Investigating the Relationship Between Diversity in Music Consumption Behavior and Cultural Dimensions: A Cross-Country Analysis

Bruce Ferwerda
Department of Computational Perception
Johannes Kepler University
Altenberger Str. 69
A-4040 Linz, Austria
bruce.ferwerda@jku.at

Markus Schedl
Department of Computational Perception
Johannes Kepler University
Altenberger Str. 69
A-4040 Linz, Austria
markus.schedl@jku.at

ABSTRACT

Diversity in recommendation lists or sets has shown to be an important feature in recommender systems as it can counteract on negative effects such as choice difficulty and choice overload. However, how much diversity a recommendation list needs to provide is not clearly defined. By analyzing music listening behavior of listeners in 47 countries, we show that diversity needs may be cultural dependent. For our analyses, we exploited a Last.fm dataset containing almost 1.1 billion single listening events. We investigated several diversity measures to identify how users in different countries apply music diversity to their listening behavior. By analyzing 53,309 Last.fm users, we found distinct diversity behavior related to several cultural dimensions of Hofstede. We show with our results that different diversity needs exist between cultures, and should be taken into account when applying diversity to a recommendation list.

CCS Concepts

•Human-centered computing → User models; •Social and professional topics → Cultural characteristics;

Keywords

diversity, cultural differences, music recommender systems

1. INTRODUCTION

By tradition, recommender systems are created to most accurately provide recommendations in line with the user's taste (i.e., output options with the highest predicted ratings). The assumption of this approach is that the higher the recommendation accuracy, the higher the attractiveness of the items for the user. However, it has been shown that by doing this two subsequent effects may occur, which are caused by recommendations that are too attractive: (i) choice difficulties [23], and (ii) choice overload [2]. One way to counteract on the negative effects of too attractive items is to introduce recommendation diversity.

The amount of diversity that a set or list of recommendations should provide has been given limited attention. Prior research has shown that personal characteristics, such as expertise, play a role in the desired amount of recommendation diversity [2]. As the shaping of behavior and preferences has shown to be influenced by culture [13], identifying diversity on a country level may already provide cues about the desired recommendation diversity.

Providing a truly personalized experience to the user is still challenging in today's recommender systems. Often there is simply not enough data available (yet) about the user. A way to solve this problem is to use questionnaires in order to get to know the user. However, this is not desirable since it is obtrusive, takes a lot of effort and time from the user, and thereby disrupts their interaction with the system. Since country information is often available through the user's profile information, identifying diversity needs on a country level could be exploited to provide users with a personalized experience. Quantifying these diversity needs and studying their relationship to cultural dimensions is the focus of the study at hand.

This study is a follow-up investigation of the one presented in [8]. A shortcoming of that study was the quite simple definition of diversity, solely based on absolute numbers of whether an artist is listened to or not, neglecting the frequency the artist is listened to. Furthermore, [8] uses genres taken from the Echonest,[1] which are very broad, thus rendering impossible fine-grained modeling and analysis. Addressing these two shortcomings, the paper at hand (i) investigates several volume- and entropy-based diversity formulations and (ii) models a population's diversity via a dictionary containing more than 2,000 genre names.

2. RELATED WORK

Recommender systems intend to create a personalized set of items that are most relevant to the user. However, highly relevant items often appear too similar to each other, resulting in recommendations that may be perceived as boring to the users. A set of items showing too much similarity (e.g., too many highly relevant items) can, in turn, cause choice overload [19]. Bollen et al. [2] and Willemsen et al. [23] inves-

[1] http://the.echonest.com

24th Conference on User Modeling, Adaptation, and Personalization (UMAP) Extended Proceedings (July 13-17, 2016, Halifax, NS, Canada): 1st Workshop on Surprise, Opposition, and Obstruction in Adaptive and Personalized Systems (SOAP)

tigated the influence of diversity on movie recommendations and found that diversity has a positive effect on the attractiveness of the recommendation set, the difficulty to make a choice, and eventually on the choice satisfaction. Bollen et al. [2] additionally identified individual differences. For example, they showed that increased expertise has positive effects on perceived item variety and attractiveness.

Besides individual user characteristics, research has shown that cultural aspects can provide useful cues too. General behavior and preferences have shown to be rooted and embodied in culture [13], hence looking at behavior on a country level may provide useful information for the desired recommendation diversity. In a comprehensive study, Hofstede et al. [11] describe national cultures among six dimensions: power distance, individualism, masculinity, uncertainty avoidance, long-term orientation, and indulgence. These dimensions describe the effects of a society's culture on the values and behaviors of its members, which we use to explain diversity differences between countries in this study. The data that describes Hofstede's cultural dimensions has been collected since 1967 and is being refined ever since.

3. METHOD

The study at hand investigates on a standardized corpus of music listening events and listener demographics on the one hand, and Hofstede's cultural dimensions on the other, the relationship between diversity of music tastes (on the genre level) and cultural aspects in a population (on the country level). In the following, we first provide details of the used dataset, its enrichment by genre terms, and the definitions for diversity that we study. Subsequently, we summarize Hofstede's cultural dimensions, against which we compare our diversity scores via correlation analysis.

3.1 Music Dataset

We used the LFM-1b dataset [18] to model diversity and perform our experiments.[2] It is a dataset containing almost 1.1 billion single listening events by more than 120,000 users of Last.fm[3] and covers over 3 million unique artists. Due to its user-generated nature, however, the data is quite noisy, e.g., metadata frequently contains typos. We therefore had to perform simple data cleaning first. Assuming that wrong artist names, which can be the result of misspellings, typos, hacking, and vandalism, etc., do not frequently occur in the dataset, we discarded all artist names that occur in the listening events of less than 10 users. This cleaning resulted in a dataset of 585,095 artists.

3.2 Modeling Diversity

Geographically, we model diversity on the country level. In order to obtain relevant results, we only consider countries with at least 100 users in the LFM-1b dataset. For detailed numbers of users, please consider Table 2.

In terms of listening behavior, we gauge diversity via scores derived from genre data of users' listening events. To this end, we first retrieve all artists' top tags provided by Last.fm via their API.[4] The resulting tags obviously contain many terms other than genre names, for which reason we index them by a dictionary of 1,998 genre and style terms extracted from Freebase.[5] We further restrict the considered tags to those that have a tag weight of at least 10 according to Last.fm's weighting scheme.[6] This eventually provides us with a set of genre tags for each artist. Statistics of the 50 most frequently occurring genres for selected countries of the dataset are provided in Table 1 for the U.S.A., Japan, and Finland. As can be seen, while there are quite a few genres that are popular among Last.fm users in all three countries (e.g., Rock, Alternative, and Metal), country-specific differences are evidenced too. For instance, J-Pop is a genre very popular in Japan, but not among the top 50 genres in any of the other countries listed here. In contrast, 3 out of the top 10 most popular genres in Finland relate to Metal.

Based on the users' demographics, as provided in the LFM-1b dataset, and the artist-related genre information, obtained as described above, we define the following volume- and entropy-based diversity measures, computed per country and reported in Table 2:

Overall volume of genre occurrences. We count the number of genre tags that appear at least once in at least one user's listening history of the respective country's user base and define it as the absolute volume of genre occurrences (indicated as *Vol. abs.* in Table 2). The relative volume is computed as the fraction of the absolute one and the number of genre tags in the dictionary (*Vol. rel.* in Table 2).

Relative listening volume exceeding one per mille. We first compute the total playcounts, i.e. number of listening events of each artist over all users in the country under investigation. Based on the artist–genre mapping, we subsequently calculate these playcounts per genre by aggregating the playcounts of all artists that are tagged by that genre. This absolute genre playcount is then normalized by the total playcount of a country, yielding an estimate of genre g's relative popularity in country c. Formally, the computation of this relative popularity $pop_c(g)$ is given in Equation 1, where G is the set of genres, U_c is the set of users in country c, A_c^g is the set of artists listened to in country c and tagged as genre g, and $pc(u,a)$ denotes the number of listening events (playcounts) of user u to artist a.[7]

$$pop_c(g) = \frac{\sum_{u \in U_c} \sum_{a \in A_c^g} pc(u,a)}{\sum_{g \in G} \sum_{u \in U_c} \sum_{a \in A_c^g} pc(u,a)} \quad (1)$$

To define diversity, we finally count the number of genres whose relative popularity exceeds one per mille of the total listening events. Again, we use this score as absolute measure and we divide it by the number of genre tags to yield a relative estimate (*Vol.* > 1‰ *abs.* and *Vol.* > 1‰ *rel.* in Table 2).

Entropy. Based on the genre-specific playcounts, computed as described in the previous paragraph, we use the nor-

[2] http://www.cp.jku.at/datasets/LFM-1b

[3] http://www.last.fm

[4] We use the API endpoint http://www.last.fm/api/show/artist.getTopTags.

[5] http://www.freebase.org

[6] Last.fm employs a non-disclosed approach to weight artist tags based on the number of users who assign the tag to the artist. While details are not provided, these weights are normalized to [0,100]. Our filtering thus discards tags infrequently used to describe the artist under consideration.

[7] We use the term le instead of pc in the formula to avoid confusions with p_c in Equation 2.

Table 1: Relative amount of listening events (playcounts PC in percent) of the 50 most frequent genres and styles for the U.S.A., Japan, and Finland.

U.S.A.		Japan		Finland	
Genre tag	PC	Genre tag	PC	Genre tag	PC
Rock	12.51	Rock	16.01	Rock	11.31
Alternative	9.63	Alternative	8.37	Metal	11.15
Alternative rock	5.86	J-pop	5.77	Alternative	7.30
Metal	4.77	Pop	4.56	Alternative rock	4.56
Pop	3.62	Metal	4.55	Hard rock	4.28
Indie	3.59	Alternative rock	4.26	Heavy metal	3.44
Hard rock	3.12	Indie	3.63	Death metal	2.74
Indie rock	3.09	Electronic	2.29	Classic rock	2.61
Classic rock	2.92	Hard rock	2.24	Pop	2.21
Electronic	2.33	Classic rock	2.23	Indie	2.13
Dance	2.21	Visual Kei	2.03	Electronic	2.00
Psychedelic	1.84	Indie rock	2.02	Indie rock	1.75
Blues	1.77	Heavy metal	1.68	Dance	1.71
Hip-Hop	1.72	Dance	1.66	Progressive rock	1.67
Punk	1.61	Punk	1.53	Nu metal	1.57
Heavy metal	1.49	Psychedelic	1.45	Progressive	1.50
Singer-songwriter	1.34	Anime	1.43	Power metal	1.46
Progressive	1.25	Electronica	1.43	Punk	1.45
Electronica	1.24	Blues	1.18	Alternative metal	1.32
Progressive rock	1.16	Japanese rock	1.17	Psychedelic	1.18
New Wave	1.08	Progressive rock	1.06	Hip-Hop	1.10
Punk rock	1.03	Pop punk	0.91	Electronica	0.90
Nu metal	0.99	Nu metal	0.86	Speed metal	0.89
Alternative metal	0.85	Progressive	0.86	Blues	0.84
Rap	0.83	New Wave	0.84	Punk rock	0.82
Post-punk	0.79	Punk rock	0.83	Viking metal	0.75
Synthpop	0.77	Singer-songwriter	0.75	Progressive metal	0.71
Pop punk	0.77	Death metal	0.75	New Wave	0.70
Rnb	0.75	Synthpop	0.67	Melodic death metal	0.69
Psychedelic rock	0.72	Hip-Hop	0.63	Thrash	0.68
Emo	0.68	Experimental	0.59	Visual Kei	0.66
Experimental	0.68	Jazz	0.59	Groove metal	0.65
Death metal	0.68	Ambient	0.59	Pop punk	0.64
Electro	0.67	Power metal	0.58	Psychedelic rock	0.64
Garage rock	0.67	Electropop	0.57	Hardcore	0.62
Blues-rock	0.66	Electro	0.52	Thrash metal	0.62
House	0.64	Alternative metal	0.52	Industrial	0.60
Techno	0.62	Post-punk	0.51	Singer-songwriter	0.59
Ambient	0.60	Speed metal	0.50	Ambient	0.58
Glam rock	0.57	Pop rock	0.47	Experimental	0.53
Folk	0.53	Instrumental	0.47	Synthpop	0.50
Indie pop	0.52	Emo	0.46	Glam rock	0.49
Art rock	0.41	House	0.44	Emo	0.49
Hardcore	0.41	Blues-rock	0.43	Symphonic metal	0.48
Funk	0.40	Funk	0.42	Metalcore	0.46
Instrumental	0.40	Glam rock	0.41	Instrumental	0.46
Speed metal	0.39	Techno	0.39	Electro	0.44
Soul	0.37	Hardcore	0.38	Technical death metal	0.44
Folk rock	0.37	Fusion	0.38	Rapcore	0.43
Industrial rock	0.36	Soul	0.38	Blues-rock	0.42

malized genre entropy as diversity measure. Formally, our adapted entropy measure is defined in Equation 2, where G is the set of all genres and $p_c(g)$ is the probability for genre g in country c. We approximate this probability as the relative frequency of genre g's playcounts among all playcounts in country c. The normalization term in the denominator ensures that the resulting diversity scores fall into the range [0,1].

$$H_c(G) = \frac{-\sum_{g \in G} p_c(g) \cdot \log_2 p_c(g)}{\log_2 |G|} \quad (2)$$

Statistics over relative genre playcounts. In addition to the volume-based diversity measures and to entropy, we investigate basic statistics of the relative genre playcounts. In particular, we compute mean and standard deviation of the elements $p_c(g)$ with $g \in G$. The corresponding scores are denoted *Vol.* μ and *Vol.* σ, respectively, in Table 2.

3.3 Modeling Cultural Dimensions

The most comprehensive framework for national cultures is considered to be Hofstede et al.'s cultural dimensions. They defined six dimensions to identify cultures [11]:

Power distance index (PDI). Defines the extent to which power is distributed unequally by less powerful members of institutions (e.g., family). High power distance indicates that a hierarchy is clearly established and executed in society. Low power distance indicates that authority is questioned and attempted to distribute power equally.

Individualism (IDV). Defines the degree of integration of people into societal groups. High individualism is defined by loose social ties. The main emphasis is on the "I" instead of the "we," while opposite for low individualistic cultures.

Masculinity (MAS). Defines a society's preference for achievement, heroism, assertiveness and material rewards for success (countries scoring high in this dimension). Whereas low masculinity represents a preference for cooperation, modesty, caring for the weak and quality of life.

Uncertainty avoidance index (UAI). Defines a society's tolerance for ambiguity. High scoring countries in this scale are more inclined to opt for stiff codes of behavior, guidelines, laws. Whereas more acceptance of differing thoughts and/or ideas are accepting for those scoring low in this dimension.

Long-term orientation (LTO). Is associated with the connection of the past with the current and future actions and/or challenges. Lower scoring countries tend to believe that traditions are honored and kept, and value steadfastness. High scoring countries believe more that adaptation and circumstantial, pragmatic problem-solving are necessary.

Indulgence (IND). Defines in general the happiness of a country. Countries scoring high in this dimension are related to a society that allows relatively free gratification of basic and natural human desires related to enjoying life and having fun (e.g., be in control of their own life and emotions). Whereas low scoring countries show more controlled gratification of needs and regulate it by means of strict social norms

4. EXPERIMENTS AND RESULTS

A correlation analysis was performed to indicate the relationship between Hofstede's cultural dimensions and the diversity measurements, cf. Table 3. Spearman correlation ($r \in$ [-1,1]) is reported as the correlation coefficient to indicate the strength of the relationship. Statistically significant results at a level of $p < 0.05$ and $p < 0.01$ are denoted by * and **, respectively.

Investigating the table, we find moderate, highly significant correlations between the cultural dimension of *individualism* and the volume-based diversity measures. This correlation is positive for absolute volume (*Vol. abs.*) and negative for the mean and standard deviation of the volume measures that take actual playcount values into account (respectively, *Vol.* μ and *Vol.* σ). It seems reasonable that listeners from cultures in which individualism is important

Table 2: Diversity scores for the top 47 countries. The columns indicate: country, total number of users, Hofstede's cultural dimensions ($r \in [0,100]$): power distance index(PDI), individualism (IDV), masculinity (MAS), uncertainty avoidance index (UAI), long-term orientation (LTO), and indulgence (IND), absolute volume of unique genre occurrences (*Vol. abs.*), relative volume of unique genre occurrences (*Vol. rel.*), absolute (*Vol.* $> 1‰$ *abs.*) and relative (*Vol.* $> 1‰$ *rel.*) volume of genre occurrences with relative listening volumes exceeding one per mille, genre entropy, mean (*Vol.* μ) and standard deviation (*Vol.* σ) of listening distributions over genres.

Country	# User	PDI	IDV	MAS	UAI	LTO	IND	Vol. abs.	Vol.rel. (%)	Vol. > 1‰ abs.	Vol. > 1‰ rel. (%)	Entropy	Vol. μ	Vol. σ
U.S.A.	10255	40	91	62	46	26	68	1111	55.55	132	6.6	0.647383	0.0009	.004580
Russia	5024	93	39	36	95	81	20	1097	54.85	141	7.05	0.665395	0.000912	.004312
Germany	4578	35	67	66	65	83	40	1100	55	138	6.9	0.662084	0.000909	.004346
Great Britain	4534	35	89	66	35	51	69	1103	55.15	132	6.6	0.64291	0.000907	.004702
Poland	4408	68	60	64	93	38	29	1077	53.85	132	6.6	0.647125	0.000929	.004696
Brazil	3886	69	38	49	76	44	59	1053	52.65	119	5.95	0.626137	0.00095	.005271
Finland	1409	33	63	26	59	38	57	1042	52.1	131	6.55	0.656833	0.00096	.004694
Netherlands	1375	38	80	14	53	67	68	1081	54.05	142	7.1	0.658458	0.000925	.004473
Spain	1243	57	51	42	86	48	44	1043	52.15	136	6.8	0.657332	0.000959	.004723
Sweden	1231	31	71	5	29	53	78	1062	53.1	124	6.2	0.649503	0.000942	.004678
Ukraine	1143	N/A	N/A	N/A	N/A	86	14	1029	51.45	139	6.95	0.665125	0.000972	.004543
Canada	1077	39	80	52	48	36	68	1056	52.8	132	6.6	0.65296	0.000947	.004637
France	1055	68	71	43	86	63	48	1045	52.25	140	7	0.66765	0.000957	.004357
Australia	976	38	90	61	51	21	71	1036	51.8	125	6.25	0.64334	0.000965	.004912
Italy	974	50	76	70	75	61	30	1031	51.55	120	6	0.645742	0.00097	.004942
Japan	806	54	46	95	92	88	42	1024	51.2	126	6.3	0.648062	0.000977	.004929
Norway	750	31	69	8	50	35	55	1028	51.4	129	6.45	0.657356	0.000973	.004700
Mexico	705	81	30	69	82	24	97	1011	50.55	137	6.85	0.655207	0.000989	.004930
Czech Republic	632	57	58	57	74	70	29	983	49.15	133	6.65	0.6687	0.001017	.004593
Belarus	558	N/A	N/A	N/A	N/A	81	15	979	48.95	140	7	0.672649	0.001021	.004543
Belgium	513	65	75	54	94	82	57	1008	50.4	142	7.1	0.66945	0.000992	.004547
Indonesia	484	78	14	46	48	62	38	842	42.1	118	5.9	0.644635	0.001188	.005790
Turkey	479	66	37	45	85	46	49	980	49	119	5.95	0.654673	0.00102	.004732
Chile	425	63	23	28	86	31	68	918	45.9	127	6.35	0.653122	0.001089	.005312
Croatia	372	73	33	40	80	58	33	940	47	129	6.45	0.665861	0.001064	.004904
Portugal	291	63	27	31	104	28	33	918	45.9	136	6.8	0.664801	0.001089	.005023
Argentina	282	49	46	56	86	20	62	927	46.35	119	5.95	0.639404	0.001079	.005586
Switzerland	277	34	68	70	58	74	66	970	48.5	132	6.6	0.66451	0.001031	.004768
Austria	276	11	55	79	70	60	63	932	46.6	140	7	0.671787	0.001073	.004804
Denmark	272	18	74	16	23	35	70	950	47.5	136	6.8	0.664297	0.001053	.004858
Hungary	272	46	80	88	82	58	31	901	45.05	137	6.85	0.687505	0.00111	.004544
Serbia	253	86	25	43	92	52	28	910	45.5	141	7.05	0.677889	0.001099	.004746
Romania	237	90	30	42	90	52	20	951	47.55	137	6.85	0.676884	0.001052	.004409
Bulgeria	236	70	30	40	85	69	16	926	46.3	143	7.15	0.681036	0.00108	.004766
Ireland	220	28	70	68	35	24	65	906	45.3	125	6.25	0.652082	0.001104	.005270
Lithuania	202	42	60	19	65	82	16	892	44.6	138	6.9	0.672913	0.001121	.004969
Slovakia	192	104	52	110	51	77	28	878	43.9	136	6.8	0.684491	0.001139	.004614
Greece	175	60	35	57	112	45	50	907	45.35	134	6.7	0.688293	0.001103	.004447
Latvia	165	44	70	9	63	69	13	904	45.2	134	6.7	0.675491	0.001106	.004787
New Zealand	164	22	79	58	49	33	75	865	43.25	134	6.7	0.672034	0.001156	.005161
China	162	80	20	66	30	87	24	847	42.35	129	6.45	0.671203	0.001181	.004991
Columbia	159	67	13	64	80	13	83	885	44.25	123	6.15	0.65439	0.00113	.005477
Iran	135	58	41	43	59	14	40	782	39.1	117	5.85	0.65695	0.001279	.005473
India	122	77	48	56	40	51	26	794	39.7	127	6.35	0.665461	0.001259	.005578
Venezuela	118	81	12	73	76	16	100	816	40.8	123	6.15	0.654646	0.001225	.005830
Estonia	107	40	60	30	60	82	16	823	41.15	125	6.25	0.672622	0.001215	.005148
Israel	100	13	54	47	81	38	N/A	830	41.5	133	6.65	0.674123	0.001205	.005378

show higher diversity in terms of numbers of distinct genres they listen to. The negative correlation to the playcount-based volume measures signifies that these listeners do not listen only to a few genres very intensely (which would result in a higher *Vol.* μ and *Vol.* σ value), but instead spread their music listening time slightly more evenly over various genres (overall, resulting in lower *Vol.* μ and *Vol.* σ scores). Interestingly, the volume measure that restricts results by the

Table 3: Spearman correlations between Hofstede's cultural dimensions and the analyzed diversity measures. Abbreviations for diversity measures as in Table 1. Results for absolute and relative diversity measures are obviously the same and therefore reported only once. Note: *p<.05, **p<.01

	Vol. abs./rel.	Vol. > 1‰ abs./rel.	Entropy	Vol. μ	Vol. σ
Power Distance	-.183	.036	.132	.183	.022
Individualism	.459**	-.167	-.117	-.459**	-.414**
Masculinity	-.073	-.115	-.133	.073	.088
Uncertainty Avoidance	.057	.301*	.218	-.057	-.174
Long-Term Orientation	.106	.443**	.442**	-.106	-.442**
Indulgence	.217	-.300**	-.558**	-.217	.225

per mille threshold does not show significant correlations to individualism. This is presumably due to the lower number of genres and styles considered in this case that does not account for a high enough amount of individualism.

For *long-term orientation*, we identify moderate positive, highly significant correlations with both volume- and entropy-based diversity measures. This can be explained by the reasonable assumption that cultures scoring high on aspects like flexibility, adaptation, and pragmatic problem-solving (according to the definition of long-term orientation) are more likely to listen to more diverse music, both in terms of unique genres listened to and entropy in their music distribution over genres. These countries' listening events are also more evenly spread over a variety of genres (lower *Vol.* σ scores).

For the cultural dimension of *indulgence*, we observe quite interesting and maybe surprising results. In fact, this dimension is highly significantly, negatively correlated to volume- and entropy-based diversity measures, in particular to the latter. Therefore, citizens of countries scoring high on indulgence, which means they tend to enjoy life and have a lot of fun, exhibit a smaller need for music diversity. This could, to some extent, be explained by a focus on certain genres that are commonly regarded as positive and happy, e.g., Pop, while avoiding music from dark genres, such as Death Metal.

The aspect of *uncertainty avoidance* is only slightly correlated to the relative volume diversity. *Power distance* and *masculinity* do not show significant correlation to any of the diversity measures.

5. CONCLUSIONS AND FUTURE WORK

In the presented study, we found distinct correlations between volume- and entropy-based music diversity measures and Hofstede's cultural dimensions, which showed to be in line with, and extended, prior results reported in [8]. We identified moderate, highly significant correlations for the aspect of individualism and volume-based diversity measures. Highly individualist societies thus listen to more diverse genres. For long-term orientation and indulgence, we also found moderate, highly significant correlations; in these cases not only for volume-based, but also for entropy-based diversity measures. For long-term orientation, this means that countries whose population can be characterized as flexible, pragmatic, and eager to adapt to changes show a higher level of diversity in their music consumption behavior. Populations characterized by high indulgence (happiness and enjoying life) in contrast show a significantly lower desire for music diversity.

Approaching diversity on a country level enables the creation of proxy measures for personalization in situations where data is limited. For example, the new user problem in recommender systems when there is not enough behavioral data yet to make personalization inferences with. To address this problem, users' personality, among other aspects, has attracted interest to make inferences for personalization [4, 9, 21]. One way to extract personality is facilitated by the increasing connectedness of applications and social media (e.g., single sign-on buttons). This allows exploitation of social media data for personality acquisition, for instance, from Facebook [1, 3, 5, 15], Twitter [10, 17], or Instagram [6, 7, 20]. However, a connection with the user's social media account is still needed. Our results could be used to make inferences about the user's diversity needs based solely on their country, which is often available through the user's account information.

Future work will investigate diversity formulations that also take into account similarities between genres. In particular when using fine-grained genre terms, an approach based on extending the one presented in [16] may yield additional interesting findings. In addition, taking into consideration similarities and affinities between countries, e.g., via Wikipedia articles [14], may allow for a more decent modeling of culture. In this study we only focused on Hofestede's cultural dimensions. However, although less comprehensive, there are other cultural dimensions (e.g., GLOBE [12] and Trompenaar's [22] cultural dimensions) available. It would be nice to investigate the consistency between the different cultural dimensions in the future.

6. ACKNOWLEDGMENTS

This research is supported by the Austrian Science Fund (FWF): P25655.

7. REFERENCES

[1] M. D. Back, J. M. Stopfer, S. Vazire, S. Gaddis, S. C. Schmukle, B. Egloff, and S. D. Gosling. Facebook profiles reflect actual personality, not self-idealization. *Psychological Science*, 21:372–374, 2010.

[2] D. Bollen, B. P. Knijnenburg, M. C. Willemsen, and M. Graus. Understanding choice overload in recommender systems. In *Proceedings of the 4th ACM Conference on Recommender Systems*, pages 63–70, Barcelona, Spain, September 2010.

[3] F. Celli, E. Bruni, and B. Lepri. Automatic personality and interaction style recognition from facebook profile pictures. In *Proceedings of the ACM International Conference on Multimedia*, pages 1101–1104, Orlando, Florida, USA, November 2014.

[4] B. Ferwerda and M. Schedl. Enhancing Music Recommender Systems with Personality Information

and Emotional States: A Proposal. In *Proceedings of the 2nd Workshop on Emotions and Personality in Personalized Services (EMPIRE)*, Aalborg, Denmark, July 2014.

[5] B. Ferwerda, M. Schedl, and M. Tkalcic. Personality traits and the relationship with (non-) disclosure behavior on facebook. In *Proceedings of the 25th International World Wide Web Conference Companion (WWW)*, Montreal, Canada, April 2016.

[6] B. Ferwerda, M. Schedl, and M. Tkalčič. Predicting Personality Traits with Instagram Pictures. In *Proceedings of the 3rd Workshop on Emotions and Personality in Personalized Services (EMPIRE)*, Vienna, Austria, September 2015.

[7] B. Ferwerda, M. Schedl, and M. Tkalčič. Using Instagram Picture Features to Predict Users' Personality. In *Proceedings of the 22nd International Conference on MultiMedia Modeling (MMM 2016)*, Miami, USA, January 2016.

[8] B. Ferwerda, A. Vall, M. Tkalčič, and M. Schedl. Exploring Music Diversity Needs Across Countries. In *Proceedings of the 24th International Conference on User Modeling, Adaptation and Personalization (UMAP)*, Halifax, Canada, July 2016.

[9] B. Ferwerda, E. Yang, M. Schedl, and M. Tkalčič. Personality Traits Predict Music Taxonomy Preferences. In *ACM CHI '15 Extended Abstracts on Human Factors in Computing Systems*, Seoul, Republic of Korea, April 2015.

[10] J. Golbeck, C. Robles, M. Edmondson, and K. Turner. Predicting Personality from Twitter. In *Proceedings of the 3rd IEEE International Conference on Social Computing (SocialCom)*, Boston, USA, October 2011.

[11] G. Hofstede, G. J. Hofstede, and M. Minkov. *Cultures and Organizations: Software of the Mind*. McGraw-Hill, USA, 3rd edition, 2010.

[12] R. J. House, P. J. Hanges, M. Javidan, P. W. Dorfman, and V. Gupta. *Culture, leadership, and organizations: The GLOBE study of 62 societies*. Sage publications, 2004.

[13] S. Kitayama and H. Park. Cultural Shaping of Self, Emotion, and Well-Being: How Does It Work? *Social and Personality Psychology Compass*, 1(1):202–222, 2007.

[14] P. Laufer, C. Wagner, F. Flöck, and M. Strohmaier. Mining cross-cultural relations from Wikipedia — A study of 31 European food cultures. In *Proceedings of the ACM Web Science Conference (WebSci)*, Oxford, UK, June–July 2015.

[15] G. Park, H. A. Schwartz, J. C. Eichstaedt, M. L. Kern, M. Kosinski, D. J. Stillwell, L. H. Ungar, and M. E. Seligman. Automatic Personality Assessment Through Social Media Language. *Journal of Personality and Social Psychology*, 108, November 2014.

[16] M. Park, I. Weber, and M. N. amd Sarah Vieweg. Understanding Musical Diversity via Online Social Media. In *Proceedings of the 9th AAAI Conference on Web and Social Media (ICWSM)*, Oxford, England, UK, May 2015.

[17] D. Quercia, M. Kosinski, D. Stillwell, and J. Crowcroft. Our twitter profiles, our selves: Predicting personality with twitter. In *Proceedings of the 3rd IEEE International Conference on Social Computing (SocialCom)*, Boston, USA, October 2011.

[18] M. Schedl. The LFM-1b Dataset for Music Retrieval and Recommendation. In *Proceedings of the ACM International Conference on Multimedia Retrieval (ICMR)*, New York, USA, June 2016.

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

[20] M. Skowron, B. Ferwerda, M. Tkalčič, and M. Schedl. Fusing Social Media Cues: Personality Prediction from Twitter and Instagram. In *Proceedings of the 25th International World Wide Web Conference Companion (WWW)*, Montreal, Canada, April 2016.

[21] M. Tkalčič, B. Ferwerda, D. Hauger, and M. Schedl. Personality Correlates for Digital Concert Program Notes. In *Proceedings of the 22nd International Conference on User Modeling, Adaptation and Personalization (UMAP)*, Dublin, Ireland, June–July 2015.

[22] F. Trompenaars and C. Hampden-Turner. *Riding the Waves of Culture — Understanding Diversity in Global Business*. McGraw-Hill Education, 3rd edition, 2012.

[23] M. C. Willemsen, B. P. Knijnenburg, M. P. Graus, L. C. Velter-Bremmers, and K. Fu. Using latent features diversification to reduce choice difficulty in recommendation lists. In *Proceedings of the RecSys 2011 Workshop on Human Decision Making in Recommender Systems*, Chicago, USA, October 2011.