Technological Forecasting & Social Change 146 (2019) 167–180

Contents lists available at ScienceDirect

Technological Forecasting & Social Change

journal homepage: www.elsevier.com/locate/techfore

Digitally forecasting new music product success via active crowdsourcing

Dennis M. Steininger[a,*], Simon Gatzemeier[b]

[a] Faculty of Business and Economics, University of Augsburg, Universitaetsstrasse 16, 86159 Augsburg, Germany
[b] Business School, University of Mannheim, Germany

ARTICLE INFO

Keywords:
Crowdsourcing
Newcomer
Forecasting
Music chart success
Music industry

ABSTRACT

Deciding which artist or song to sign and promote has always been a challenge for recording companies, especially when it comes to innovative newcomer singers without any chart history. However, the specifics of a creative industry such as the hedonic nature of music, socio-network effects, and ever fastening fashion cycles in combination with digitalization have made the recording industry even more competitive and these initial decisions even more crucial. With respect to the ongoing digital transformation and shift in power from organizations to consumers, we leverage digitally mediated wisdom of the crowd to build a forecasting model for better understanding chart success. Therefore, we draw on the literature of hedonic and experiential goods to investigate the relationship between crowd evaluations based on listening experience and popular music chart success. We track 150 song positions in reported music charts and also evaluate these songs via the crowd. Our model indicates that the wisdom of the crowd can improve forecasting chart success by almost 30% relatively to factors that have been earlier identified in the literature. However, this forecasting relevance is bound to certain conditions, namely the composition of the crowd, the underlying chart and market mechanisms, and the novelty of the musical material. In sum we find that crowd-based mechanisms are especially suited to forecast the performance of novel songs from unknown artists, which makes them a powerful yet very affordable decision support instrument for very uncertain contexts with limited historical data available. These findings can support recording companies to address the challenge of signing newcomers and thereby further enable the innovation system of the industry.

1. Introduction

It has always been crucial for music companies and labels to sign the right newcomer artists and select the best songs for a release in order to ensure a constant revenue stream from music product sales (Cameron, 2016; Dewan and Ramaprasad, 2014; Ordanini, 2006). These decisions are by no way perfect even today, which is reflected in revenues that are distributed according to the power law: revenues of few commercial successes compensate for the losses of many failures (Cameron, 2016; Strobl and Tucker, 2000). The reasons for this high number of failures and their challenges for firm survival can be found in the complex and uncertain market characteristics of the music publishing industry as well as in ever changing technologies.

First, the consequences of the digital transformation, such as music piracy, lower per unit revenues of digital content, and new business models have diminished the total profitability of the industry (Aguiar and Waldfogel, 2015; Benner and Waldfogel, 2016; Dobusch and Schüßler, 2014; Hiller, 2016; Lam and Tan, 2001). Second, the selection criteria of the recording companies are based on informal heuristics since the hedonic nature of music prevents an objective measurement of its quality. Third, direct and large-scale consumer engagement has been expensive. In order to increase the chances of reaching the consumer, the products are therefore developed and bundled according to the expected consumer taste formulated by the media gatekeepers (Hirsch, 1972; Ordanini, 2006). Fourth, consumer taste is susceptible to fashion cycles, socio-network effects and decision difficulties due to increasing choice of consumption, which result in diversity and uncertainty about expected demand. Fifth, recording firms select from the vast supply of creative content, bundle the selection into marketable entertainment products, and distribute to the consumers under high promotional efforts with the hope of passing the media gatekeepers, such as radio or television (Young and Collins, 2010). Once a product has passed these gatekeepers it will be repeatedly exposed to the largest possible audience, whose attention is then focused on this elevated set and ideally transformed into commercial transactions. However, the selection criteria of media gatekeepers are arbitrary and opaque from the perspective of artists and recording companies (Ordanini, 2006). Hence, there is high uncertainty

* Corresponding author.
E-mail address: dennis.steininger@wiwi.uni-augsburg.de (D.M. Steininger).

https://doi.org/10.1016/j.techfore.2019.04.016
Received 20 February 2018; Received in revised form 27 September 2018; Accepted 19 April 2019

about the outcome of the production process in the recording industry. Consequently, cultural producers, this means all actors involved in creating, producing, and marketing an experiential good, such as musicians, writers, directors, and record labels (Seiter, 1986) need more accurate and complete information about the potential of their creative outputs.

Existing research tries to accommodate this information need mainly by developing econometric prediction models in order to find patterns and regularities of successful experiential goods based on their original value, stemming from the intrinsic (e.g., song features such as tempo, loudness) or extrinsic features (e.g., chart history of an artist, consumer awareness) (e.g., Bhattacharjee et al., 2007; Dewan and Ramaprasad, 2014; Strobl and Tucker, 2000). Nevertheless, objective approaches fail to take into account the derivative value arising from the interaction between consumers and music (Lorenzen and Frederiksen, 2005; Prince, 1972). We tie in on this notion and intend to close the gap by building a forecasting model based on large-scale crowd evaluations of musical products (Holbrook, 2006). Crowdsourcing, i.e. giving usually small and simple tasks to an anonymous mass of people (the crowd) via the Internet, is seen as a tool to support management decisions due to its ability to leverage the wisdom of the crowd by aggregating results of these small tasks (Bonabeau, 2009; Budescu and Chen, 2014; Surowiecki, 2005). Research provides evidence that decisions made via the wisdom of the crowd of uneducated can be better than that of elite experts in a specific field (e.g., Budescu and Chen, 2014; Poetz and Schreier, 2012; Surowiecki, 2005). Furthermore, we argue that the crowd is a powerful supportive tool because it is also the targeted audience of popular music products. As commercial success in the music industry is determined through the number of listeners and record sales which is mirrored in chart rankings as a proxy (Dewan and Ramaprasad, 2014; Ordanini and Nunes, 2016), we formulate the following research question using these mechanisms: *Can the digitally mediated wisdom of the crowd improve forecasting of popular music chart success and thereby support artist investing decisions in the music industry?* We contribute valuable insights about the forecasting ability of the crowd for experiential hedonic products (i.e., music) via crowdsourcing to the body of research. We thereby also derive implications for human computation and decision support systems. Furthermore, our approach has high practical relevance to strengthen the position of recording companies. To support practical use of our approach, we propose best practices for each step of a crowd forecasting project that we have gained during this work.

2. Theoretical foundations and literature review

We began our endeavor through a qualitative interview with an expert from the recording industry. These first insights guide our literature review on popular music chart success, the decision-making process for new songs in the recording industry, the inherent features of music as a hedonic experiential product, and crowdsourcing to harness the wisdom of the crowd.

2.1. Performance and decision-making in the recording industry

Commercial success in the recording industry is tracked in charts, reflecting the relative performance of creative input in comparison to other material. Being successful in the charts is the most important key performance indicator of the industry (Dewan and Ramaprasad, 2014; Ordanini and Nunes, 2016). Recording companies track peak positions and duration on the charts in order to get insights about the quality and lifecycle of their musical material (Bhattacharjee et al., 2007). Since this material rather has a short lifecycle, promotional resources are allocated to the initial release in order to elevate the entry position, which is also often the peak position, because the material then falls to lower ranks and eventually drops out of the charts (Anand and Peterson, 2000).

The recording industry is considered a cultural industry system, which is characterized by a flow of cultural goods from the technical subsystem with the input producers (e.g., singers, songwriters) along a chain of mediating actors such as recording companies (managerial subsystem) and filtering/influencing actors such as radio stations (institutional subsystem) to the consumers and finally (hopefully) chart success (Hirsch, 1972). The managerial subsystem (i.e., recording industry) is characterized by a de facto oligopolistic structure with few multinational organizations, the major labels, and a long tail of small and medium independent labels (Rayna and Striukova, 2009). These organizations bundle and transform creative raw material into entertainment products. Thereby, the core actor is the Artist & Repertoire (A&R) department. Similar to the function of R&D departments in organizations of other industries, the A&R department is responsible for the discovery, selection, development, and management of talents (Mali, 2008; Neelamegham and Jain, 1999; Negus, 1992; Zwaan and ter Bogt, 2009). Depending on the corporate strategy, special composed A&R teams are screening the vast amount of creative input (e.g., demo records/tapes) for promising and trending talents to invest in, which will complement the existing portfolio. Nowadays, A&R managers mostly rely on their professional network in order to find new talents, since recommendations from peers and acquainted opinion leaders communicate credibility (Mali, 2008; Neelamegham and Jain, 1999; Ordanini, 2006; Zwaan and ter Bogt, 2009). The initial step is then the evaluation of the talent's marketability. Another important criterion is the quality of music. But there are no established indicators for the quality of hedonic experiential goods. Therefore, A&R managers use surrogate measures to describe the quality level, such as the innovativeness, uniqueness, and authenticity of a song.

In addition, to better circumscribe the quality, A&R managers often anchor the talents to examples of similar and already established artists (Negus, 1992; Ordanini, 2006; Zwaan and ter Bogt, 2009). Furthermore, intuition is commonly applied in ascertaining the potential of talents also in major labels. A&R managers trust their gut feeling, which they have acquired from professional experience, knowledge, and personal interest in music (Frith, 1996; Negus, 1992). Also, the career maturity of the talent is incorporated into the selection decision. An already established or mature artist bears less risk than does a newcomer because for newcomers it's even more problematic to foresee potential chart success. An interviewed expert from the recording industry stated: *"The hardest thing is to decide whether or not to sign a real newcomer without any chart history of whom you have received a demo tape only"*. This leads to decreasing innovation with new artists particularly in the dominating major labels and puts the entire recording industry system at risk (Benner and Waldfogel, 2016; Ordanini, 2006). Finally, criteria with respect to the personal characteristics of the artist are taken into account, such as the artistic skill level, the working attitude, the charisma, and the visual appearance (Negus, 1992). Based on these criteria, selection decisions are made. However, there are some inherent biases and drawbacks in this process (Lampel et al., 2000).

The decision to select a potential talent is not based on the actual ascertainment of consumer taste but on the perceived expectations of consumers' taste by media gatekeepers. Even though the creative output passes the media gatekeeper, it might therefore fail to be appreciated by the audience. The full responsibility lies at the A&R manager then (Hirsch, 1972; Young and Collins, 2010). In addition, the process of sorting out the vast quantity of input is very time-consuming. Also, past economic failures are exerting pressure on future selections (Benner and Waldfogel, 2016). The likelihood increases that A&R managers focus on skimming short-term potential by relying on proven properties and imitating successful competitors instead of developing a long-term career for the new artist and creating a unique style (Steinkrauß et al., 2008). Even though it is backed up by experience in the recording industry, decision-making based on intuition is always accompanied by informal heuristics, such as wishful thinking or pure randomness (Seifert and Hadida, 2006). Hence, large sums are invested in artists without the certainty of a return. Therefore scoring a commercial success is very speculative. Increasing the amount of accurate and complete information during this process would strengthen the position of the managerial subsystem. Hence, there is a need for new models and innovations to overcome this information bottleneck (Steinkrauß et al., 2008).

2.2. Music as an intangible, experiential and hedonic good

Music is an intangible experiential good (Dewan and Ramaprasad, 2014) that is "directed at a public of consumers, for whom [it] generally serve[s] an [a]esthetic or expressive, rather than a clearly utilitarian function" (Hirsch, 1972, p. 642). In addition, music is an integral part of human communication and fulfills emotional, social, or cognitive functions (Lacher and Mizerski, 1994). Research has criticized the traditional view on consumption and proposed the alternative hedonic consumption paradigm (Alba and Williams, 2013; Holbrook and Hirschman, 1982; Voss et al., 2003). "Consumption has begun to be seen as involving a steady flow of fantasies, feelings, and fun encompassed by what we call the 'experiential view'. This experiential perspective is phenomenological in spirit and regards consumption as a primarily subjective state of consciousness with a variety of symbolic meanings, hedonic responses, and [a]esthetic criteria" (Holbrook and Hirschman, 1982, p. 132). Pöyry et al. (2013) find that consumers of services which are inherently hedonic in nature have a higher willingness to participate in such services. The decision process when buying hedonic products is also substantially different to that of utilitarian products because consumers need to experience and 'feel' hedonic products (Okada, 2005). Consequently, since the evaluation of music is bound to the experience, there is no standard measure to compare music quality, because the derived value is due to the subject (Styvén, 2007). This makes it particularly difficult for decision makers in the A&R departments to select the 'right' songs that will be liked by the audience and commercially successful at the end.

Existing research has taken several paths to explain the quality of popular music and associated chart success for music-extrinsic and music-intrinsic features as depicted below.

2.3. Music-extrinsic features influencing chart success

Extrinsic features of music are concerned with those properties of a song, which do not stem from the song itself. These are all strategic and commercial factors associated to and manipulated by the recording industry. For example, the association with a major label highly influences the resources and capabilities that are available to promote a song or artist and subsequently also impacts chart success (Cameron, 2016; North and Oishi, 2006). Furthermore, the power of media gatekeepers was found to skew attention and awareness of the audience towards successful artists and songs in order to maximize the profit from advertising (Ahlkvist and Faulkner, 2002). Consequently, it has to be assumed that songs, which have been listed in official charts before, were successful in passing these gatekeeping mechanisms. Thus, artists with prior chart history release mainstream songs whereas unknown artists without prior chart history release newcomer songs (DiMaggio, 1977). Prior chart success of an artist was therefore also identified as not only increasing the chances to be signed by a major label but also as an explaining variable of future chart success for songs of that specific artist (Benner and Waldfogel, 2016; Ordanini, 2006).

Further, awareness is found to capture the omnipresence of an artist or song in the evoked set of consumers (Hoyer and Brown, 1990). Since consumers are rather passive in acquiring new information, choice decisions are made on the basis of this evoked set and it is very important for artists to position their material accordingly (Macdonald and Sharp, 2000). Therefore, awareness is a point-in-time observation of the current "buzz" about a song or artist in comparison to other songs, which is also used as an indicator for marketing effectiveness (Goel et al., 2010; Lavidge and Steiner, 1961). As a result, extant literature argues that the higher the awareness for a particular song (e.g., buzz on social media), the better will be its chart success (Dewan and Ramaprasad, 2014). Finally, repertoire refers to the origin of the respective song. Domestic repertoire is defined as musical productions created in the respective market, whereas international repertoire has foreign origin. Existing research suggests that the spill over effects and popularity of international repertoire increases the likelihood of chart success (Legrand, 2012).

2.4. Music-intrinsic features influencing chart success

In contrast to the extrinsic features, the intrinsic features are concerned with the elementary properties of a song, such as tone, tempo, or rhythm. As introduced above, human beings do not evaluate music based on its individual technical features. Instead, the song is evaluated as a whole on a higher semantic layer (Celma and Serra, 2008). Also, current popular music styles are very similar in terms of their physical and emotional tone, and preferences of naïve listeners can be approximated to certain extend by using the basic audio features evaluated by experienced listeners (Dainow, 1977). Hence, we may assume that in terms of mass audience taste, there are common underlying audio features responsible for the chart success (Rumsey et al., 2005; Sloboda, 1991). Following this avenue, a team of researchers from Bristol University has developed a machine learning algorithm that can calculate a score for estimating the hit potential of a song based on its audio features (Ni et al., 2011).

2.5. Harnessing the wisdom of the crowd to improve decision-making

Extant literature has elaborated on several mechanisms such as prediction markets (e.g., Berg et al., 2009; Decker et al., 2011; Wolfers and Zitzewitz, 2004), expert ratings (e.g., Förster and von der Gracht, 2014; Keller and von der Gracht, 2014), or simulations (e.g., Palmer et al., 2015; Parker et al., 2015; Zhao et al., 2016) to foresight future outcomes and events. Only recently research has gained interest in the so called 'wisdom of the crowd' that is enabled by digital platforms. The inventor of the term crowdsourcing, Jeff Howe, posits that "given the right set of conditions, the crowd will almost always outperform any number of employees" (Howe, 2009, p. 11) also for complex tasks. Others argue that decisions made via the wisdom of the crowd of uneducated can be better than that of elite experts in a specific field (e.g., Budescu and Chen, 2014; Poetz and Schreier, 2012; Surowiecki, 2005). Recent research has built on these ideas and harnessed the wisdom of the crowd to improve forecasting for very complex and uncertain situations and support decision-making of managers (Bonabeau, 2009).

Besides analyzing the power of (passive) crowd wisdom via numbers of posts, which were not intentionally made to forecast (Briscoe et al., 2016), in social media to predict flu waves or political elections (e.g., Bollen et al., 2011; Forlines et al., 2014; Franch, 2013; Jin et al., 2010; Paul et al., 2014; Tumasjan et al., 2011), one interesting research stream harnesses crowdsourcing to recruit non-experts with the specific aim to find new innovations (de Mattos et al., 2018; e.g., Garcia Martinez, 2017; Majchrzak and Malhotra, 2013; Mortara et al., 2013; Schemmann et al., 2016) or forecast future outcomes (active crowd wisdom).

For example, extant research provides evidence that crowdsourced sales forecasts of packaged consumer goods can be better than traditional company internal methods in specific contexts (e.g., Lang et al., 2016; Mortara et al., 2013). Stock market recommendations by amateurs are also argued to outperform other measures of prediction (e.g., Gottschlich and Hinz, 2014) as well as weather forecasting via the crowd (e.g., Niforatos et al., 2014). Some research projects look at the use of crowd wisdom to enhance intelligence and secret services (e.g., Halman, 2015) while other authors not directly look at forecasting outcomes but analyze how crowdsourced forecasting can be included in internal firm processes and strategies (e.g., de Mattos et al., 2018; Hosseini et al., 2015). Others contribute methodological insights on performance differences between several forecasting tools and crowdsourcing (e.g., Flostrand, 2017) or on crowd specifics (e.g., Budescu and Chen, 2014; Prpić et al., 2015; Turner and Steyvers, 2011).

However, to the best of our knowledge, neither of the existing research looks at the active wisdom of the crowd to forecast sales of hedonic and experiential goods nor at musical products particularly. As introduced above, music is an experiential good and therefore no objective measures on good or bad can be applied but we rather have to analyze and aggregate subjective emotions, feelings, and intentions of its consumers (Holbrook and Hirschman, 1982). We argue that this makes active wisdom of the crowd particularly interesting to forecast future success of musical products because it is very difficult even for a small group of experts in A&R departments to objectively elaborate on music quality. Crowdsourcing can help by collecting subjective feelings, intentions, and emotions of listeners via open online marketplaces and aggregate the results in order to accommodate an information need in the recording industry. Since crowdsourcing is not restricted to a certain industry but the collection of crowd wisdom can be tailored to the specific information needs of experiential products, we suggest that it is well suited to solve the pending need for more accurate and complete information in the recording industry, namely the extension of the managerial subsystem with an external support subsystem (Hirsch, 1972; Mason and Suri, 2011). Filtering and ranking the vast amount of incoming creative material can support A&R managers in their decision-making by reflecting emotions, the taste, and intentions of the crowd and resembling a de facto objective standard for quality evaluation. Consequently, the application of a crowd support subsystem is an alternative (bottom-up) approach to disintermediate traditional taste-making and gatekeeping mechanisms. Instead of releasing songs and hoping that song releases will break even, A&R managers could rely on the crowd to evaluate the potential of this material prior to conducting large investments in terms of marketing and artist development. However, such a crowdsourced forecasting instrument needs to be tailored to the specifics of music as an experiential product.

2.6. Hypotheses development

Based on the ideas introduced above, our suggested model centers around explaining the dependent variable "chart success". To do so, we derived the music-extrinsic (e.g., chart history of an artist) and music-intrinsic features (e.g., tone or tempo) that existing literature has commonly shown to be predictors of chart success as controls. They build our base model that we then try to improve in its predictive performance by adding our aggregated crowd evaluation variables (see Fig. 1). The latter were developed based on the literature of consumer behavior for hedonic products as will be explained in further depth in the following.

As there are no official global popular music charts, the definition of chart success is bound to a certain music market. Thus, based on the summary of the top 20 music markets in the 2011 IFPI annual report of the global music industry, we have selected a relevant set of markets with a trade value greater than one billion US dollars (IFPI, 2011). The resulting markets are the United States ($4.1677 bn), Japan ($3.9586 bn), Germany ($1.4122 bn) and the United Kingdom ($1.3785 bn). We decided to investigate the rather conservative German market, where the tipping point of the digital transformation has not yet been reached (as of this study) and the majority of value is still generated by physical sales (BVMI, 2017; Lam and Tan, 2001). Furthermore, the music charts for this market are solely compiled on the basis of physical and digital sales not incorporating music streaming at the time of data collection. Hence, the uncertainty of outcome is higher, since traditional gatekeeping mechanisms are mostly intact, meaning consumers have less choice to influence the selection of musical material. Finding a relationship between German chart success and active crowd evaluations would yield valuable information especially for the recording companies in the managerial subsystem. Given the experiential nature of music, we need a framework that reflects the hedonic consumption paradigm introduced above and can be connected to chart success.

In addition, the construct for the crowd evaluation has to allow crowd workers to discriminate the quality of the respective songs based on the post-consumption evaluation of the listening experience. Thereby, we would like to focus our attention solely on the evaluation of the musical stimuli. Even though it has been acknowledged to influence chart success, we are exempting any visual appearance of the cultural producer from our study (Hargreaves and North, 1997). Since the aggregated crowd evaluation is calculated on the basis of individual consumer evaluations we have screened the consumer research literature for adequate constructs. Predicting success in consumer research is concerned with the performance or non-performance of admired behavior. As data of observed and actual behavior is absent in most research studies, it seems reasonable to suggest, that the widely used intention to perform certain behavior should serve as a surrogate for and predictor of actual behavior (Morwitz and Schmittlein, 1992). Behavioral intention is defined by Ajzen (1991, p. 181) as an "[indication] of how hard people are willing to try, of how much an effort they are planning to exert in order to perform the behavior". Thus, the stronger an individual's motivation to perform a specific behavior, the more likely should the individual convert the intention into actual behavior. There is an ongoing debate about behavioral intention constructs due to empirical concerns about the predictive validity (Kalwani

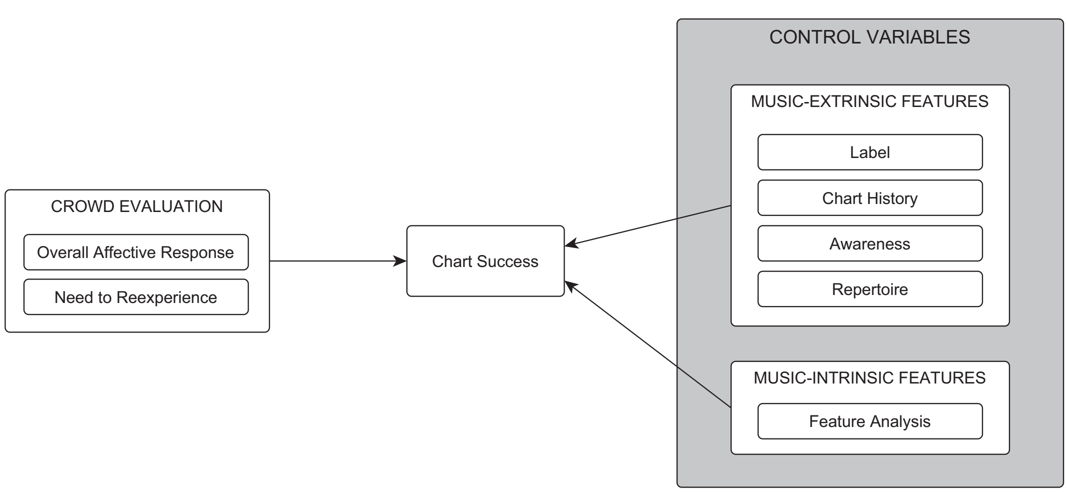

Fig. 1. Conceptual model.

and Silk, 1982; Sheeran, 2002; Sutton, 1998). Several researchers argue extensively about the discrepancy in the intention-behavior relationship, meaning that consumers intending to perform certain behavior essentially do not act afterwards (Sheeran, 2002). We therefore decided to include an additional evaluative post-consumption measure (the need to re-experience) into the model.

We adopt the hedonic consumption model by Lacher and Mizerski (1994) because it attempts to explain why consumers purchase music and therefore connects a detailed model of the hedonic consumption paradigm to behavioral intention measures. Even though this model has been established prior to the digital transformation of the recording industry, Dilmperi et al. (2011) have successfully applied it to explain individual differences in the profiles of music pirates and genuine music consumers. Further, the hedonic consumption model focuses on the listening experience solely evoked by the musical stimuli, which is in accordance with our intention. The model contains a global evaluative construct as well. Consequently, we adopt the two relevant constructs from the model, namely the need to re-experience and the overall affective response that were found to influence consumers' music purchase intentions.

According to Zajonc (1980), there are three aspects to consider when it comes to the judgmental character of the overall affective response. First, the value of hedonic experiential goods is holistic. This 'gestalt approach' refers to the notion that the value is derived from the overall impression about the object and its global features rather than from the individual ones (Agarwal and Malhotra, 2005). Second, affective judgements about an object are reflecting the inner self of the individual, meaning the respective values and personality (Zajonc, 1980). Third, individuals are not able to verbalize the reasons behind their affective judgments (Mittal, 1988). Hence, we define the overall affective response as the global valence judgement of a crowd worker about liking or disliking a piece of music based on the listening experience (Lacher and Mizerski, 1994). Incorporating the above-mentioned selection criteria of the A&R managers, the subjective evaluation of a crowd worker about the hedonic value of a song constitutes a performance benchmark in comparison to other songs and therefore shows the unique selling point of a song. In addition to the assumed relationship in the hedonic consumption model, there have been other researchers confirming the significant relationship to purchase intention (e.g., Mizerski et al., 1988). We generalize these findings to behavioral intentions related to music consumption and formulate the following hypothesis:

H1. The aggregated affective response from the crowd is positively related to chart success of a song.

The need to re-experience is an individual's desire to consume a piece of music again (Lacher and Mizerski, 1994). However, the consumer does not necessarily have to purchase the music in order to control the re-experience nowadays, since there are several choices for consuming music. Therefore, we do not emphasize the temporal control as being the major aspect of this construct, but the multitude of behavioral intentions related to music consumption. Consumers might directly influence chart success by re-experiencing the music or indirectly influence it by recommending music to friends. This construct is a self-reflection about the personal fit of the song with the lifestyle and status expressions of the consumer, since only credible and meaningful information is recommended to others (Lacher and Mizerski, 1994). Electronic word-of-mouth is the most important promotion strategy for artists, since on the one hand, music is evolving more and more into a service and on the other hand, it is an indicator for growth in artist equity (Howe, 2009; Styvén, 2007). Positive word-of-mouth leads to a potential increase in the fan base of the artist, which in turn increases the attention of and exposure to the mass audience as well as increases the likelihood of reaching the critical mass for entering the charts (Hayes, 2008). We assume that converting crowd workers into consumers and advocates of a particular artist is an indicator for the chart success potential of a song. Therefore, we define the need to re-experience as a consumer's desire to re-experience and recommend a song based on the listening experience and formulate the following hypothesis:

H2. The aggregated need to re-experience from the crowd is positively related to chart success of a song.

As mentioned above, there are several factors in the production process of a piece of music that may not be controlled by the cultural producers. Hence, we do not only control for traditional items such as age, gender, country but also for mood, music experience, musical affinity, formal music training, music employment, music preferences, purchase preferences, and consumption preferences of crowd workers but also take into account factors that have been identified earlier. Existing research suggests such criteria for prediction models (e.g., Asai, 2008) based on two main areas. Hence, we control for extrinsic factors such as the awareness that a song has due to marketing efforts or 'buzz' (Goel et al., 2010), the label association (major vs. independent) (Steinkrauß et al., 2008), the chart history of the song's artist (DiMaggio, 1977), repertoire (domestic vs. international) (Legrand, 2012), and the positive musical features (intrinsic features) that are concerned with the elementary properties of a song, such as tone, tempo or rhythm (Ni et al., 2011).

Based on prior studies that provide evidence of the predictive power of the crowd, we suggest that our crowd constructs can improve the forecasting of popular music chart success in comparison to the established elements from the literature (our controls). We therefore posit:

H3. Adding the crowd evaluation constructs to the model will improve the overall predictive power to explain chart success of a song.

3. Method

We use structural equation modelling (SEM) since this enables us to transform our set of hypotheses into a path diagram with latent variables to test and estimate the relationship between the crowd workers' song evaluations through a survey on our constructs, controls, and chart success at the same time. We are more interested in the forecasting relevance of these evaluations than in the accurate estimation of the respective model parameters. We also have a rather small number of songs and therefore apply the PLS-SEM approach (Cepeda Carrión et al., 2016; Ringle et al., 2012).

3.1. Operationalization

We operationalize our dependent variable chart success with the peak position as the single-item manifest indicator for each song. For the period of three months, we are provided with weekly rankings consisting of the top 100,000 sold digital and physical single tracks by the world's leading entertainment data provider. These rankings have to be weighted in order to reflect a composite approximation of the actual chart position. The weighting scheme is defined according to the digital and physical share of the German market's trade value (IFPI, 2011). Nevertheless, we only have to calculate the composite chart position for songs from the sample where physical as well as digital sales ranks are available (see Eq. (1)). Hence, the endogenous variable is measured on an ordinal level and is operationalized as follows:

$$\mathrm{Comp}_{ij} = \begin{cases} x_{ijd},\ x_{ijd} > 0,\ x_{ijp} = 0, \\ (w_p * x_{ijp}) + (w_d * x_{ijd}),\ x_{ijd},\ x_{ijp} > 0 \\ x_{ijp},\ x_{ijp} > 0,\ x_{ijd} = 0 \\ 0,\ x_{ijd},\ x_{ijp} = 0 \end{cases} \tag{1}$$

with:

$i = 1 \ \dots$ N (number of song); j = 1 … M (number of week); x_{ijd} = digital sales rank of song i in week j; x_{ijp} = physical sales rank of song i in week j; $w_d = 0.13$ (digital share of trade value); $w_p = 0.81$ (physical share of trade value); $Comp_{ij}$ = composite approximated chart position of song i in week j.

Since the overall affective response is concerned with a global feeling and subjective evaluation of a sample song after consumption, we intended to find a construct best suited for this kind of judgement. Nevertheless, the construct should also reflect all the different facets of the selection criteria of the A&R managers and should cover the hedonic evaluation of the listening experience. Consequently, as in the case of the Hedonic Consumption Model, we have adopted the global index of evaluation (Lacher and Mizerski, 1994). This construct is a summary of bipolar adjective pairs, which are correlating strongly with the "good-bad" pair and thus have an affective overtone. There are nine bipolar adjective pairs, which are reflectively measured on a 7-point Semantic Differential scale, namely bad-good, distasteful-tasty, dull-exciting, tasteless-tasteful, unimaginative-creative, untalented-talented, unpleasant-pleasant, forgettable-memorable and boring-interesting (Osgood et al., 1957). Thus, the level of measurement is interpreted as an interval scale.

The operationalization of the need to re-experience shall display the crowd worker's self-reflection about her/his support for the artist based on the listening experience. Thus, according to Hayes (2008) post-consumption behaviors, such as the recommendation intention, will ideally reflect a crowd worker's advocacy for the artist. Therefore, we decide to use the existing measurement model from the hedonic consumption model as well, because it also has high reliability (Coefficient Alpha of 0.90) and covers there-experience as well as recommendation aspect (Lacher and Mizerski, 1994). This construct measures three reflective items on a 7-point Likert scale (strongly disagree/strongly agree), such as "I would enjoy listening to this song again", "I would like to play this song for my friends" and "I want to be able to listen to this song whenever I feel like it" (Likert, 1932). In the same way, the scale is considered to be of interval level of measurement (also see Table 1 for an overview). The operationalization of the control variables can be distinguished in two groups. On the one hand, we have dichotomized (binary) variables indicating the presence or absence of a certain attribute. This applies for the label affiliation (LAB), which attests, whether a major label releases a sample song or not. We acquired the label information from the iTunes Music store and compared them with the information provided in the sales rankings. Further, the chart history (CHH) reflects the career status of an artist of being either a newcomer or mainstream artist, based on the previous chart listings and achievements. Hence, we queried several online chart databases for any previous chart listings of the artists.

Besides, we included the origin of the repertoire (REP). Therefore, we assigned samples to the international repertoire category based on the appearance on the iTunes Music stores of multiple countries. Further, a sample is regarded as domestic repertoire, if it is only listed in the local iTunes Music store. On the other hand, there are variables, which are measured on ordinal scales. For the awareness (AWA) variable, we have chosen the Google Trends search volume index, which is used to indicate the relative importance of a certain search term in comparison to other search terms in the same category and period of time. Finally, we measure the potential of the intrinsic features (FAN) of the respective sample song with the hit score computed by Ni et al. (2011).

3.2. Data collection

We collect data on 150 songs and let each of them be evaluated by 20 crowd workers on Amazon Mechanical Turk resulting in a total of 2.852 observations. We use the averaged values of the crowd evaluations for each song. Since there is strong cultural proximity of the German music market to the Anglo-American markets in terms of popular music culture (i.e., chart rankings of the US and Germany often show large overlaps) and songs in English language dominate the German top-10 lists (Achterberg et al., 2011), we focus on crowd workers originating in Anglo-American countries with similar popular music culture (i.e., US, UK, Canada). The crowd on Amazon Mechanical Turk is seen as roughly resembling the Internet population (Ipeirotis, 2010b). We use 90-s samples of the songs. Furthermore, we are facing the same uncertainty as the A&R managers in determining which new releases will make it into the charts. Thus, in order to determine sample eligibility, we have to follow a lenient strategy. First, we restrict the language parameter to "English" due to its characteristic as global language and wide application in international repertoire. Second, we include only songs from most popular genre dimensions (Rentfrow and Gosling, 2003). As we have mentioned in the introduction, previous research focused their attention mainly on historic data in order to build their prediction models (e.g., Asai, 2008; Gazley et al., 2011; Ni et al., 2011). However, our intention is to forecast chart success based on evaluations of newly released songs. Even more vital is to assure that the crowd workers have not yet been exposed or stumbled upon the respective song since this familiarity bias may flaw the true evaluation of the song (Zajonc, 1980). Thus, we have to test the samples on the day of the release and actively control for familiarity with the particular song and/or artist.

We follow recommended practices of previous research studies in order to establish appropriate quality mechanisms prior, during, and after the task execution. Apart from limiting to workers from western countries prior to participation, they are randomly selected. Crowd workers initially had to pass the qualification test, since we have to ensure that they are physically able to hear. Therefore we asked them to listen to several animal sounds and to tick the corresponding animal

Table 1
Operationalization of main variables.

Variable	Operationalization	Type	Cite/source
Chart success	Composite peak position (digital/physical)	Weighted score (1–100,000)	Media Control GfK Int.
Overall affective response	Bad-good, distasteful-tasty, dull-exciting, tasteless-tasteful, unimaginative-creative, untalented-talented, unpleasant-pleasant, forgettable-memorable, boring-interesting	7-Point semantic differential	Holbrook and Huber (1979)
Need to re-experience	I would enjoy listening to this song again, I would like to play this song for my friends, I want to be able to listen to this song whenever I feel like it	7-Point Likert scale	Lacher and Mizerski (1994)
Label	Major vs. independent	1 major, 0 otherwise	iTunes Music Store
Chart history	Mainstream vs. newcomer	1 mainstream, 0 otherwise	Chart databases
Repertoire	International vs. domestic	1 international 0 otherwise	iTunes Music Store
Awareness	Google Trends search volume index	Score (0 – 100)	www.google.com/trends
Feature analysis	Scoreahit potential score	Weighted score (0 – 12)	www.scoreahit.com

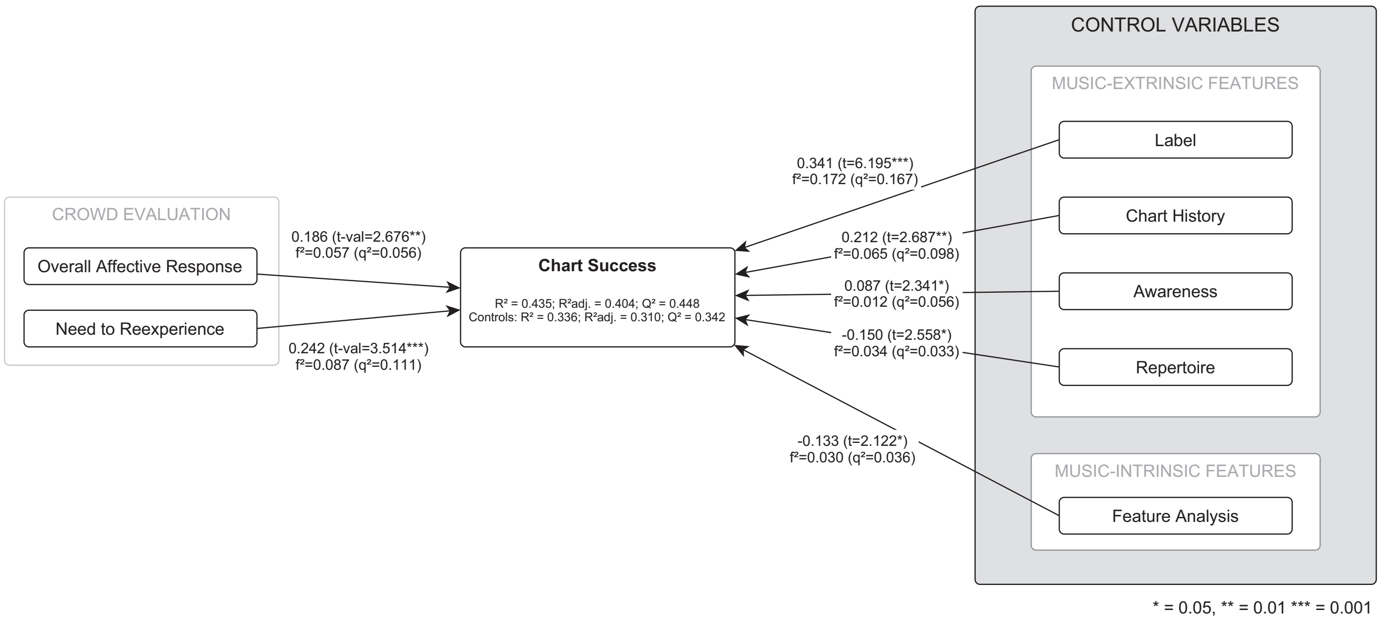

Fig. 2. Model estimation.

(Sprouse, 2010). To further protect our survey from satisficing, we also implement attention checks during execution, which are randomly positioned among the pages and scale items. Further, after evaluating the sample song, we check for too short listening and answering times as well as suspicious answering patterns. In order to indicate that the crowd workers actually completed the survey, they are presented a unique completion code or a similar generated token which they have to paste into the respective textbox (Mason and Suri, 2011).

4. Data analysis and estimation results

After HIT publication we recorded 3890 attempts to work on our task. The incorporation of the chosen quality management mechanisms proved valuable, because it separated the useful submissions from the useless ones. Consequently, the final 2852 approved submissions resemble 73.32% of the initial attempts. The entire song (150 songs) evaluation task was accomplished within 3.32 days. As a consequence of the quality check, the average ratings per song reduced from 20 to 19.01. The average completion time for a HIT corresponds to 7.22 min and is about half of the initially allotted time. Based on the average time, the effective hourly wage increased to $3.33, which complies with related work on Amazon Mechanical Turk (Berinsky et al., 2012; Ipeirotis, 2010a; Jia et al., 2017). In terms of the composition of the crowd, 74% of the participating crowd workers are in the age groups between 18 and 34, which are highly relevant for the recording industry (Steinkrauß et al., 2008). Further, 60% of the crowd workers are female and 96% of them earn less than or equal to $60,000 per year. However, to ensure that the crowd has enough influence on chart success, we analyzed the percentage of music product buyers within our sample, which is a bit higher than for the internet population in general.

With the focus laid on crowd workers from Western countries with similar popular music culture, we already expected that a large share (97%) of them originates from the US. The crowd is also very active on AMT. 40% of crowd workers spend between 5 and 15 h per week on AMT and another 22% works between 15 and 30 h per week on AMT.

The absence of crowd workers with profound musical background let us assume that the appraisal of music seems to be a universal phenomenon (Kazai, 2010). Further, the crowd workers mainly consume music by digitally accessing (37%) and digitally owning (51%) the music. On the aggregated song-level, we have observed that the values for the dependent variable are primarily determined by the digital peak position, because the majority of songs (78%) did not enter physical sales ranks. We have also found that songs can be classified into three chart ranges, namely the Top 500 (43%), the long tail (47%) and the non-entries (10%) (Anderson, 2006). In general, major label material is dominating the dataset (59%), especially the Top 500 (86%). This is in accordance with the division of power in the global recording industry, where the market shares of the big four labels outperform the remaining labels (Sernoe, 2005).

Also, mainstream artists (73%) outweigh the newcomer artists (27%) in the dataset, whereby the mainstream artists are leading the Top 500 (83%) and the long tail (64%). Approximately nine out of ten samples (86%) are associated as being of international repertoire. This ratio is also consistent throughout all chart ranges. The average awareness of the Top 500 (15.56) is higher than for the long tail (4.81), which in turn is higher than for the average awareness for the non-entries (0.00). In terms of the feature analysis score, our dataset shows roughly a normal distribution around the mean score of 5.81. Hence, the audio features of the majority of songs are very similar but do not necessarily comprise hit potential. The average score for the non-entries (6.69) was greater than for the Top 500 (6.03), which was greater than the score from the long tail (5.60).

For the evaluation of our model, we excluded the non-entries from the estimation, because they did neither enter the digital nor the physical sales ranks and therefore do not provide information for the dependent variable. With the remaining sample songs (n = 134), we focused on evaluating the reliability and validity of the measurement model of our two crowd constructs, because the single-item control variables are not part of the core model (Ringle et al., 2012). A preliminary factor analysis resulted in two factors with eigenvalues > 1, whereas the first factor accounts for 59% of the variance and resembles the overall affective response. The second factor explains another 21% of the variance and is associated with the need to re-experience (Fig. 2).

All indicator loadings were high on their intended factor and exceeded the suggested threshold (0.700). Further, the values of Cronbach's Alpha (OAR: 0.962; NTR: 0.956) and the composite reliability (OAR: 0.967; NTR: 0.972) surpass the required threshold values for acceptable results (0.800) (Cronbach, 1951). The values for the average extracted variance of the overall affective response are considerably lower than the ones of the need to re-experience. Nevertheless, it is still above the critical value (0.500) (Vinzi et al., 2010).

Internal consistency is also reflected with indicator reliabilities ranging from 0.527 to 0.971 and exceeding the suggested decision criterion (0.400) (Bagozzi and Yi, 1988). In addition, all indicators have high significance. Consequently, we assume that both latent variables have very high levels of internal consistency and convergent validity. Moreover, we tested the Fornell-Larcker criterion against our data in order to ascertain discriminant validity. Thereby, the squared correlation between overall affective response and the need to re-experience (0.045) is considerably lower than the AVE of both variables (Fornell and Larcker, 1981). Thus, we also assume discriminant validity of our measurement model.

In contrast to the exemption from the analysis of the measurement model, we have incorporated all paths and variables in the analysis of the structural model (see Fig. 1). The execution of the PLS algorithm with the parameters set to path weighting scheme and standardized values showed that our proposed model has an explanatory power (R2) of 43.5% of the variance in chart success. Corrected for possible biases stemming from the number of predictors, the adjusted explanatory power (R2adj.) reduces to 40.4%. Further, the blindfolding procedure with an omission distance of 7 resulted in a predictive relevance (Q2) of 44.8% (Fornell and Bookstein, 1982).

Compared to an estimation of a model consisting only of the control variables, the inclusion of the crowd evaluation constructs increases the explanatory power by approximately 30% relative to the explained variance of the established control variables from the literature only. The main predictors of the model seem to be the label affiliation, prior chart history, and our two crowd evaluation constructs. This shows that our hypothesis H3 holds true.

5. Discussion and implications

5.1. Discussion of implications for chart success forecasting

In the beginning, we posed the research question, whether the digitally mediated crowd is able to improve the forecasting of popular music chart success. Based on our results, we have found significant and positive relationships between the crowd evaluation constructs and popular music chart success. The results from the estimation have shown, that the path between the overall affective response and chart success has predictive relevance (${q_{OAR}}^2 = 0.056$). In addition, the path between the need to re-experience and chart success has predictive relevance as well (${q_{NTR}}^2 = 0.111$). By summing up both values, the predictive relevance of the crowd is greater than zero (${Q_{crowd}}^2 = 0.056 + 0.111$) with the need to re-experience having twice as much predictive power than the overall affective response. Hence, we have acquired empirical evidence for the rejection of the null hypothesis ($0.166 > 0$) and are able to answer our research question: Yes, the crowd can support the forecasting of chart success.

However, even though we almost replicated the reliability of the measurement models of the latent constructs from the hedonic consumption model, this statement has to be qualified with respect to the overall impact of our model. It seems that there are several aspects influencing the magnitude of this forecasting ability. First, the degree of industry manipulation is confounding the impact of the crowd, meaning that even though the songs are rated on an equal and unbiased basis, traditional gatekeeping mechanisms and mass promotion dynamics are still intact (Hirsch, 1972). There is an imbalance in the allocation of resources and exposure, which is not in the crowd's locus of control but influences its forecasting relevance. With increasing chart position the impact of the crowd evaluation becomes less important and industry manipulation is dominant in these chart ranges. Hence, we have had a closer look at the degree of industry manipulation per chart range, whereby we distinguish between songs that have received any manipulation, songs with prior chart history, songs with major label affiliation and songs with prior chart history and major label affiliation.

We find that the crowd shows discriminatory power in evaluating songs without any industry manipulation (i.e., for newcomers). In this case, we observed that the better the evaluation of the crowd the higher is the chart position. In fact, there were two hidden gems in the dataset, which were neither affiliated to a major label nor previously exposed to the masses but highly appreciated by the crowd and ranking high in the Top 500 chart range (25th and 368th position). Targeting investments and marketing budgets towards these songs might have been far more profitable than diversifying the investment on a myriad of failures while blowing up the marketing budgets in order to distinguish the songs from other releases.

The German music market may be regarded as conservative, because it is still dominated by physical sales which is also reflected in its chart rankings (BVMI, 2017; Lam and Tan, 2001). However, due to the on-going digital transformation, individuals nowadays have an increased choice of consuming music without the necessity of purchasing physical records or digital downloads (Benner and Waldfogel, 2016; Fox and Wrenn, 2001). In fact, there are several other revenue streams, which may have the potential to supplant these sales in the future, such as revenues generated from subscription services like Spotify or live performances (Hiller, 2016). In our study, we found that the crowd already largely consists of digital natives, since digital consumption patterns (88%) are dominating their physical counterparts (12%). In addition, the focus of consumption also switches from owning (62%) to accessing (streaming) music (38%).

Hence, even though a song is positively evaluated and there is a strong need to re-experience, the predictive impact of the crowd might not be as high as it could be because the mechanisms of the German music charts at the point of data collection does not take into account the consumption preferences (streaming music) from the crowd. This implies that updates in the chart ranking calculation mechanisms to respond to digitalization might further align the crowd with the used dependent variable and thereby increase predictive relevance. Nonetheless, as being a market information regime for the recording industry, the charts should reflect all possible consumption choices in order to determine the true overall popularity of a piece of music (Anand and Peterson, 2000). The large share of digital-only releases in our dataset supports this view.

5.2. Implications for crowdsourced forecasting in general

We summarize implications and learnings for crowdsourced forecasting below (see also Table 2). Our study provides evidence that the development of a crowd forecasting instrument that is specifically tailored to a context can be helpful in predicting future outcomes. We particularly elaborated on the specifics of the product under study (i.e., popular music) to develop our forecasting model. Theoretical grounding can support this, for example by looking at the product type (e.g., experiential, hedonic good).

Carefully selecting the crowdsourcing platform as well as participants will probably also influence forecasting quality. We argue that results should be best for products and crowd evaluation instruments that are aligned with the average crowdsourcing participants. Crowdsourcing platforms commonly offer several ways to limit to certain groups of participants (e.g., countries, age, qualifications). One such very important tool is the limitation to participants that have successfully completed several tasks (500 are often seen as a useful level here) and also have an approval rate above a certain threshold. We experienced 95% percent or above as a helpful level (Jia et al., 2017).

When setting up a crowdsourced forecasting project, some thoughts should also be spent on anti-satisficing mechanisms and attention checks. We found that around 30% of the participants show suspicious patterns or fail included checks. This is similar to other results not only for crowdsourced but also for panel surveys (Berinsky et al., 2012; Jia et al., 2017). Besides standard anti-satisficing mechanisms and time

Table 2
Suggestions for the setup of a crowd forecasting project.

Step	Suggestions
1. Development of crowd evaluation constructs	● Develop crowd forecasting constructs based on established theory of the specific field to be forecasted (i.e., popular music in our case). ● Tailor the constructs to the inherent nature of the object under evaluation by the crowd (e.g., experiential versus utilitarian product or service). ● Include and collect important control variables into the forecasting model that have been identified in earlier literature (e.g., awareness of artist). ● If applicable, introduce measures to rule out social desirability answers of workers.
2. Platform selection	● Select the platform with regards to its population and the products target groups (e.g., cultural background, language, demographics).
3. Participant selection	● Limit participants on the selected platform to languages/cultures/countries matching your inquiry (i.e., US in our case). Some platforms offer geo-location here, which proofed helpful to limit based on the real location of a worker. ● Amazon Mechanical Turk (MTurk) is the most prominent platform and should match most western and English-speaking crowd forecasting projects. Its population stems mostly from the US and India. See Berinsky et al. (2012) and Huff and Tingley (2015) for detailed analyses of MTurk demographics. ● Use qualification management provided by the selected platform. For example, MTurk offers limiting the participation in a task (HIT) to specific workers. ● Limit to workers that have > 500 approved tasks and a general task approval rate of over 95% (Jia et al., 2017). ● If applicable, limit to workers that have an adult content qualification. We applied this due to potential adult language in the songs. ● If applicable, limit to one time participation per worker only. ● Set up payment according to the length of your evaluation and time pressure (also see Buhrmester et al., 2011).
4. Quality measures during crowd evaluation	● Start the evaluation in the morning of the workers' countries (we found that starting at 8 am in the specific country works well while starting at night introduces a lot of issues with drunk or tired workers). ● Introduce anti-satisficing and attention checks throughout your evaluation. Our experience from the project shows that around 20–30% of the completed tasks show satisficing behavior. ● Too simple anti-satisficing checks showed to not work well as some workers seem to be already trained for them. A well-working check in our project hints in the description texts on some of the survey pages to not click on the next button but on a headline instead, which proofed very helpful. Another useful check is to include some items twice and later check for strong differences between answers. See Oppenheimer et al. (2009) for more details on these anti-satisficing techniques. ● Introduce further quality checks pertaining to your evaluation object, if applicable. For example, crowd workers had to play different animal sounds and mark corresponding animals to ensure that they were technically and physically able to hear music. ● Track times (e.g., for survey completion, listening to music, watching video). ● Generate a completion code at the end of your evaluation survey.
5. Quality measures after crowd evaluation	● Exclude data of workers that failed anti-satisficing and quality checks. ● Exclude data of workers that took < 50% of the average time. ● Exclude data of workers that did not enter the correct completion code. ● Analyze answer patterns for abnormalities (see Mason and Suri, 2011). ● Analyze and include control variables (e.g., we included measures whether the workers were already familiar with a song, worked in the music industry, or had a certain music taste).

tracking, we introduced another check that was explicitly developed for our specific setting. Participants had to listen to different animal sounds and then tick the animal name to ensure that they were able to technically and physically hear the music.

After data collection via the crowd, it is important to undertake data cleansing for eliminating biased or fraudulent data (Jia et al., 2017). General quality measures such as the removal of extremely short survey cycles or, in our case, participants only listening to a fraction of a song should be applied. Furthermore, participants that failed checks should be eliminated from the data set as well as suspicious answer patterns (e.g., only strongly agree ticked for all answers). For details on such pattern analysis see also Mason and Suri (2011). A comprehensive overview for the suggested steps and quality measures of a crowd-sourced forecasting project can be found in Table 2.

5.3. Implications for the recording industry system

As introduced above, the basic production process in the recording industry can be structured into a net of related subsystems. Combining the discussed results from the model estimation with our suggested application of crowdsourcing for the recording industry, we suggest adding an external support subsystem in the form of an open marketplace to the industry. Thereby, the integration of the support subsystem is re-intermediating the recording industry bottom-up relying on the principles of crowdsourcing (Steinkrauß et al., 2008). In addition, the support subsystem could also be built on a hybrid approach by using crowd evaluations and further metrics, such as automatic feature analysis. Based on this notion, we are able to formulate implications for the major stakeholders in this research.

The crowd can be used to filter out the hidden gems from the vast amount of creative input. Thereby, the crowd may either be used to filter the input for the A&R department of recording companies from the managerial subsystem or to filter the input and to provide feedback for the do-it-yourself artists favoring the direct distribution to the consumers over the Internet. Nonetheless, internal resources of cultural producers can be reallocated to more value adding functions then, such as bundling or promoting the musical material and the crowd filter can be applied as on demand service.

However, the crowd can also be helpful in the bundling and development process by providing instant feedback about the quality and marketability of a certain song or artist. Hence, the creative outputs from the managerial subsystem may be tailored to the needs of the consumers. Consequently, the support subsystem offers the potential to acquire cost-efficient market intelligence. This could even be done by implicitly analyzing the listening habits of the crowd workers instead of actively asking for the rating. In addition, the crowd can be helpful in the same way for the institutional subsystem by delivering more accurate market intelligence, i.e. for scheduling and programming the playlists of the radio stations.

Moreover, in contrast to usual push marketing strategy of the managerial subsystem, there could be also a crowd-induced pull strategy in order to weaken the decision power of the media gatekeepers and to increase the certainty that musical productions actually reach the consumers and enter the charts. Establishing a loyal and committed crowd may enable this pulling force.

Hence, there is a wide area of application for the support subsystem along the value chain of the recording industry. However, the recording companies should not lose the pace, because, similar to the digitalization of music, there are new services and organizations rushing into the industry, such as Musicxray, Slicethepie, or Reverbnation. Furthermore, the support subsystem could also take over a different crowdsourcing model, such as a crowdfunding platform where the crowd decides about the support of the artist. For example, Amanda Palmer[1] has successfully crowd-funded the production of her album and recently scored a chart success. This is inter alia due to the increased interaction between Amanda Palmer and her fan base. Thus again, this is a proof, that advocacy in the form of recommendation or re-experience is very important for the success of an artist.

6. Limitations and future research

Apart from the implications, there are also some limitations to our study. First, the necessity to restrict the analysis to a certain music market requires the adaptation of the research design to the underlying peculiarities and mechanisms in this respective market. In our cases this led to the use of a chart ranking as dependent variable that does not include some dimensions of music success in a digitized world (e.g., song streaming). Second, the crowd forecasts should become even more accurate with increasing sample size, which is only accomplished through higher investments of cost and time. Even though we found the means to be already pretty stable with around 20 crowd evaluations per song, this should be further investigated in future research because it largely impacts the cost associated with the suggested crowd evaluation. Furthermore, aligning the origin of the crowd workers with the country where chart success is to be forecasted should further improve the accuracy of the results. Future research could look into this by investigating the influence of regional boundaries for crowd-based forecasting. Third, even though the sample is more diverse than traditional convenience samples, the focus on Internet users and digital natives still limits the general representativeness. However, to ensure the influence of our crowd sample's music product buying behavior on chart success, we also checked for the percentage of buyers of music products within our sample, which is a bit larger in our sample compared to the Internet population. Fourth, since we have faced the same uncertainty about the future chart performance of the sample songs as A&R managers in the managerial subsystem, we had to follow a lenient sampling strategy for the song selection, which resulted in in a dataset containing more songs from known than from unknown artists (i.e., newcomers), however, our approach seems to work best for newcomers. Fifth, for reasons of data availability, the point in time for the data collection was chosen to be the day of the release instead of a point in time prior to the release.

Relating to these limitations of our study, there are several further directions for future research. First, cooperating with one of the major labels could enable access to songs prior to the release date. This would allow replicating our study with unreleased songs instead of songs from the release day. Second, the validity of our findings should also be tested with further datasets (e.g., only novel songs from unknown artists). Third, the applicability of the findings could be tested in other cultural industries as well. An example of the German movie industry already indicates this potential. The platform “flimmer.de” pays consumers for previewing trailers of upcoming movies. Fourth, the findings from the study may be compared with other music markets in order to find general patterns of crowdsourced success forecasting or to prove our assumption that the predictive relevance of the crowd will be larger in already digitally transformed music markets. Fifth, we could also imagine a prediction model based on the chart lifecycle analysis by the crowd. Finally, a more complex hierarchical prediction model would also yield insights into the motives and determinants of the crowd evaluations.

7. Conclusion

We have conducted this study in order to find out, whether the disruptive potential of crowdsourcing can be used for the success forecasting of popular music in the German recording industry. Therefore, we aimed at finding empirical evidence for the causal relationship between aggregated crowd song evaluations and the peak chart position. We have used two constructs from the hedonic consumption model to reflect the crowd forecast. Together with control variables from the literature related to the music industry, such as label affiliation or prior chart history, we have applied PLS to build, test and estimate our model, whereas we have used Amazon Mechanical Turk to build the song evaluation tasks. Our results show that we are able to explain large parts of chart success through our model ($R^2 = 0.435$). We find evidence for a reliable and valid measurement model and argue that our structural model generally has explanatory power and predictive relevance. However, digging deeper into the details, we find that this forecasting power seems to be the highest for novel songs from unknown artists, which might help the music industry in approaching the very challenging task to innovate with newcomers.

Overall, we have found that the inclusion of our crowd forecasting constructs increased the explained variance by approximately 30% relative to the variance that is explained by the established controls only. Thus, opening up the value chain of the recording industry may result in more complete and accurate market information about the potential of musical material. We recommend integrating a support subsystem in the value chain of the recording industry in order to reduce the uncertainty by filtering the vast amount of creative input.

Appendix A

Table 3
Operationalization of demographics and further controls.

	Item	Operationalization	Source
Age			
AGE1	18–24	Drop-Down Categories	Ipeirotis (2010a)
AGE2	25–34	Drop-Down Categories	Ipeirotis (2010a)
AGE3	35–44	Drop-Down Categories	Ipeirotis (2010a)
AGE4	45–54	Drop-Down Categories	Ipeirotis (2010a)
AGE5	55–64	Drop-Down Categories	Ipeirotis (2010a)
AGE6	65 +	Drop-Down Categories	Ipeirotis (2010a)
Gender			
GEN	Male vs. female	1 Male|0 otherwise	www.surveygizmo.com

(*continued on next page*)

[1] See: www.amandapalmer.net

Table 3 (*continued*)

	Item	Operationalization	Source
Country of origin			
COO	Pre-made country list	Drop-Down Categories	www.surveygizmo.com
Income			
INC1	< $10,000	Drop-Down Categories	Ipeirotis (2010a)
INC2	$10,000–$20,000	Drop-Down Categories	Ipeirotis (2010a)
INC3	$20,000–$30,000	Drop-Down Categories	Ipeirotis (2010a)
INC4	$30,000–$40,000	Drop-Down Categories	Ipeirotis (2010a)
INC5	$40,000–$60,000	Drop-Down Categories	Ipeirotis (2010a)
INC6	> $60,000	Drop-Down Categories	Ipeirotis (2010a)
Mechanical Turk activity			
MTA1	< 1 hpw	Drop-Down Categories	Ipeirotis (2010a)
MTA2	1–5 hpw	Drop-Down Categories	Ipeirotis (2010a)
MTA3	5–15 hpw	Drop-Down Categories	Ipeirotis (2010a)
MTA4	15–30 hpw	Drop-Down Categories	Ipeirotis (2010a)
MTA5	> 30 hpw	Drop-Down Categories	Ipeirotis (2010a)
Motivation			
MOT1	The task is fun	7-Point Likert (agr.-disagr.)	Ipeirotis (2010a)
MOT2	It is a fruitful way to spend time	7-Point Likert (agr.-disagr.)	Ipeirotis (2010a)
MOT3	For secondary income purposes	7-Point Likert (agr.-disagr.)	Ipeirotis (2010a)
MOT4	This is my primary source of income	7-Point Likert (agr.-disagr.)	Ipeirotis (2010a)
MOT5	I could use some extra money	7-Point Likert (agr.-disagr.)	Adopted from Kaufmann et al. (2011)
MOT6	It is a meaningful way of doing work	7-Point Likert (agr.-disagr.)	Adopted from Kaufmann et al. (2011)
Community board activity / independence			
CBA	How frequent do you interact on Community boards concerned with Amazon Mechanical Turk? (i.e. Turkernation.com)	7-Point Likert (never-every time)	Own
Musical affinity			
MAF	Music is important to me	7-Point Likert (agr.-disagr.)	North and Oishi (2006); Rentfrow and Gosling (2003)
Formal music training			
FMT1	None	Drop-Down Categories	Adopted from Prince (1972)
FMT2	1–5 years	Drop-Down Categories	Adopted from Prince (1972)
FMT3	6–10 years	Drop-Down Categories	Adopted from Prince (1972)
FMT4	11–15 years	Drop-Down Categories	Adopted from Prince (1972)
FMT5	16–20 years	Drop-Down Categories	Adopted from Prince (1972)
FMT6	> 20 years	Drop-Down Categories	Adopted from Prince (1972)
Musical employment			
MEM1	None	Drop-Down Categories	Adopted from Prince (1972)
MEM2	1–5 years	Drop-Down Categories	Adopted from Prince (1972)
MEM3	6–10 years	Drop-Down Categories	Adopted from Prince (1972)
MEM4	11–15 years	Drop-Down Categories	Adopted from Prince (1972)
MEM5	16–20 years	Drop-Down Categories	Adopted from Prince (1972)
MEM6	> 20 years	Drop-Down Categories	Adopted from Prince (1972)
Music preferences/song genre			
MPREF1	Pop	7-Point Likert (like-dislike)	Rentfrow and Gosling (2003)
MPREF2	Rock	7-Point Likert (like-dislike)	Rentfrow and Gosling (2003)
MPREF3	Alternative	7-Point Likert (like-dislike)	Rentfrow and Gosling (2003)
MPREF4	Dance/electronic	7-Point Likert (like-dislike)	Rentfrow and Gosling (2003)
MPREF5	Rap/hip-hop	7-Point Likert (like-dislike)	Rentfrow and Gosling (2003)
Purchase preferences			
PPREF1	1–3	Drop-Down Categories	Preiser and Vogel (2002)
PPREF2	4–9	Drop-Down Categories	Preiser and Vogel (2002)
PPREF3	10–19	Drop-Down Categories	Preiser and Vogel (2002)
PPREF4	20 +	Drop-Down Categories	Preiser and Vogel (2002)
Consumption preferences			
CPREF1	Digital vs physical	7-Point Likert (physical-digital)	Own
CPREF2	Accessing vs. owning	7-Point Likert (owning-accessing)	Own

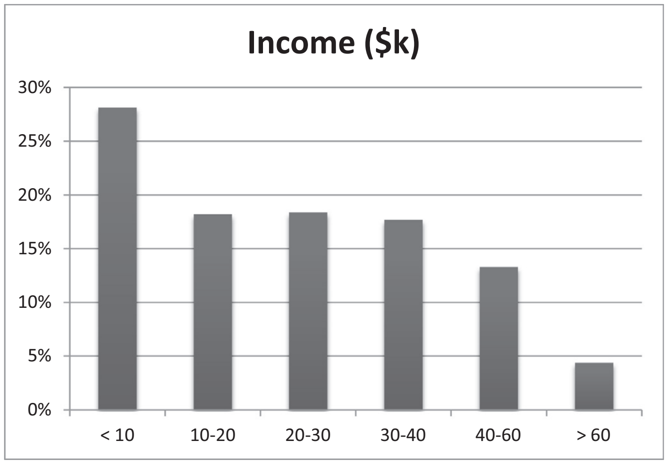

Fig. 3. Income distribution of the crowd workers.

References

Achterberg, P., Heilbron, J., Houtman, D., Aupers, S., 2011. A cultural globalization of popular music? American, Dutch, French, and German popular music charts (1965 to 2006). Am. Behav. Sci. 55, 589–608. https://doi.org/10.1177/0002764211398081.

Agarwal, J., Malhotra, N.K., 2005. An integrated model of attitude and affect. J. Bus. Res. 58, 483–493. https://doi.org/10.1016/S0148-2963(03)00138-3.

Aguiar, L., Waldfogel, J., 2015. Streaming reaches flood stage: does Spotify stimulate or depress music sales? (NBER working paper no. 21653). National Bureau of Economic Research. https://doi.org/10.3386/w21653.

Ahlkvist, J.A., Faulkner, R., 2002. “Will this record work for us?”: managing music formats in commercial radio. Qual. Sociol. 25, 189–215.

Ajzen, I., 1991. The theory of planned behavior. Organ. Behav. Hum. Decis. Process. 50, 179–211.

Alba, J.W., Williams, E.F., 2013. Pleasure principles: a review of research on hedonic consumption. J. Consum. Psychol. 23, 2–18. https://doi.org/10.1016/j.jcps.2012.07.003.

Anand, N., Peterson, R.A., 2000. When market information constitutes fields: sensemaking of markets in the commercial music industry. Organ. Sci. 11, 270–284. https://doi.org/10.1287/orsc.11.3.270.12502.

Anderson, C., 2006. The Long Tail: Why the Future of Business is Selling Less of More. Hyperion, New York, NY, USA.

Asai, S., 2008. Factors affecting hits in Japanese popular music. Journal of Media Economics 21, 97–113. https://doi.org/10.1080/08997760802069895.

Bagozzi, R.P., Yi, Y., 1988. On the evaluation of structural equation models. J. Acad. Mark. Sci. 16, 74–94. https://doi.org/10.1007/BF02723327.

Benner, M.J., Waldfogel, J., 2016. The song remains the same? Technological change and positioning in the recorded music industry. Strategy Science 1, 129–147. https://doi.org/10.1287/stsc.2016.0012.

Berg, J.E., Neumann, G.R., Rietz, T.A., 2009. Searching for Google's value: using prediction markets to forecast market capitalization prior to an initial public offering. Manag. Sci. 55, 348–361.

Berinsky, A.J., Huber, G.A., Lenz, G.S., 2012. Evaluating online labor markets for experimental research: Amazon.com's mechanical Turk. Polit. Anal. 20, 351–368. https://doi.org/10.1093/pan/mpr057.

Bhattacharjee, S., Gopal, R.D., Lertwachara, K., Marsden, J.R., Telang, R., 2007. The effect of digital sharing technologies on music markets: a survival analysis of albums on ranking charts. Manag. Sci. 53, 1359–1374. https://doi.org/10.1287/mnsc.1070.0699.

Bollen, J., Mao, H., Zeng, X., 2011. Twitter mood predicts the stock market. Journal of Computational Science 2, 1–8. https://doi.org/10.1016/j.jocs.2010.12.007.

Bonabeau, E., 2009. Decisions 2.0: the power of collective intelligence. MIT Sloan Manag. Rev. 50, 45–52.

Briscoe, E., Appling, S., Schlosser, J., 2016. Technology futures from passive crowdsourcing. IEEE Transactions on Computational Social Systems 3, 23–31. https://doi.org/10.1109/TCSS.2016.2563460.

Budescu, D.V., Chen, E., 2014. Identifying expertise to extract the wisdom of crowds. Manag. Sci. 61, 267–280. https://doi.org/10.1287/mnsc.2014.1909.

Buhrmester, M., Kwang, T., Gosling, S.D., 2011. Amazon's mechanical Turk a new source of inexpensive, yet high-quality, data? Perspect. Psychol. Sci. 6, 3–5. https://doi.org/10.1177/1745691610393980.

BVMI, 2017. Music Industry Yearbook Germany 2016. Bundesverband Musikindustrie e.V, Hamburg.

Cameron, S., 2016. Past, present and future: music economics at the crossroads. J. Cult. Econ. 40, 1–12. https://doi.org/10.1007/s10824-015-9263-4.

Celma, Ò., Serra, X., 2008. FOAFing the music: bridging the semantic gap in music recommendation. Web Semant. Sci. Serv. Agents World Wide Web 6, 250–256. https://doi.org/10.1016/j.websem.2008.09.004.

Cepeda Carrión, G., Henseler, J., Ringle, C.M., Roldán, J.L., 2016. Prediction-oriented modeling in business research by means of PLS path modeling: introduction to a JBR special section. J. Bus. Res. 69, 4545–4551. https://doi.org/10.1016/j.jbusres.2016.03.048.

Cronbach, L.J., 1951. Coefficient alpha and the internal structure of tests. Psychometrika 16, 297–334. https://doi.org/10.1007/BF02310555.

Dainow, E., 1977. Physical effects and motor responses to music. J. Res. Music. Educ. 25, 211–221. https://doi.org/10.2307/3345305.

Decker, C., Welpe, I.M., Ankenbrand, B.H., 2011. How to motivate people to put their money where their mouth is: what makes employees participate in electronic prediction markets? Technol. Forecast. Soc. Chang. 78, 1002–1015. https://doi.org/10.1016/j.techfore.2011.01.010.

Dewan, S., Ramaprasad, J., 2014. Social media, traditional media, and music sales. MIS Q. 38, 101–121.

Dilmperi, A., King, T., Dennis, C., 2011. Pirates of the web: the curse of illegal downloading. J. Retail. Consum. Serv. 18, 132–140. https://doi.org/10.1016/j.jretconser.2010.12.004.

DiMaggio, P., 1977. Market structure, the creative process, and popular culture: toward an organizational reinterpretation of mass-culture theory. J. Pop. Cult. 11, 436–452. https://doi.org/10.1111/j.0022-3840.1977.00436.x.

Dobusch, L., Schüßler, E., 2014. Copyright reform and business model innovation: regulatory propaganda at German music industry conferences. Technol. Forecast. Soc. Chang. 83, 24–39. https://doi.org/10.1016/j.techfore.2013.01.009.

Flostrand, A., 2017. Finding the future: crowdsourcing versus the Delphi technique. Business Horizons 60, 229–236. https://doi.org/10.1016/j.bushor.2016.11.007.

Forlines, C., Miller, S., Guelcher, L., Bruzzi, R., 2014. Crowdsourcing the future: predictions made with a social network. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '14. ACM, New York, NY, USA, pp. 3655–3664. https://doi.org/10.1145/2556288.2556967.

Fornell, C., Bookstein, F.L., 1982. Two structural equation models: LISREL and PLS applied to consumer exit-voice theory. J. Mark. Res. 19, 440–452.

Fornell, C., Larcker, D., 1981. Evaluating structural equation models with unobservable variables and measurement error. J. Mark. Res. 18, 39–50.

Förster, B., von der Gracht, H., 2014. Assessing Delphi panel composition for strategic foresight — a comparison of panels based on company-internal and external participants. Technol. Forecast. Soc. Chang. 84, 215–229. https://doi.org/10.1016/j.techfore.2013.07.012.

Fox, M., Wrenn, B., 2001. A broadcasting model for the music industry. Int. J. Media Manag. 3, 112–119. https://doi.org/10.1080/14241270109389954.

Franch, F., 2013. Wisdom of the crowds: 2010 UK election prediction with social media. Journal of Information Technology & Politics 10, 57–71. https://doi.org/10.1080/19331681.2012.705080.

Frith, S., 1996. The A&R Men. In: Gillett, C. (Ed.), The Beat Goes on: The Rock File Reader. Pluto Press, East Haven, CT, USA.

Garcia Martinez, M., 2017. Inspiring crowdsourcing communities to create novel solutions: competition design and the mediating role of trust. Technol. Forecast. Soc. Chang. 117, 296–304. https://doi.org/10.1016/j.techfore.2016.11.015.

Gazley, A., Clark, G., Sinha, A., 2011. Understanding preferences for motion pictures. J. Bus. Res. 64, 854–861. https://doi.org/10.1016/j.jbusres.2010.09.012.

Goel, S., Hofman, J.M., Lahaie, S., Pennock, D.M., Watts, D.J., 2010. Predicting consumer behavior with web search. Proc. Natl. Acad. Sci. 107, 17486–17490. https://doi.org/10.1073/pnas.1005962107.

Gottschlich, J., Hinz, O., 2014. A decision support system for stock investment recommendations using collective wisdom. Decis. Support. Syst. 59, 52–62. https://doi.org/10.1016/j.dss.2013.10.005.

Halman, A., 2015. Before and beyond anticipatory intelligence: assessing the potential for crowdsourcing and intelligence studies. Journal of Strategic Security 8. https://doi.org/10.5038/1944-0472.8.3S.1468.

Hargreaves, D.J., North, A.C., 1997. The Social Psychology of Music. (New York, NY, USA).

Hayes, B.E., 2008. Measuring Customer Satisfaction and Loyalty: Survey Design, Use, and Statistical Analysis Methods. ASQ Quality Press, Milwaukee, WI, USA.
Hiller, R.S., 2016. Sales displacement and streaming music: evidence from Youtube. Inf. Econ. Policy 34, 16–26. https://doi.org/10.1016/j.infoecopol.2015.12.002.
Hirsch, P.M., 1972. Processing fads and fashions: an organization-set analysis of cultural industry systems. Am. J. Sociol. 77, 639–659.
Holbrook, M.B., 2006. Audience judgments as the potential missing link between expert judgments and audience appeal: an illustration based on musical recordings of "my funny valentine". J. Acad. Mark. Sci. 34, 8–18. https://doi.org/10.1177/0092070305281627.
Holbrook, M.B., Hirschman, E.C., 1982. The experiential aspects of consumption: consumer fantasies, feelings, and fun. J. Consum. Res. 9, 132–140.
Holbrook, M.B., Huber, J., 1979. Separating perceptual dimensions from affective overtones: an application to consumer aesthetics. J. Consum. Res. 5, 272–283.
Hosseini, M., Moore, J., Almaliki, M., Shahri, A., Phalp, K., Ali, R., 2015. Wisdom of the crowd within enterprises: practices and challenges. Comput. Netw. 90, 121–132. https://doi.org/10.1016/j.comnet.2015.07.004.
Howe, J., 2009. Crowdsourcing: Why the Power of the Crowd Is Driving the Future of Business. Crown Business, New York, NY, USA.
Hoyer, W.D., Brown, S.P., 1990. Effects of brand awareness on choice for a common, repeat-purchase product. J. Consum. Res. 17, 141–148.
Huff, C., Tingley, D., 2015. "Who are these people?" evaluating the demographic characteristics and political preferences of MTurk survey respondents. Research & Politics 2, 2053168015604648. https://doi.org/10.1177/2053168015604648.
IFPI, 2011. The Recording Industry in Numbers (Industry Report). IFPI.
Ipeirotis, P., 2010a. Analyzing the Amazon mechanical Turk marketplace. XRDS: Crossroads 17, 16–21. https://doi.org/10.1145/1869086.1869094.
Ipeirotis, Panagiotis, 2010b. Demographics of Mechanical Turk. NYU Center for Digital Economy Research Report, CeDER-10-01 1–14.
Jia, R., Steelman, Z., Reich, B., 2017. Using mechanical Turk data in IS research: risks, rewards, and recommendations. Commun. Assoc. Inf. Syst. 41.
Jin, X., Gallagher, A., Cao, L., Luo, J., Han, J., 2010. The wisdom of social multimedia: using Flickr for prediction and forecast. In: Proceedings of the 18th ACM International Conference on Multimedia. ACM, pp. 1235–1244.
Kalwani, M.U., Silk, A.J., 1982. On the reliability and predictive validity of purchase intention measures. Mark. Sci. 1, 243–286. https://doi.org/10.1287/mksc.1.3.243.
Kaufmann, N., Schulze, T., Veit, D., 2011. More than fun and money. Worker motivation in crowdsourcing–a study on mechanical Turk. Presented at the Americas Conference on Information Systems (AMCIS), Detroit, MI, USA, pp. 1–11.
Kazai, G., 2010. An Exploration of the Influence That Task Parameters Have on the Performance of Crowds. Presented at the CrowdConf, San Francisco, CA, USA. pp. 1–9.
Keller, J., von der Gracht, H.A., 2014. The influence of information and communication technology (ICT) on future foresight processes — results from a Delphi survey. Technol. Forecast. Soc. Chang. 85, 81–92. https://doi.org/10.1016/j.techfore.2013.07.010.
Lacher, K.T., Mizerski, R., 1994. An exploratory study of the responses and relationships involved in the evaluation of, and in the intention to purchase new rock music. J. Consum. Res. 21, 366–380.
Lam, C.K.M., Tan, B.C.Y., 2001. The internet is changing the music industry. Commun. ACM 44, 62–68.
Lampel, J., Lant, T., Shamsie, J., 2000. Balancing act: learning from organizing practices in cultural industries. Organ. Sci. 11, 263–269. https://doi.org/10.1287/orsc.11.3.263.12503.
Lang, M., Bharadwaj, N., Di Benedetto, C.A., 2016. How crowdsourcing improves prediction of market-oriented outcomes. J. Bus. Res. 69, 4168–4176. https://doi.org/10.1016/j.jbusres.2016.03.020.
Lavidge, R.J., Steiner, G.A., 1961. A model for predictive measurements of advertising effectiveness. J. Mark. 25, 59–62.
Legrand, E., 2012. Monitoring the Cross-border Circulation of European Music Repertoire Within the European Union (Industry Report). European Music Office.
Likert, R., 1932. A technique for the measurement of attitudes. Arch. Psychol. 22, 1–55.
Lorenzen, M., Frederiksen, L., 2005. The management of projects and product experimentation: examples from the music industry. Eur. Manag. Rev. 2, 198–211. https://doi.org/10.1057/palgrave.emr.1500044.
Macdonald, E.K., Sharp, B.M., 2000. Brand awareness effects on consumer decision making for a common, repeat purchase product: a replication. J. Bus. Res. 48, 5–15.
Majchrzak, A., Malhotra, A., 2013. Towards an information systems perspective and research agenda on crowdsourcing for innovation. J. Strateg. Inf. Syst. 22, 257–268. https://doi.org/10.1016/j.jsis.2013.07.004.
Mali, V.V., 2008. An alternative operating model for the record industry based on the development and application of non-traditional financial models. UCLA Ent. L. Rev. 15, 127–138.
Mason, W., Suri, S., 2011. Conducting behavioral research on Amazon's mechanical Turk. Behav. Res. Methods 44, 1–23. https://doi.org/10.3758/s13428-011-0124-6.
de Mattos, C.A., Kissimoto, K.O., Laurindo, F.J.B., 2018. The role of information technology for building virtual environments to integrate crowdsourcing mechanisms into the open innovation process. Technol. Forecast. Soc. Chang. https://doi.org/10.1016/j.techfore.2017.12.020.
Mittal, B., 1988. The role of affective choice mode in the consumer purchase of expressive products. J. Econ. Psychol. 9, 499–524. https://doi.org/10.1016/0167-4870(88)90016-5.
Mizerski, R., Pucely, M.J., Perrewe, P., Baldwin, L., 1988. An experimental evaluation of music involvement measures and their relationship with consumer purchasing behavior. Pop. Music Soc. 12, 79–96. https://doi.org/10.1080/03007768808591325.
Mortara, L., Ford, S.J., Jaeger, M., 2013. Idea competitions under scrutiny: acquisition, intelligence or public relations mechanism? Technol. Forecast. Soc. Chang. 80, 1563–1578. https://doi.org/10.1016/j.techfore.2013.01.008.
Morwitz, V.G., Schmittlein, D., 1992. Using segmentation to improve sales forecasts based on purchase intent: which "intenders" actually buy? J. Mark. Res. 29, 391–405.
Neelamegham, R., Jain, D., 1999. Consumer choice process for experience goods: an econometric model and analysis. J. Mark. Res. 36, 373–386.
Negus, K., 1992. Producing Pop: Culture and Conflict in the Popular Music Industry. E. Arnold; Distributed in the USA by Routledge. Chapman, and Hall, New York, NY, USA.
Ni, Y., Santos-Rodríguez, R., McVicar, M., De Bie, T., 2011. Hit Song Science Once Again a Science? Melia Sierra Nevada & Melia Sol y Nieve, Sierra Nevada, ES, pp. 1–2.
Niforatos, E., Vourvopoulos, A., Langheinrich, M., Campos, P., Doria, A., 2014. Atmos: A hybrid crowdsourcing approach to weather estimation. In: Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication, UbiComp '14 Adjunct. ACM, New York, NY, USA, pp. 135–138. https://doi.org/10.1145/2638728.2638780.
North, A.C., Oishi, A., 2006. Music CD purchase decisions. J. Appl. Soc. Psychol. 36, 3043–3084. https://doi.org/10.1111/j.0021-9029.2006.00142.x.
Okada, E.M., 2005. Justification effects on consumer choice of hedonic and utilitarian goods. J. Mark. Res. 42, 43–53. https://doi.org/10.1509/jmkr.42.1.43.56889.
Oppenheimer, D.M., Meyvis, T., Davidenko, N., 2009. Instructional manipulation checks: detecting satisficing to increase statistical power. J. Exp. Soc. Psychol. 45, 867–872. https://doi.org/10.1016/j.jesp.2009.03.009.
Ordanini, A., 2006. Selection models in the music industry: how a prior independent experience may affect chart success. J. Cult. Econ. 30, 183–200. https://doi.org/10.1007/s10824-006-9013-8.
Ordanini, A., Nunes, J.C., 2016. From fewer blockbusters by more superstars to more blockbusters by fewer superstars: how technological innovation has impacted convergence on the music chart. Int. J. Res. Mark. 33, 297–313. https://doi.org/10.1016/j.ijresmar.2015.07.006. The Entertainment Industry.
Osgood, C.E., Suci, G.J., Tannenbaum, P.H., 1957. The Measurement of Meaning. University of Illinois Press, Urbana, IL, USA.
Palmer, J., Sorda, G., Madlener, R., 2015. Modeling the diffusion of residential photovoltaic systems in Italy: an agent-based simulation. Technol. Forecast. Soc. Chang. 99, 106–131. https://doi.org/10.1016/j.techfore.2015.06.011.
Parker, A.M., Srinivasan, S.V., Lempert, R.J., Berry, S.H., 2015. Evaluating simulation-derived scenarios for effective decision support. Technol. Forecast. Soc. Chang. 91, 64–77. https://doi.org/10.1016/j.techfore.2014.01.010.
Paul, M.J., Dredze, M., Broniatowski, D., 2014. Twitter improves influenza forecasting. PLoS Curr 6. https://doi.org/10.1371/currents.outbreaks.90b9ed0f59bae4ccaa683a39865d9117.
Poetz, M.K., Schreier, M., 2012. The value of crowdsourcing: can users really compete with professionals in generating new product ideas? J. Prod. Innov. Manag. 29, 245–256. https://doi.org/10.1111/j.1540-5885.2011.00893.x.
Pöyry, E., Parvinen, P., Malmivaara, T., 2013. Can we get from liking to buying? Behavioral differences in hedonic and utilitarian Facebook usage. Electronic commerce research and applications. Social Commerce- Part 2 (12), 224–235. https://doi.org/10.1016/j.elerap.2013.01.003.
Preiser, J., Vogel, A., 2002. The Music Industry in the 21st Century (Industry Report). (Screen Digest).
Prince, W.F., 1972. A paradigm for research on music listening. J. Res. Music. Educ. 20, 445–455.
Prpić, J., Shukla, P.P., Kietzmann, J.H., McCarthy, I.P., 2015. How to work a crowd: developing crowd capital through crowdsourcing. Business Horizons 58, 77–85. https://doi.org/10.1016/j.bushor.2014.09.005.
Rayna, T., Striukova, L., 2009. Monometapoly or the economics of the music industry. Prometheus 27, 211–222. https://doi.org/10.1080/08109020903127778.
Rentfrow, P.J., Gosling, S.D., 2003. The do re mi's of everyday life: the structure and personality correlates of music preferences. J. Pers. Soc. Psychol. 84, 1236–1256. https://doi.org/10.1037/0022-3514.84.6.1236.
Ringle, C.M., Sarstedt, M., Straub, D.W., 2012. A critical look at the use of PLS-SEM in MIS quarterly. MIS Q. 36, iii–xiv.
Rumsey, F., Zielinski, S., Kassier, R., Bech, S., 2005. Relationships between experienced listener ratings of multichannel audio quality and Naïve listener preferences. J. Acoust. Soc. Am. 117, 3832–3840. https://doi.org/10.1121/1.1904305.
Schemmann, B., Herrmann, A.M., Chappin, M.M.H., Heimeriks, G.J., 2016. Crowdsourcing ideas: involving ordinary users in the ideation phase of new product development. Res. Policy 45, 1145–1154. https://doi.org/10.1016/j.respol.2016.02.003.
Seifert, M., Hadida, A.L., 2006. Facilitating talent selection decisions in the music industry. Manag. Decis. 44, 790–808. https://doi.org/10.1108/00251740610673341.
Seiter, E., 1986. Stereotypes and the media: a re-evaluation. J. Commun. 36, 14–26. https://doi.org/10.1111/j.1460-2466.1986.tb01420.x.
Sernoe, J., 2005. "Now we're on the top, top of the pops": the performance of "non-mainstream" music on billboard's albums charts, 1981–2001. Pop. Music Soc. 28, 639–662. https://doi.org/10.1080/03007760500142670.
Sheeran, P., 2002. Intention—behavior relations: a conceptual and empirical review. Eur. Rev. Soc. Psychol. 12, 1–36. https://doi.org/10.1080/14792772143000003.
Sloboda, J.A., 1991. Empirical studies of emotional response to music. In: Jones, M.R., Holleran, S. (Eds.), Cognitive Bases of Musical Communication. American Psychological Association, Washington, DC, US, pp. 33–46.
Sprouse, J., 2010. A validation of Amazon mechanical Turk for the collection of acceptability judgments in linguistic theory. Behav. Res. Methods 43, 155–167. https://doi.org/10.3758/s13428-010-0039-7.
Steinkrauß, N., Gmelin, H., Günnel, S., 2008. Wettbewerbsanalyse. In: Clement, M., Schusser, O., Papies, D. (Eds.), Ökonomie Der Musikindustrie. Gabler Verlag,

Wiesbaden, GER, pp. 27–44.
Strobl, E.A., Tucker, C., 2000. The dynamics of chart success in the U.K. pre-recorded popular music industry. J. Cult. Econ. 24, 113–134. https://doi.org/10.1023/A:1007601402245.
Styvén, M., 2007. The intangibility of music in the internet age. Pop. Music Soc. 30, 53–74. https://doi.org/10.1080/03007760500503442.
Surowiecki, J., 2005. The Wisdom of Crowds. New ed. Trade Paperback, Anchor, New York, NY.
Sutton, S., 1998. Predicting and explaining intentions and behavior: how well are we doing? J. Appl. Soc. Psychol. 28, 1317–1338. https://doi.org/10.1111/j.1559-1816.1998.tb01679.x.
Tumasjan, A., Sprenger, T.O., Sandner, P.G., Welpe, I.M., 2011. Election forecasts with twitter: how 140 characters reflect the political landscape. Soc. Sci. Comput. Rev. 29, 402–418. https://doi.org/10.1177/0894439310386557.
Turner, B., Steyvers, M., 2011. A wisdom of the crowd approach to forecasting. In: 2nd NIPS Workshop on Computational Social Science and the Wisdom of Crowds.
Vinzi, V.E., Chin, W.W., Henseler, J., Wang, H., 2010. Handbook of Partial Least Squares Concepts, Methods and Applications. Springer, Berlin; New York.
Voss, K.E., Spangenberg, E.R., Grohmann, B., 2003. Measuring the hedonic and utilitarian dimensions of consumer attitude. J. Mark. Res. 40, 310–320. https://doi.org/10.1509/jmkr.40.3.310.19238.
Wolfers, J., Zitzewitz, E., 2004. Prediction markets. J. Econ. Perspect. 18, 107–126. https://doi.org/10.1257/0895330041371321.
Young, S., Collins, S., 2010. A view from the trenches of music 2.0. Pop. Music Soc. 33, 339–355. https://doi.org/10.1080/03007760903495634.
Zajonc, R.B., 1980. Feeling and thinking: preferences need no inferences. Am. Psychol. 35, 151–175. https://doi.org/10.1037/0003-066X.35.2.151.
Zhao, R., Zhou, X., Han, J., Liu, C., 2016. For the sustainable performance of the carbon reduction labeling policies under an evolutionary game simulation. Technol. Forecast. Soc. Chang. 112, 262–274. https://doi.org/10.1016/j.techfore.2016.03.008.
Zwaan, K., ter Bogt, T.F.M., 2009. Research note: breaking into the popular record industry: an Insider's view on the career entry of pop musicians. Eur. J. Commun. 24, 89–101. https://doi.org/10.1177/0267323108098948.

Dennis M. Steininger is an assistant professor in information systems, digital entrepreneurship, and business modeling at the Faculty of Business and Economics of the University of Augsburg, Germany. Moreover, he is a research fellow and guest lecturer of the Institute for SME and Entrepreneurship Research at the University of Mannheim, Germany. He holds a diploma and PhD in business from the University of Mannheim and has visited the London School of Economics and Political Science for parts of his studies. His research focuses on the digitalization of innovation and business models and the related impacts of crowdsourcing, crowdfunding, and social media. His work has been published in the *Journal of Management Information Systems, Information Systems Journal, Electronic Markets, e-Service Journal*, and others.

Simon Gatzemeier is a senior data analyst and senior consultant and a research fellow of the University of Mannheim. He received his M.Sc. in Business Informatics the University of Mannheim and Singapore Management University. His research has been presented at conferences.