193

Marin, M. M., Lampatz, A., Wandl, M., & Leder, H. (November 2016). Berlyne revisited: Evidence for the multifaceted nature of hedonic appreciation of paintings and music. *Frontiers in Human Neuroscience, 10*, Article 536, 1-20. doi: 10.3389/fnhum.2016.00536

Mather, E. (August 2013). Novelty, attention, and challenges for developmental psychology. *Frontiers of Psychology, 4,* Article 491, 1-4. doi: 10.3389/fpsyg.2013.00491

McCall, R. B., & McGhee, P. E. (1977). The discrepancy hypothesis of attention and affect in infants. In I. Č. Užgiris & F. Weizmann (Eds.) *The structuring of experience* (pp. 179-210). New York: Plenum. doi: 10.1007/978-1-4615-8786-6_7

McCormack, J. & d'Inverno, M. (Eds.) (2012). *Computers and creativity.* New York: Springer.

Mure, E. (1831). *The story of the three bears.* https://digitalarchive.tpl.ca/objects/371300/the-story-of-the-three-bears--metrically-related-with-illus

Solomon, R. L. 1980. The opponent-process theory of acquired motivation: The costs of pleasure and benefits of pain. *American Psychologist, 35*(8), 691-712. doi: 10.1037/0003-066X.35.8.691

Solomon, R. L. & Corbit, J. D. (1974). An opponent-process theory of motivation: I. Temporal dynamics of affect. *Psychological Review, 81*(2), 119-

145. doi: 10.1037/h0036128

Sousa, A. M., Jussara, K., Almeida, M., & Figueiredo, F. (August 2019). Analyzing and modeling curiosity in online content consumption: A LastFM case study. *IEEE/ACM International Conference on Advances in Social Networks, Analysis, and Mining.* 426-431. doi: 10.1145/3341161.3342917

Sreenivasan, S. (2013). Quantitative analysis of the evolution of novelty in cinema through crowdsourced keywords. *Scientific Reports, 3*, Article 2758, 1-10. doi: 10.1038/srep02758

Steele, F. A. (1918). *English Fairy Tales.* London: Macmillan & Co.

Walker, E. L. (1981). The quest for the inverted U. In H. I. Day (Ed.) *Advances in Intrinsic Motivation and Aesthetics* (pp. 39–70). New York: Plenum.

Wundt, W. (1874). *Grundzüge der physiologischen psychologie,* 1st ed., Leipzig, Germany: Verlag von Wilhelm Engelmann.

Yerkes, R. & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit formation. *Journal of Comparative Neurology and Psychology, 18*(5), 459-482. doi: 10.1002/cne.920180503