Amsterdam University Press

COMPUTATIONAL COMMUNICATION RESEARCH 5.1 (2023) 1–32
HTTPS://DOI.ORG/10.5117/CCR2023.1.15.GONG

Media selection is highly predictable, in principle

Xuanjun Gong
Department of Communication, University of California, Davis
Department of Statistics, University of California, Davis

Richard Huskey
Department of Communication, University of California, Davis
Cognitive Science Program, University of California Davis
Center for Mind and Brain, University of California Davis

Abstract
Media research is, in part, interested in accurately explaining and predicting people's media selection. Explanation is an accurate description of the causal mechanisms that govern media selection whereas prediction is focused on making accurate inferences about unobserved data. However, meta-analyses demonstrate that existing media selection theories and models have limited explanatory accuracy. The predictive accuracy of these theories and models is unknown. Our project bridges this gap by empirically specifying how predictable, in principle, media selection is. To achieve this ambition, we articulate key conceptual distinctions between explanation and prediction. Subsequently, we report three empirical studies that examine prediction accuracy as a function of model complexity and estimate the theoretical maximum predictability of people's music-listening and web-browsing behaviors. Approximately 80% of music selection and 60% of web-browsing behaviors are predictable. Moreover, a simple Markov Chain model that uses information about people's prior media selection can achieve about 20% prediction accuracy for music selection and 10% accuracy in predicting web-browsing. By estimating the maximum predictability of people's media selection behavior, we gain a first-ever benchmark by which media selection theories and models can be compared. More broadly, we show how simple models that account for the sequential dependency in media selection lend new insights and suggest novel directions for future theory development.

Keywords: prediction, explanation, media selection, computational modeling, sequential modeling

Introduction

Media research is, in part, interested in identifying mechanisms and developing theories to explain people's complex media use behaviors. By explanation, we mean that a theory specifies the causal mechanisms that give rise to an observable behavior. A commonly used research paradigm for developing and testing explanations is to build diagrams of constructed variables connected by directional arrows, and then use statistical Null-Hypothesis-Significance-Testing (NHST) to answer research questions about whether an expected effect exists in the collected dataset of self-reported or behavioral responses (Levine, Weber, Hullett, et al., 2008; Levine, Weber, Park, & Hullett, 2008; Pearl & Mackenzie, 2018).

These explanatory approaches are common in the social sciences, including the field of Communication. And they have advanced our understanding of media use behaviors. This includes, but is not limited to, entertainment media choices (Knobloch, 2003), informational media consumption (Knobloch-Westerwick, 2014), media multitasking (Ophir et al., 2009), social media usage (Wang et al., 2012), individual differences in media use (Weaver, 1991), and more. Under different theoretical perspectives and assumptions, media usage research continues to flourish by testing and updating numerous models and theories, many of which are related to media selection, such as uses and gratifications (Katz et al., 1973), mood management theory (Zillmann, 1988), and information utility theory (Hastall, 2009).

These developments have created a massive landscape of explanatory models and theories. They have also received fierce criticism for their (in)ability to replicate (Dienlin et al., 2021; LeBel et al., 2018), lack of integrative theoretical framework (Huskey et al., 2020; Lang & Ewoldsen, 2010), and most critically, their low explanatory power (Lang, 2013). In fact, meta-analytic work shows that effect sizes for explanatory studies focused on media selection range from r = 0.07–0.24 (Rains et al., 2018). This indicates that, at best, 5.8% of the variation of media selection behaviors can be explained by existing theory[1]. These issues suggest that our existing theories are insufficient for explaining a majority of the variation in media selection

[1] Explainability is best indicated by the total explained variation (e.g., R squared or Eta squared) of the statistical model including all variables. However, model explainability is often either unreported or not-interpreted given the hypothesis-oriented nature of explanatory studies, which typically require reporting and interpreting effect size and p-values for a single hypothesis test related to a specific explanatory variable of interest. For this reason, we can only estimate the total explained variation with the highest effect size reported in existing meta-analytic work.

behaviors.

Communication researchers also want to build theories that predict. By predict, we mean a theory's ability to make accurate inferences about unobserved data. Analytically, this means fitting (training) a statistical model on one set of data, and applying (testing) that same model and its parameter estimates on a new set of data, with high accuracy. Theoretically, explanatory and predictive accuracy are identical. This may explain a common perception among social scientists is that theories with high explanatory accuracy will also have high predictive accuracy. However, this often simply is not true. In fact, for explanatory theories and models, predictive accuracy is often substantially lower than explanatory accuracy (Yarkoni & Westfall, 2017). As already mentioned, the current state of affairs for our ability to explain people's media selection is quite low. What about our prediction accuracy?

We do not know. We do not currently have an idea of how predictable media selection should be, in principle. As a consequence, we also do not know how explainable people's media selection should be. Answering this question is vital. Theoretically, an answer gives us insight into how *knowable* people's media selection behavior is. It also gives us some hints for developing explanations that increase knowledge. Practically, people's lives are impacted by predictive media selection models. These include recommendation algorithms developed to enhance people's experiences when streaming video, listening to music, scrolling through a social media feed, and browsing the internet. Right now, existing predictive models that govern recommendation algorithms achieve accuracies that are substantially higher than state-of-the art media selection theories (e.g., Covington et al., 2016). By discovering the theoretical maximum predictability of people's media selection, we gain a new benchmark for evaluating both explanatory and predictive theories and models. We also gain insight into how to bridge the accuracy gap between explanatory and predictive frameworks.

In this article, we show that media selection is highly predictable, in principle. Accordingly, it should be possible, in principle, to develop theories that substantially improve on existing explanatory accuracy. We further argue that gaps between predictive and explanatory accuracy can be bridged by estimating, evaluating, and comparing the *predictability* of derived explanatory models, which eventually complement the traditional explanation of media behaviors. In the following sections, we first articulate key distinctions between explanation and prediction, then illustrate the benefits and cautions of building predictive models. We illustrate these distinctions by an empirical study that evaluates the predictability and explainability of simple as compared to complex models on dichotomous preferential movie selection behavioral data. Consistent with prior meta-analytic work, and regardless of model complexity, these results show relatively low model accuracy. Accordingly, we ask what is the theoretical upper bound for model accuracy? To answer this question, we present two empirical studies that measure the theoretical maximum predictability of people's real-world sequential media selection behaviors using a large-scale music listening dataset, as well as a large-scale web-browsing dataset. We show that people's music listening behavior is 80% predictable in principle. Similarly, web-browsing has 60% maximum predictability. We then apply a simple Markov Chain (MC) predictive model aimed at achieving the maximum predictability of music listening and web-browsing behaviors. This analysis shows that a simple model that accounts for temporal dependency in the data can achieve up to $\sim$20% prediction accuracy, which substantially outpaces the accuracy of state-of-the-art media selection theories. We conclude with a discussion of our results and their implications for theory.

Explanation, Prediction, And Why A Model's Explanatory Accuracy Often Tells Us Very Little About Its Predictive Accuracy

Recently, discussions that distinguish between prediction and explanation have been raised in multiple social science disciplines including psychology (Yarkoni & Westfall, 2017), behavioral economics (Peterson et al., 2021), political science (Lin, 2015), computational social science (Hilbert et al., 2018; Hofman et al., 2021), as well as media research (Fisher & Hamilton, 2021). However, predictive modeling methods have been rarely or incorrectly implemented in media research. Thus, the following section aims to clear misconceptions about what predictive modeling is, give suggestions for building predictive models, and provide an exemplary study for measuring the predictability of media behaviors.

What Distinguishes Explanatory and Predictive Models?

Earlier we discussed the conceptual distinctions between explanation and prediction. As a reminder, explanation is the application of statistical models for testing causal hypotheses about relationships between theoretical constructs, whereas prediction is the application of statistical models or data mining techniques to generate predictions for new or future observations (Shmueli, 2010). In this section, we discuss the applied implications of these distinctions.

First, explanation and prediction serve different proximal research goals. In lay language, explanation aims to *explain* and prediction aims to *predict*. More complicated, as illustrated by Shmueli (2010), explanatory studies try to explain the true associations or causal mechanisms (F) using a hypothetical relationship (f) between theoretical constructs. Explanatory studies collect independent variables (X) and dependent variables (Y), estimate statistical models ($\hat{f}(X)$), and use NHST to verify if $\hat{f}(X) = f(X)$ is statistically plausible. Here the proximal goal of explanation is testing the hypothetical relationship f, with the collected data (X and Y) and statistical models ($\hat{f}(X)$) as tools. In contrast, predictive studies aim to generate good predictions ($\hat{Y}$) for new observations of target variables (Y) by estimating a model ($\hat{f}(X)$). Thus the proximal goal of prediction is the targeted variables (Y), with the collected predictors (X) and hypothetical relationships (f) as the tool. Even though explanation and prediction serve different proximal goals, both should pursue the same ultimate goal to approximate the true relationships (F) by updating hypothetical relationships (f) and minimizing the errors in the hypothetical models.

The second conceptual distinction between explanation and prediction involves model errors and the bias-variance trade-off. Precisely, model errors, measured as the sum of errors, can be decomposed into three types of errors, that are: model bias, which refers to the errors resulting from the misspecification of the hypothetical models, model variance, which refers to the errors resulting from the variation of collected samples from the true population, and the irreducible or true errors, which refers to the errors resulting from the stochastic data generating process and are irreducible by any modeling methods (Hastie et al., 2009). Explanation and prediction reduce model bias and variance errors in distinct ways. To illustrate, imagine a study testing the relationship between people's comedy movie preferences and their age.

Model bias exists because a correlational relationship between comedy movie preference and people's age is unlikely to be the true explanatory relationship. Why? Is age itself the true explanatory mechanism? Or is age some correlate of the true mechanism. A true explanatory model specifies the mechanism, and not its correlate. Identifying these mechanisms is a core goal of lifespan-based media selection research specifically (Mares et al., 2008; Shifriss et al., 2015), and explanatory research more generally.

To reduce model bias, both explanatory and predictive studies need to add relationships between the outcome variable and independent variables into the model to capture the variation of outcome variables. Explanation and prediction methods reduce model bias differently, in a way such that explanation modeling methods emphasize the correctness of models while predictive modeling emphasizes the completeness of models. Specifically, explanatory studies propose theory-driven hypothetical relationships, reject null relationships, and retain statistically significant relationships in the model. As a result, to ease the procedures of statistical testing and interpretation of the testing results, explanatory studies usually consider a small size of simple linear relationships when testing hypothetical relationships. However, these simple linear relationships, such as the positive or negative association between age and media preference, usually fail to capture the complete variation of media preferences. On the other hand, prediction modeling methods use data-driven approaches with complex models, by adding all possible hypothetical relationships into the model and increasing the complexity of the model to capture variation in outcome variables as much as possible. An example, neural networks usually do a better job in function approximation compared to simple explanatory linear models (Peterson et al., 2021), thus performing better in predicting outcome variables. However, the complexity of these models often makes them difficult or even impossible to interpret.

Model bias can be reduced by increasing model complexity. However, doing so often results in overfitting issues (Yarkoni & Westfall, 2017). In short, overfitting increases model variance. What is model variance? Model variance is the result of a variation of estimated model parameters due to the variation of sampling distributions. Imagine that ten different research groups collect ten independent datasets. Each group then fits the same statistical model, as defined by theory, for each of the ten independent datasets. We will observe model variance. This is because of the variance in the sampling distribution that comprises each collected dataset. When each unique dataset is used to fit a regression model, the estimated regression coefficients will differ across each different dataset as a function of the variance in the sampling distribution (for a concrete example of the impact of model variance on issues of statistical inference and reproduibility, see Marek et al., 2022).[2] As model complexity increases to reduce model bias,

[2]For linear models, model variance is calculated as $p * \sigma^2$ where p = the number of parameters in the model. Accordingly, model variance increases as the number of model parameters increases. Of course, when the number of model parameters is held constant, model variance decreases as the sample size in each sampling distribution increases.

model variance increases. This is known as the bias-variance tradeoff.

Model variance matters for explanatory and predictive models in different ways. For predictive models, model variance emerges due to high model complexity, and works directly against predictive accuracy. Predictive accuracy decreases as model variance increases, and this relationship is hard to counteract, especially with complex and sophisticated models because the larger size of model parameters means higher model variance (Domingos, 2012). To deal with high model variance and overfitting issues, predictive models usually use regularization techniques. For instance, a common approach is shrinkage methods, such as ridge regression or principal component regression, to shrink or eliminate parameters to avoid model variance and improve predictions.

For explanatory models whose main research focus is not predictive accuracy, model variance emerges due to procedural overfitting, which can explain why some replication attempts are successful, and why some fail (Yarkoni & Westfall, 2017), particularly in replication attempts with larger sampling distributions that more accurately reflect the true population distribution (Ebersole et al., 2016; Marek et al., 2022; Open Science Collaboration, 2015). Explanatory studies usually ignore the existence of model variance. As explained above, high model variance results in high model errors. However, this increase of model errors can hardly be recognized by explanatory studies, because only one sample is collected to estimate and evaluate the model without any regularization technique to overcome overfitting problems (Yarkoni & Westfall, 2017). Statistical tools exist for explanatory modeling to control model variance by model comparison using evaluation metrics such as adjusted R squared (R_a^2), Akaike information criterion (AIC; Akaike, 1998), Bayesian information criterion (BIC; Schwarz, 1978). However, these tools are not always implemented in explanatory studies. Moreover, these model performance metrics are usually limited as a generic method to evaluate models, because (1) they rely on a set of complicated assumptions on testing models and data, and (2) they are computed by the parameter counts and the maximized likelihood which might be unavailable to some models (Burnham & Anderson, 2004; Vrieze, 2012; Yarkoni & Westfall, 2017).

In summary, explanation and prediction have different focuses on model errors. Explanation focuses on building a *correct* model which reduces the bias of the estimated models. But prediction aims to build a *good* model which decreases the model bias while also controlling the model variance (Shmueli, 2010). These different focuses on model error lead to different modeling approaches. Explanatory modeling usually use NHST to reject non-significant hypothetical relationships in an effort to reduce the model bias. On the other hand, predictive models usually maintain a large size of possible hypothetical relationships to create a complex model in order to reduce model bias, and use shrinkage methods to restrict model variance.

Finally, explanation uses *in-sample* data to evaluate a hypothesized model, while prediction uses *out-of-sample* data to assess model performance. For explanatory models, this means gathering data and fitting a model on the singular dataset. For predictive models, this means gathering two or more independent datasets. A commonly used technique for achieving this aim is to split one large dataset into two or more smaller independent datasets. Subsequently, one dataset is used to fit (train) the model and the other dataset is used to evaluate (test) the model. This is known as *k-fold* cross-validation and the number of folds (k) is $\leq n$ where $n =$ number of observations in the dataset. In *k-fold* cross-validation, the model and parameter estimates trained on one or more fold(s) are then fitted to the data in the subsequent fold(s). Mean square error (MSE) and other model performance parameters (e.g., accuracy, precision, recall, f1) are then calculated to evaluate the model's performance on each fold, and a distribution of model evaluation metrics is obtained. The best performing models achieve low MSE scores and high model performance when fitted to multiple unique test datasets.

Most linear modeling techniques, when applied to explanatory models, fit a model that minimizes MSE as best as possible. For example, a regression model estimates parameters that are optimized to reduce the model's MSE (this is also known as the line of best fit). Predicting into an out-of-sample dataset is more difficult than parameter estimation for an in-sample dataset. This is because the expected MSE for an out-of-sample data point can be decomposed into model variance, model bias, and irreducible (or true) error (Hastie et al., 2009). Thus, even though $MSE_{training} = MSE_{test}$ is theoretically possible, in most applications, $MSE_{training} \ll MSE_{test}$. Therefore evaluations that use a model's accuracy (e.g., effect size), calculated using in-sample data, often overestimate the model's predictive accuracy.[3] This is why we earlier said that good explanatory models are often not also good predictive models.

If empirical explanatory models and meta-analyses of explanatory models possibly overestimate a model's predictive accuracy, a different solution

[3]This also is true for meta-analyses conducted on explanatory research using in-sample model estimation. This means that the effect size reported in such a meta-analysis likely overestimates predictive accuracy.

is necessary. In what follows, we first examine the discrepancy between explanatory and predictive accuracy in empirical models of media selection that examine static media choices. We also recognize that people regularly select media sequentially, and recent theorizing has considered this temporal dependency (Gong & Huskey, in press). Accordingly, we examine how much predictive information temporal dependencies in selection data offer, and describe an approach for estimating the theoretical maximum predictability of people's sequential media selection. We then report the results of a simple Markov Chain model designed to accurately predict people's sequential media choices by accounting for temporal dependency in the data.

Method

Open Science Practices

In accordance with calls for open practices in communication science (Bowman & Keene, 2018; Dienlin et al., 2021; Lewis, 2020), the code necessary to reproduce these analyses is posted to GitHub[4]. The data used in this project are already open and can be accessed from their relevant repositories.

Building Models of Static and Sequential Media Selection

Media selections are discrete responses, meaning the possible value space for media choices is separate individually. Thus model building for media choices requires a probabilistic or deterministic specification of the relationship between discrete media options with independent variables (Gong et al., 2023). How to specify this relationship depends on research questions and the media selection contexts. For instance, a generalized linear model can predict media choices for new users by assuming each media selection is independent with each other selection, with explanatory variables such as user characteristics, situational factors, or media option characteristics. On the other hand, a sequential model can predict media choices for new episodes of existing media users, with predictive variables such as media choice histories and previous evaluations of media content.

In the following section, we introduce both types of models in their simple format. For static media selection, we discuss two types of models: linear logistic models (LLM), a type of generalized linear model, and Support Vector Machine (SVM), a type of black-box machine learning model.

[4]https://github.com/cogcommscience-lab/media_selection_predictability

For sequential media selection, we discuss: entropy based approaches for estimating maximum model predictability, an approach that examines the randomness in sequential data, and Markov Chain (MC) models, a type of model that accounts for temporal dependency in data.

Generalized Linear Model: Linear Logistic Model

LLM assumes each media choice observation is independently made by media users without sequential dependencies, meaning future choices do not depend on previous choices. It estimates the probability of choosing a specific media option with a sigmoid function transforming the linear combination of independent variables (x_i) and regression coefficients (β_i) into a probability value (P) ranging from zero to one (Equation 1).

$$P = \frac{1}{1 + exp(-\beta_0 - \sum_i^m \beta_i x_i)} \tag{1}$$

Black Box Model: Support Vector Machine

Distinct from generalized linear models, which estimate the probability of a given option being selected, SVM deterministically classifies media choices by constructing the max-margin hyperplane separating media choice observations into distinct classes depending on independent variables (James et al., 2013). With kernel methods that construct the high-dimensional space of all given observations, SVM captures the higher-order non-linear relationships between independent variables and media selection choices. Thus SVM is capable of predicting media selection as resultant of hard-to-explain complex mechanisms. However, due to its high opaque complexity, SVM is usually difficult to interpretively explain the mechanisms of how independent variables lead to media selection.

Estimating Maximum Predictability of Sequential Media Selection

The accessibility of digital trace data provides new opportunities for researchers to develop theories and models to predict and understand people's behaviors. In principle, people's behaviors are highly predictable from their previous behavioral histories. For example, with mobile geo-tracking data, studies found that people's mobility behaviors have 93% potential predictability, which indicates the non-randomness of high regularity in human physical mobility patterns (Song et al., 2010). Moreover, studies have shown that this theoretical predictability upper bound is empirically achievable via MC models (including the classic MC model, hidden MC models, and mixed MC models), neural networks, and Bayesian networks (Lu et al., 2013). Similarly, studies found that people's dyadic communication processes in open source collaborative communication platforms (e.g., Github, Wikipedia) have about 80% potential predictability, and one fourth of the predictability of the collaboration sequences originates from the dynamic structure of the dyadic communication processes, while three-fourths of the predictability comes from the static frequencies (Hilbert et al., 2018).

When examining sequential selection, we need models that can account for temporal dependencies in the data. Accordingly, we will focus on estimating the theoretical maximum predictability of people's sequential media selection behavior, and its empirical predictability using MC modeling. We define sequential media selection as the trajectory of media content that a user traverses in a media space. Said differently, our study explores how people select media over time and investigates the theoretical upper boundary for how predictable media selection is. We do this for both music listening and web-browsing. We motivate this analysis using music listening, but the analytical logic is identical when applied to web-browsing.

Maximum predictability can be estimated by measuring the entropy of trajectories, which is a discrete sequence of digital traces (for a full demonstration of the logic underpinning equations 2 - 7, see Song et al., 2010). This sequence can be denoted as:

$$X_j = \{X_1, X_2, ..., X_T\} \tag{2}$$

Here, X_1 encodes the music track an individual is listening to at time 1, and X_T denotes the track at time T. Entropy measures the disorderness and randomness of the trajectory, thus higher entropy will lead to lower predictability. Specifically, empirical entropy ($S_{empirical}$) can be calculated from the true trajectories observed as specified in equation 3:

$$S_{empirical} = -\sum_{X_i' \subset X_i} P(X_i') \times log_2[P(X_i')] \tag{3}$$

In equation 3, $P(X_i')$ is the probability of finding a particular time-ordered subsequence X_i' in the trajectory X_i.

This approach also requires an appropriate null model. We can calculate two types of null entropy from null models of the trajectory. A random entropy (S_{rand}) approach assumes visited tracks are uniformly distributed and is given by equation 4, where N is the number of unique tracks in the trajectory.

$$S_{rand} = log_2(N) \tag{4}$$

By comparison, uncorrelated entropy (S_{unc}) assumes that the order of visited tracks can be randomly shuffled, and is given by equation 5, where P_k denotes the probability of track k being visited.

$$S_{unc} = -\sum_{k=1}^{N} p_k \times log_2(P_k) \tag{5}$$

With the estimated entropies for each trajectory, the upper bound of the predictability (Π) can be given by solving Fano's inequality (equation 6 and 7):

$$S = H(\Pi) + [1 - H(\Pi)] \times log_2(N - 1) \tag{6}$$

$$H(\Pi) = -\Pi log_2(\Pi) - (1 - \Pi) log_2(1 - \Pi) \tag{7}$$

Thus, we can calculate the maximum predictability ($\Pi_{empirical}$) of the empirical trajectories with the empirical entropy ($S_{empirical}$), maximum predictability (Π_{rand}) of the null uniformly distributed trajectories with random entropy (S_{rand}), and maximum predictability (Π_{unc}) of the null shuffled trajectories with uncorrelated entropy (S_{unc}). These values are calculated for each trajectory for each participant in each database (music listening, web-browsing). Inference testing is then conducted on using the empirical and null distributions using an independent samples t-test with α = .05 (two-tailed).

Sequential Model: Markov Chain Model

MC based models are often used to interrogate temporal dependency in choices. For instance, Lu et al. (2013) found that temporal dependencies can be modeled to achieve high predictability of human physical mobility. In the current study, as a preliminary attempt to investigate people's sequential media selection, we built an order one MC model in an effort to approach theoretical maximum predictability. Essentially, the order one MC model assumes each media selection event (i.e., choosing a song to listen to or a website to visit) only depends on the adjacent previous choice and is independent of every other previous choice. Thus, given the choice C_i at time

$t-1$, the MC model will predict the probability of choosing a media option C_j at time t as the conditional probability $P(C_j|C_i)$, which is estimated as in equation 8, where N_{c_i,C_k} denotes the number of occurrences of choosing C_k after choosing C_i, and m denotes the total number of unique options.

$$P(C_j|C_i) = \frac{N_{C_i,C_j}}{\sum_{k=1}^{m} N_{C_i,C_k}} \quad (8)$$

Data Description and Pre-Processing

In this study, we use three datasets. The first dataset (see Dichotomous Movie Selection section) allows us to apply models of various complexity in order to examine how model complexity shapes explanation and prediction accuracy (given the same set of user and media features). The second and third datasets (see Sequential Music and Website Selection section) provide temporal information which allows us to examine sequential dependencies in the data. One constraint of the first dataset is that it lacks temporal features, whereas the second and third datasets lack user and media features. When examined holistically, it becomes possible to evaluate which features have the strongest contribution to model accuracy while also demonstrating that temporal dependenices in sequential selection can yield more accurate models.

Dichotomous Movie Selection Dataset

We obtained a movie selection dataset from a dichotomous movie selection task, where participants (N = 301) repeatedly (n = 140) chose their preferential movie option among two movie options that systematically varied in valence and arousal (Gong et al., 2023). In this study, participants were also placed into different mood states that systematically varied in valence and arousal. This dataset contains a total of 42,140 (301×140) movie selection observations. From this dataset, in addition to selection observations, we extracted media movie features (i.e., valence and arousal difference between two options), media user characteristics (i.e., race, gender, age), and situational factors (i.e., mood valence and arousal).

In order to evaluate model predictability, i.e., prediction accuracy for out-of-sample media choices, we split the dataset into train (80%) and test (20%) datasets by randomly drawing media choice observations for each participant separately. Accordingly, the training dataset includes 33,712 (301×112) media choices, while the test dataset includes 8,428 (301×28) media choices.

Model explainability was evaluated by examining within-sample model fit for the overall dataset, as is standard procedure for explanatory models.

Sequential Music and Website Selection Dataset

To examine maximum predictability and MC predictability of people's sequential media selection, we obtained two large-scale digital trace datasets: a music-listening dataset from *Last.fm*[5] (Celma, 2010), and a web-browsing dataset from the *Web History Repository*[6]. The music listening dataset was collected from *Last.fm* (a music recommendation platform) using the platform's API. The dataset contains 19,098,862 listening records for 992 users. The web browsing dataset was collected through voluntary submissions of web-browsing histories from contributors, which include a total of 5,155,149 web-page visits from 524 users.

For the music listening trace data, we excluded music listeners who: (a) listened to more than 20,000 unique tracks or less than 200 unique tracks (n = 64), (b) only listened to the same tracks (n = 1), or (c) only listened to novel tracks (n = 31)[7]. We excluded people with a small number of listening records because entropy estimation is unstable when the sequence is short. People with a large number of unique tracks were excluded to make sure the computing time is reasonable. We also excluded people with only one unique track and with all unique tracks, because their predictability should artificially be 1 and 0 respectively. The preprocessed music listening data includes 17,779,387 listening records from 920 individuals.

For the same reasons, we excluded web browsers who: (a) visited more than 20k unique websites or less than 20 unique websites (n = 48), (b) only visited the same website (n = 1), or (c) only listened to novel websites (n = 11), from the web browsing data[8]. The final web browsing data includes 1,971,227 web browsing records from 465 individuals.

Results

Predicting and Explaining Independent Media Selection

In order to predict and explain people's movie choices, we estimated several models ranging from simple to complex, including simple and complex

[5]http://ocelma.net/MusicRecommendationDataset/lastfm-1K.html
[6]http://webhistoryrepository.l3s.uni-hannover.de/download.php
[7]These filtering criteria result in overlapping filtered subjects.
[8]Same as in music listening dataset, filtering criteria result in overlapping filtered subjects.

LLMs, ridge logistic models of complex relationships, and a black-box machine learning SVM model. These LLMs are capable of generating explanations and predictions for new users or new episodes of existing users, because these models assume that each media selection event is independent of each other. We estimated LLMs using the *lme4* package (Bates et al., 2015), the ridge logistic model using the *glmnet* package (Tay et al., 2023), and the SVM model using the *scikit-learn* package (Pedregosa et al., 2011). Then, we evaluated model predictability (i.e., percentage of accurate predictions) for out-of-sample media selection in the test dataset, and explainability (i.e., percentage of accurate estimates) of in-sample media selection in the training dataset (Figure 1).

We found that, compared to chance accuracy at 50%, the simple explanatory LLM with only movie features (i.e., movie valence and movie arousal) predicts 59.0% out-of-sample choices and explains 58.9% in-sample choices. After adding situational factors (i.e., mood valence and mood arousal) and the two-way interactions between situational factors and movie features, the simple explanatory LLM predicts 59.0% out-of-sample choices and explains 59.0% in-sample choices. Furthermore, the simple explanatory LLM, which includes the interaction between movie features and media user race predicts 59.1% out-of-sample choices and explains 59.2% in-sample choice. Similarly, simple explanatory LLM with movie features and media user gender predicts 60.3% out-of-sample choices and explains 60.0% in-sample choices. Simple explanatory LLM with movie features and media user age predicts 60.7 out-of-sample choices and explains 60.5% in-sample choices. Finally, the LLM models which include all two-way interactions between media features and other independent variables (i.e., mood, race, gender, age) predicts 61.7% out-of-sample choices and explains 61.9% in-sample choices.

We also estimated more complex predictive models and evaluated their predictability and explainability. We found that the complex LLM with seven-way interaction terms of all independent variables (i.e., movie features, user characteristics, and situational factors) predicts 62.4% out-of-sample choices and explains 64.1% in-sample choices. After applying regularization techniques (i.e., L2 norm regularization and cross validation), the ridge logistic model of the seven-way interactions predicts 62.6% out-of-sample choices and explains 63.8% in-sample choices. Finally, the black box SVM model predicts 63.4% out-of-sample choices and explains 65.3% in-sample choices.

In summary, for simple models, as the models become more and more

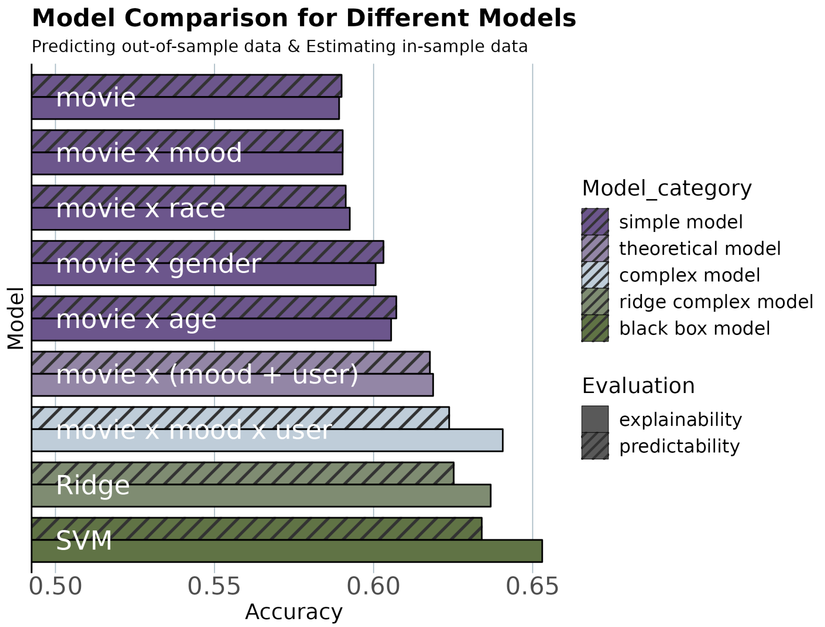

Figure 1: Accuracy evaluations for different models based on their predictability for out-of-sample choices and explainability for in-sample choices. The X-axis encodes predictive accuracy, and starts from 0.5 because the base chance predictive accuracy is 0.5 for binary prediction. Here, the theoretical model represents the model based on Mood Management Theory as specified in Gong et al. (2023).

complex by adding linear terms into the LLM, the model's predictability and explainability increase simultaneously. However, at a certain point when the model is becoming too complex to interpret (e.g., the seven-way interaction LLM), the model starts overfitting and shows a large discrepancy between predictability and explainability due to high model variance. Regularization techniques used in the ridge model help to address the overfitting issue by slightly increasing predictability and lowering explainability. Finally, the black box SVM model achieved the highest predictability. In general, our results reported here supported our arguments about (1) the existence of the distinction between predictability for out-of-sample data and explainability for in-sample data, (2) the bias-variance trade-off phenomenon and overfitting issues that emerge when a model becomes increasingly complex, (3) regularization techniques can help address the overfitting issues when building predictive models.

Maximum Predictability for Sequential Media Selection

In the above section, we examined how model complexity, coupled with theoretical variables of interest, shapes model accuracy when predicting and explaining static media selection. The results were largely consistent with existing meta-analytic (Rains et al., 2018) work that shows similarly low levels of accuracy. Is this the best that can be done? In this section, we draw on recent theorizing that temporal dependencies in sequential media selection offer important signal (Gong & Huskey, in press). With this focus on examining temporal dependency, we ask the following questions: what is the theoretical upper maxima for model accuracy, and what is the discrepancy between empirical model's accuracy and the theoretical maximum accuracy? Said differently, how predictable is real-world sequential media selection, and how well can we currently predict? To answer these questions, we estimated the theoretical maximum predictability for sequential music listening and web-browsing behaviors. The measured entropies and maximum predictability for each music listener (n = 920) and each web-browser (n = 465) were estimated by *scipy* optimization tools (Virtanen et al., 2020).

Since $S_{rand} \geq S_{unc} \geq S_{empirical}$, it can be shown that $\Pi_{rand} \leq \Pi_{unc} \leq \Pi_{empirical}$. Based on these results, the theoretical upper bound for people's media listening has maximum predictability at about 80% (Figure 2A). Second, the maximum predictability of empirical music listening trajectories ($M = 0.787, SD = 0.124$) is significantly higher than both uncorrelated listening trajectories ($M = 0.237, SD = 0.096; t(919) = 178.330, p < 0.001$), and random listening trajectories ($M = 0.000, SD = 0.000; t(919) = 191.296, p < 0.001$). Third, by comparing the predictability differences between the two null models, we noticed that there is only a small increase between the predictability of the shuffled null trajectories and the random null trajectories, from 0 to 0.237, compared to the increase of predictability by considering the order in music listening trajectories (Figure 2B). This means that each functions as a suitable null distribution.

Similarly, the theoretical upper bound for people's web-browsing has a maximum predictability at about 60% (Figure 3A). The maximum predictability of empirical web-browsing trajectories ($M = 0.628, SD = 0.096$) is significantly higher than both uncorrelated browsing null trajectories ($M = 0.173, SD = 0.085; t(464) = 139.064, p < 0.001$), and random browsing null trajectories ($M = 0.003, SD = 0.004; t(464) = 182.218, p < 0.001$). The difference in predictability of the shuffled and random null trajectories was also small (Figure 3B).

In summary, people's sequential media selection patterns have high

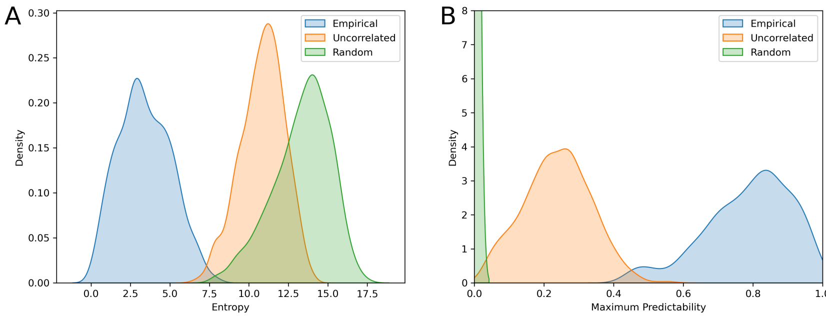

Figure 2: For the Last.fm dataset, these distributions show (A) measured entropy or S and (B) maximum predictability or Π.

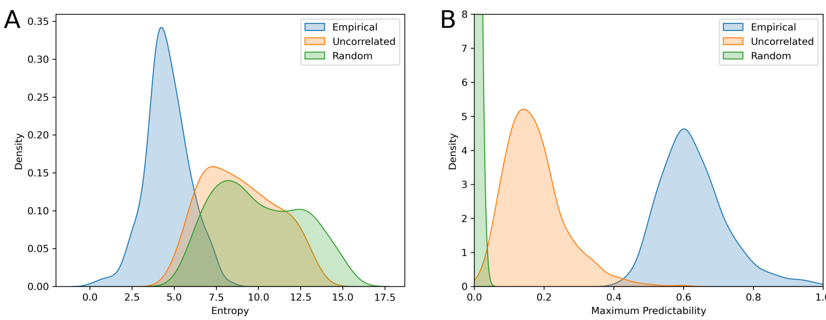

Figure 3: For the Web History Repository, these distributions show (A) measured entropy or S and (B) maximum predictability or Π.

regularity, such that we can possibly predict a maximum of approximately 80% of music listening and 60% of web-browsing behaviors. Interestingly, frequency (S_{unc}) of listening to a music track or visiting a website offers substantially lower predictability scores when compared to information that encodes the order ($S_{empirical}$) in which music tracks were listened to or websites were visited. Thus, it is critical to point out that, for future model building to explain or predict people's media selection, theories should focus more on the order of tracks being visited compared to the frequency of each track being visited.

Markov Chain Model Predictability for Sequential Media Selection

The above results emphasize the importance of order (temporal dependency) in estimating theoretical maximum predictability for music listening and web-browsing. Accordingly, we fitted order one MC models for each of the music listening (n = 920) and web-browsing (n = 465) trajectories. Next, following Lu et al. (2013), for each step in the music listening trajectories, we predict the subjects' next step by maximizing the conditional probability of moving from the current track/site to the next track/site. Finally, we compute the accuracy of the MC predictions, defined as the percentage of correct predictions.

For music listening, the results suggest that, with a simple Markov Chain, we can achieve prediction accuracy at about 20% ($M = 0.212, SD = 0.206$; Figure 4A). There is also a strong positive relationship between the prediction accuracy of the MC model and the estimated maximum predictability, and the MC model prediction accuracy is strictly bounded by the maximum predictability (Figure 4B). Finally, the difference between the MC prediction accuracy and the maximum predictability is still surprisingly large ($M = 0.575, SD = 0.106$; Figure 4A).

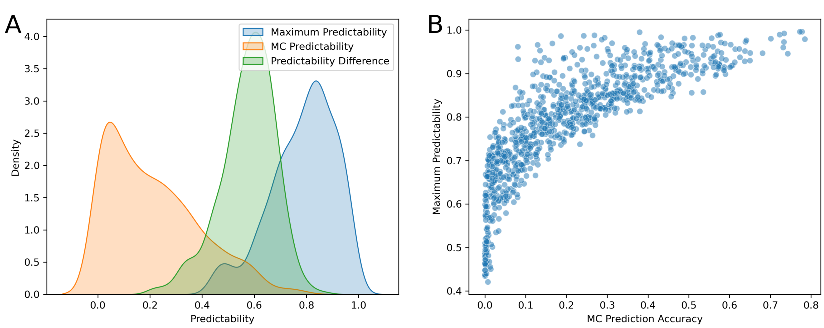

Figure 4: Results of MC model of music listening.

For web-browsing, we obtained similar results as music listening. The MC model achieves at about 10% prediction accuracy ($M = 0.096, SD = 0.080$; Figure 5A). And MC prediction accuracy has a strong positive relationship with maximum predictability (Figure 5B). The difference between the MC prediction accuracy and the maximum predictability is also large ($M = 0.533, SD = 0.045$; Figure 5A).

In summary, with the knowledge of an individual's current media choice (listening track/website visit), we can successfully predict 21.2% of the next

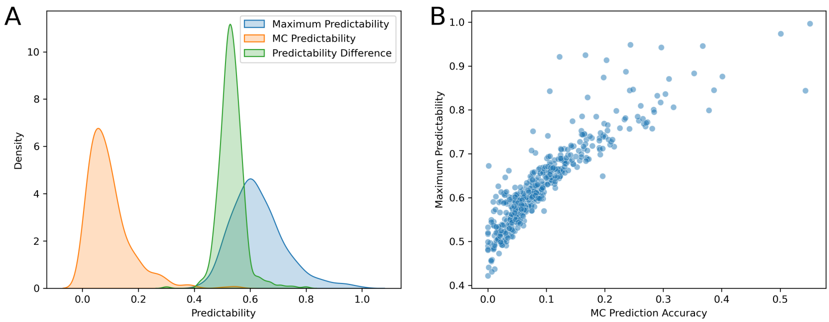

Figure 5: Results of MC model of web browsing.

listening track and 9.6% of the next website visit with a MC model. This high predictability indicates that people's media behaviors follow certain rules and order, which have not been well-exploited in previous media selection studies. The high potential predictability should strengthen communication researchers' confidence in theory building to explain and predict media behaviors. Moreover, the positive relationship between the MC prediction accuracy and the maximum predictability indicates the convergent validity of the theoretical maximum predictability measures. Lastly, the large gap between MC prediction accuracy and maximum predictability suggests that there is still a large space for future theoretical and modeling improvement to capture the variation of music listening choices.

Discussion

In the current study, we articulated the key theoretical distinction between explanation and prediction, and potential benefits of applying predictive methods in media selection research as complements to traditional explanatory approaches. Our empirical results showcased (1) that it is possible for predictive models to achieve higher accuracy compared to traditional explanatory models by increasing model complexity and adopting regularization techniques; (2) people's real-world media selection is highly predictable, in principle, and can be predicted by sequential predictive models with high accuracy, and (3) in addition to theoretical variables that are thought to predict static media selection, temporal dependencies in the data offer strong signal that can be used to make accurate predictions (Bialek, 2022) about sequential media selection (Gong & Huskey, in press).

Traditional media selection studies often focus on exploring how media features, user characteristics, and situational factors influence media selection at the population level. This approach views media selection in a snapshot scenario and assumes each media selection is an observation independent of its antecedents. However, real-world media selection happens in a dynamic environment, where future media choices depend on preceding media consumption in a regular, habitual, and routine way (Schnauber-Stockmann et al., 2023). Thus, the snapshot view of media selection necessarily represents an incomplete theoretical framework to guide explaining and predicting media selection, where sequential dependencies among media choices are non-trivial. Furthermore, our maximum predictability results show that empirical media selection, such as music listening and web browsing, can only be more accurately predicted after considering the sequential dependencies among media choices. This result suggests that, in addition to user characteristics, media features, and situational factors, future media selection research should consider sequential dependencies in theoretical, statistical, and computational modeling (for an example of what this might look like, see Gong & Huskey, in press). Doing so will account for sequential dependencies and dynamic variation in media selection.

Though uplifting to show a simple MC model can predict people's sequential media selection with high accuracy, it is still worthwhile to note that the MC model is a data-driven model with a large number of parameters (number of unique tracks squared). This presents a clear interpretability problem. What *was* it about a prior track that made it so successful at predicting a future track? In short, we simply do not know. The predictive accuracy of our MC models was high, but their explanatory interpretability is quite low. Looking to the future, there are at least two strategies for improving predictive accuracy: lean into explanatory ambiguity by building more complex data-driven models, or find a way to use data-driven predictive models to build elegant theory-driven models. In what follows, we discuss both.

Using Predictive Models to Inform Explanatory Models

Predictive models can help researchers navigate the exploration of complex relationships between theoretical constructs and aid in the discovery of novel constructs and measurements. This most certainly is true for media selection research. For instance, using a computational reward-learning model, researchers were able to predict the sequential timing of people's social media usage (Lindström et al., 2021) and find the relationship between past social media rewards and future social media usage. Similarly, by building a generative predictive model of Wiki-page browsing behaviors, researchers were able to measure behavioral tendencies (i.e., reinforcement and regularity) from predictive models, and then discover novel relationships between people's curiosity-related personalities and the estimated behavioral tendencies (Lydon-Staley et al., 2020). Furthermore, large-scale datasets of people's media use behaviors are increasingly accessible (including the ones we used in this study), as is mounting research interest in understanding media use behaviors (e.g., a recent call for digital trace studies in *Computational Communication Research*). These datasets contain a vast trove of data regarding people's media use and behavior including information about what media content/platform people choose and when people made the choice. Compared with traditional explanatory studies, predictive modeling is usually advantageous in modeling these complex behavior-generating processes.

At the same time, building highly accurate predictive models is a core goal of computer science and media and entertainment industries. Ideally, media research should help inform these efforts. However, that is not really the current situation. Currently, most models predicting media selection in the machine learning industry (e.g., recommendation systems) are highly complex black-box models with a large number of parameters, such as deep neural networks, neural collaborative filtering models, recurrent neural networks, or transformer models. These models take enormous media content features and large-scale media consumption data, and generate highly accurate predictions of media content preferences or behaviors, but with little insight on how the model works or why it works. Consequently, issues emerge due to the opaque and complex nature of these models. For instance, the high complexity of these predictive models results in computing difficulties in model training, especially with large-scale datasets. Moreover, these models lose their generalizability because the parameters are often overfitted for specific media consumption behaviors and, therefore, can only be deployed for domain-specific purposes. For instance, a model trained to predict movie selection on Netflix can hardly be used to predict video consumption behaviors on YouTube. Even though the underlying media selection processes are very similar to each other.

These explanatory deficits in predictive models can be addressed by taking insights from cognition and communication theories and by bridging the gap between explanatory methods and predictive methods. In fact, recent advances in recommendation systems have started to notice the importance of understanding explanatory mechanisms of media selection behaviors. For instance, studies have found that the knowledge of the media users' surprise and curiosity can help to build a curiosity-driven recommendation system. Importantly, this model achieves better performance compared to models that do not account for individual user personality characteristics (Al-Doulat, 2018; Shrestha et al., 2020). Similarly, evidence suggests that, when fit with information about emotional features of media content and media user emotional state, recommendation systems can produce more accurate and dynamic music recommendation effectiveness (Moscato et al., 2021). These approaches take inspiration from explanatory studies, and lean into the construction of more elegant predictive models, which results in better model performance.

In summary, there exists a concurrent communicative gap between explanatory studies and predictive studies, which share a similar research interest in investigating media selection behaviors. We argue that both explanatory research and predictive research can benefit by bridging these two separate lines of research focuses and methods, in an age of information where media consumption data are highly accessible to researchers.

Benefits of Prediction in Media Research

Before we discuss the benefits of predictive models, we want to caution readers that we are not arguing that explanation is not important anymore. Instead, we believe that explanation is and should always be the central goal of scientific media research. But, predictive modeling can and should complement existing explanatory methods. Why? Combining the two will help researchers build and test better theory.

First, predictive modeling can help media researchers to improve the practical implications of their scientific findings (Lin, 2015). Current behavioral sciences are deficient in predicting human behaviors (Yarkoni & Westfall, 2017). Developing predictive models of empirical behavior is of theoretical importance as it serves as a way to verify the extent to which a theory truly contributes to advancing knowledge. Communication and media research also values practical contributions (Krcmar et al., 2016) and predictive modeling of people's media selection has real-world industry applications (e.g., recommendation systems). Recommender systems help improve people's lives by introducing people to media they are likely to enjoy. If social scientific contributions can be used to enhance people's lives, and predictive modeling helps in this endeavor, then this should be considered a top priority for the discipline (Lin, 2015).

Second, predictions can help studies avoid replication crises (Open Science Collaboration, 2015) and p-hacking issues (Simmons et al., 2011) through the collection of large-scale datasets, reducing model variance, and controlling parameter overfitting (Yarkoni & Westfall, 2017). Evaluating predictive metrics restricts the researchers' degree of freedom to change analytical approaches (Yarkoni & Westfall, 2017) and flexibility to accommodate explanations after obtaining unsatisfying results (Szollosi & Donkin, 2021). This is because even though p-hacking practices, such as changing model specification after data collection, could increase explanatory power, it does not benefit predictive power because it induces model overfitting issues.

Third, predictions provide principled methods, i.e., measuring prediction accuracy, to distinguish, compare and integrate different theories (Hofman et al., 2021), which guides the falsification or updating of theories. As an example, Peterson et al. (2021) evaluated the existing decision-making theories based on the predictive performance of the corresponding model on a large-scale experimental dataset. They identified key problems in the theory building of previous decision-making research and pointed out new directions of studies to consider the context of decisions. Combined with formal modeling methods (Smaldino, 2017), researchers will be encouraged to develop integrated theoretical frameworks that generate hypotheses across diverse domains (Muthukrishna & Henrich, 2019). Especially media selection research, which is beginning to integrate with the decision-making literature (Fisher & Hamilton, 2021).

A Note On Effect Size

There is an increasing discussion about the distinction between significant and substantial (or practical or meaningful) effect sizes (Cumming, 2013; Funder & Ozer, 2019). One key element of this discussion asks: are small but significant effect sizes meaningful, and if so, under what circumstances? In response to this question, researchers are now called on to theoretically specify the smallest effect size of interest for NHST (SESOI; Lakens et al., 2018; Weber and Popova, 2012) or region of practical equivalence for Bayesian analyses (ROPE; Kruschke, 2013). Exactly what the SESOI or ROPE is depends, in part, on if an effect size is understood to be cumulative, or not[9]. Small non-cumulative effects may not be substantial, whereas small but cumulative effects may be substantial.

[9] Cigarette smoking is an often-used example to demonstrate this point. The effect of smoking just one cigarette on a healthy individual's mortality is very small and therefore, not substantial. However, the effect of continuously smoking cigarettes across the lifespan is cumulative such that smoking has a substantial effect on morality.

Our project speaks to this discussion. Existing meta-analytic work shows that empirical media-selection accuracy varies between r = 0.07–0.24 (Rains et al., 2018). Our own models of static media selection in this study also show small empirical accuracies (range = 8.9%–15.3%; Figure 1) when temporal dependencies in the selection data are not modeled. These small accuracies hold, even when complex black-box models that should maximize accuracy are employed. What are we to make of this effect? Is there any evidence of a cumulative effect?

Our sequential models offer one way of interpreting the effect via comparison. These models show that, even when only the temporal dependency in the data is accounted for, model accuracy can be improved, and dramatically so (range 9.6%–21.2%; Figures 4 and 5). This means that the prediction accuracy achieved solely by accounting for the previously selected media choice (a type of cumulative effect) exceeds both theory-driven and black-box models of static selection. In summary, by modeling temporal dependencies in our data, we clearly show medium-to-large cumulative effects, even in a literature characterized by relatively small static effects. A best case scenario would combine theoretical variables of interest with temporal dependency data. To our knowledge, no such dataset exists publicly. We are currently working to generate such a dataset for evaluating media selection prediction and explanation.

Limitations

There are a few limitations in our project. First, our empirical studies recognized high potential predictability and MC prediction accuracy for two large-scale media selection datasets for music listening and web browsing behaviors. We expect our findings can be generalized to other types of media selection behaviors, such as movie selection, video selection or news selection. However, the actual generalizability of our findings is still unknown without empirical tests of other types of media selection. Therefore, we encourage future researchers to consider collecting large-scale media selection datasets and applying predictive modeling methods to verify the generalizability of our findings.

Second, we recognize limitations regarding potential influences that come from algorithmic recommendation systems. Media consumption is biased by music or video streaming services or the web ranking systems of search engines, and these biases are difficult to measure without an explicit knowledge of the recommendation system algorithms. Admittedly, we do not know to which extent we are predicting people's autonomous media choices or recommendation systems' choices. Surely we cannot assume that an individual's media selection is free of any external influence. However, we cannot also consider media users as robots who are fully dependent on recommenders' decisions, either. Subjects have the freedom to make higher level decisions and adjust their own media consumption trajectories. Consider a driver that relies on digital maps recommending driving routes from home to working locations. The recommendation system might influence the driver's driving traces about where and when to make a turn, but ultimately the driver is deciding the destinations and whether or not to accept or refuse recommendations. Similarly, media users make free decisions to select lower-ranked media content, skip unfavored video/music, or to jump out of recommended sessions. Therefore, investigating the predictability of empirical media consumption trajectories generated by this hybrid decision-process between media users and the recommendation systems is an important step to initiate the line of research explaining people's autonomous decision mechanisms and their underlying motivation under algorithmic influences (Kleinberg et al., 2022).

Moreover, the current study offers empirical implications for studies interested in real-life media selection behaviors. Media scholars, or generally social scientists, need to accept the fact that nowadays people's behaviors are highly influenced by algorithms, and discover ways of studying algorithmic influence (e.g., Hilbert et al., 2019). Ignoring this algorithmic influence or subtracting human behaviors from environmental influences might induce unrealistic inferences or predictions for real-life behaviors, which is a deviation from one of the core goals of social sciences, that is to understand human behaviors or social dynamics in the real world.

Finally, and unlike other published projects investigating sequential behavior in non-media use contexts, we found that the order one MC model is incapable of providing sufficient prediction accuracy relative to the potential predictability for media behaviors, and the current study is limited to offer a satisfactory explanation for why. Indeed, the MC model is a simple data-driven predictive model which only uses one previously consumed media content to predict one future media content. Possible additional factors for future studies to consider include: (1) people's individual differences and temporary mental states (e.g., mood); (2) the affective or cognitive features of media content (e.g., emotion, novelty, morality, etc.); (3) a reinforcement learning decision mechanism, where people adjust their future media selection based on feedback of previous media choices in a way such that people reinforce high-rewarding media choices and avoid low-rewarding media choices. We anticipate these directions will help future researchers to improve the prediction accuracy and help us better understand media behaviors (for an extended discussion, see Gong and Huskey, in press).

Conclusion

How and why people select media has been a major source of theoretical inquiry since the 1940s. Numerous theoretical accounts exist including, but not limited to, Mood Management Theory (Zillmann, 1988), selective exposure (Knobloch-Westerwick, 2014), uses and gratifications (Rubin, 2009), and so on (for a review, see Hartmann, 2009). Typically, media selection theories explain rather small effect sizes (Rains et al., 2018). Said differently, most of our current theories have very low explanatory accuracy and therefore they also have low prediction accuracy. Our current results suggest that it should be possible, in principle, for theories of media selection to explain substantial variance in people's media selection behaviors. Our results point to a promising path forward. Most media selection theories are primarily focused on individual and media characteristics and theorize that selection is often driven by an interaction between the two.

In our sequential selection studies, even though we know nothing about the individual, and little about the media (beyond listening frequency and listening order), we show that it is nevertheless possible, in principle to predict 60–80% of people's media selection behavior. As the old adage goes, past behavior is a strong predictor of future behavior. Our research shows that media researchers would do well to remember that in their theorizing. With that said, one limitation of our approach is that it is rather silent when it comes to explaining *why* these particular patterns exist. This represents a fundamental challenge to media researchers; that is, developing theoretical *why* explanations that offer sufficient explanatory power to capture the high predictability (in principle) of media selection.

References

Akaike, H. (1998). Information Theory and an Extension of the Maximum Likelihood Principle. In E. Parzen, K. Tanabe, & G. Kitagawa (Eds.), *Selected Papers of Hirotugu Akaike* (pp. 199–213). Springer. https://doi.org/10.1007/978-1-4612-1694-0_15

Al-Doulat, A. (2018). Surprise and Curiosity in A Recommender System. *2018 IEEE/ACS 15th International Conference on Computer Systems and Applications (AICCSA)*, 1–2. https://doi.org/10.1109/AICCSA.2018.8612897

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bialek, W. (2022). On the dimensionality of behavior. *Proceedings of the National Academy of Sciences*, *119*(18), e2021860119. https://doi.org/10.1073/pnas.2021860119

Bowman, N. D., & Keene, J. R. (2018). A Layered Framework for Considering Open Science Practices. *Communication Research Reports*, *35*(4), 363–372. https://doi.org/10.1080/08824096.2018.1513273

Burnham, K. P., & Anderson, D. R. (2004). Multimodel Inference: Understanding AIC and BIC in Model Selection. *Sociological Methods & Research*, *33*(2), 261–304. https://doi.org/10.1177/0049124104268644

Celma, Ò. (2010). Music Recommendation. In Ò. Celma (Ed.), *Music Recommendation and Discovery: The Long Tail, Long Fail, and Long Play in the Digital Music Space* (pp. 43–85). Springer. https://doi.org/10.1007/978-3-642-13287-2_3

Covington, P., Adams, J., & Sargin, E. (2016). Deep Neural Networks for YouTube Recommendations. *Proceedings of the 10th ACM Conference on Recommender Systems*, 191–198. https://doi.org/10.1145/2959100.2959190

Cumming, G. (2013). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis*. Routledge.

Dienlin, T., Johannes, N., Bowman, N. D., Masur, P. K., Engesser, S., Kümpel, A. S., Lukito, J., Bier, L. M., Zhang, R., Johnson, B. K., Huskey, R., Schneider, F. M., Breuer, J., Parry, D. A., Vermeulen, I., Fisher, J. T., Banks, J., Weber, R., Ellis, D. A., … de Vreese, C. (2021). An Agenda for Open Science in Communication. *Journal of Communication*, *71*(1), 1–26. https://doi.org/10.1093/joc/jqz052

Domingos, P. (2012). A few useful things to know about machine learning. *Communications of the ACM*, *55*(10), 78–87. https://doi.org/10.1145/2347736.2347755

Ebersole, C. R., Atherton, O. E., Belanger, A. L., Skulborstad, H. M., Allen, J. M., Banks, J. B., Baranski, E., Bernstein, M. J., Bonfiglio, D. B. V., Boucher, L., Brown, E. R., Budiman, N. I., Cairo, A. H., Capaldi, C. A., Chartier, C. R., Chung, J. M., Cicero, D. C., Coleman, J. A., Conway, J. G., … Nosek, B. A. (2016). Many Labs 3: Evaluating participant pool quality across the academic semester via replication. *Journal of Experimental Social Psychology*, *67*, 68–82. https://doi.org/10.1016/j.jesp.2015.10.012

Fisher, J. T., & Hamilton, K. A. (2021). Integrating Media Selection and Media Effects Using Decision Theory. *Journal of Media Psychology*, *33*(4), 215–225. https://doi.org/10.1027/1864-1105/a000315

Funder, D. C., & Ozer, D. J. (2019). Evaluating effect size in psychological research: Sense and nonsense. *Advances in Methods and Practices in Psychological Science*, *2*(2), 156–168.

Gong, X., & Huskey, R. (in press). Consider the time dimension: Theorizing and formalizing sequential media selection. *Human Communication Research*.

Gong, X., Huskey, R., Eden, A., & Ulusoy, E. (2023). Computationally modeling mood management theory: a drift-diffusion model of people's preferential choice for valence and arousal in media. *Journal of Communication*, *73*(5), 476–493. https://doi.org/10.1093/joc/jqad020

Hartmann, T. (Ed.). (2009, April). *Media Choice: A Theoretical and Empirical Overview*. Routledge. https://doi.org/10.4324/9780203938652

Hastall, M. R. (2009). Informational Utility as Determinant of Media Choices. In *Media Choice*. Routledge.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning*. Springer. https://doi.org/10.1007/978-0-387-84858-7

Hilbert, M., James, R. G., Gil-Lopez, T., Jiang, K., & Zhou, Y. (2018). The Complementary Importance of Static Structure and Temporal Dynamics in Teamwork Communication. *Human Communication Research*, *44*(4), 427–448. https://doi.org/10.1093/hcr/hqy008

Hilbert, M., Liu, B., Luu, J., & Fishbein, J. (2019). Behavioral Experiments With Social Algorithms: An Information Theoretic Approach to Input–Output Conversions. *Communication Methods and Measures*, *13*(4), 267–286. https://doi.org/10.1080/19312458.2019.1620712

Hofman, J. M., Watts, D. J., Athey, S., Garip, F., Griffiths, T. L., Kleinberg, J., Margetts, H., Mullainathan, S., Salganik, M. J., Vazire, S., Vespignani, A., & Yarkoni, T. (2021). Integrating explanation and prediction in computational social science. *Nature*, *595*(7866), 181–188. https://doi.org/10.1038/s41586-021-03659-0

Huskey, R., Bue, A. C., Eden, A., Grall, C., Meshi, D., Prena, K., Schmälzle, R., Scholz, C., Turner, B. O., & Wilcox, S. (2020). Marr's Tri-Level Framework Integrates Biological Explanation Across Communication Subfields. *Journal of Communication*, *70*(3), 356–378. https://doi.org/10.1093/joc/jqaa007

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning – with applications in r* (Vol. 103). Springer. https://doi.org/10.1007/DOI

Katz, E., Blumler, J. G., & Gurevitch, M. (1973). Uses and gratifications research. *Public Opinion Quarterly*, *37*, 509–523. https://doi.org/10.1086/268109

Kleinberg, J., Mullainathan, S., & Raghavan, M. (2022, June). The Challenge of Understanding What Users Want: Inconsistent Preferences and Engagement Optimization. https://doi.org/10.48550/arXiv.2202.11776

Knobloch, S. (2003). Mood Adjustment via Mass Communication. *Journal of Communication*, *53*(2), 233–250. https://doi.org/10.1111/j.1460-2466.2003.tb02588.x

Knobloch-Westerwick, S. (2014, July). *Choice and Preference in Media Use: Advances in Selective Exposure Theory and Research*. Routledge.

Krcmar, M., Ewoldsen, D. R., & Koerner, A. (2016, May). *Communication Science Theory and Research: An Advanced Introduction*. Routledge.

Kruschke, J. K. (2013). Bayesian estimation supersedes the t test. *Journal of Experimental Psychology: General*, *142*(2), 573.

Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. *Advances in Methods and Practices in Psychological Science*, *1*(2), 259–269.

Lang, A. (2013). Discipline in crisis? the shifting paradigm of mass communication research. *Communication Theory*, *23*(1), 10–24. https://doi.org/https://doi.org/10.1111/comt.12000

Lang, A., & Ewoldsen, D. (2010, January). Beyond effects: Conceptualizing communication as dynamic, complex, nonlinear, and fundamental.

LeBel, E. P., McCarthy, R. J., Earp, B. D., Elson, M., & Vanpaemel, W. (2018). A Unified Framework to Quantify the Credibility of Scientific Findings. *Advances in Methods and Practices in Psychological Science*, *1*(3), 389–402. https://doi.org/10.1177/2515245918787489

Levine, T. R., Weber, R., Hullett, C., Park, H. S., & Lindsey, L. L. M. (2008). A critical assessment of null hypothesis significance testing in quantitative communication research. *Human Communication Research*, *34*(2), 171–187. https://doi.org/10.1111/j.1468-2958.2008.00317.x

Levine, T. R., Weber, R., Park, H. S., & Hullett, C. R. (2008). A Communication Researchers' Guide to Null Hypothesis Significance Testing and Alternatives. *Human Communication Research*, *34*(2), 188–209. https://doi.org/10.1111/j.1468-2958.2008.00318.x

Lewis, N. A. (2020). Open Communication Science: A Primer on Why and Some Recommendations for How. *Communication Methods and Measures*, *14*(2), 71–82. https://doi.org/10.1080/19312458.2019.1685660

Lin, J. (2015). On Building Better Mousetraps and Understanding the Human Condition: Reflections on Big Data in the Social Sciences. *The ANNALS of the American Academy of Political and Social Science*, *659*(1), 33–47. https://doi.org/10.1177/0002716215569174

Lindström, B., Bellander, M., Schultner, D. T., Chang, A., Tobler, P. N., & Amodio, D. M. (2021). A computational reward learning account of social media engagement. *Nature Communications*, *12*(1), 1311. https://doi.org/10.1038/s41467-020-19607-x

Lu, X., Wetter, E., Bharti, N., Tatem, A. J., & Bengtsson, L. (2013). Approaching the Limit of Predictability in Human Mobility. *Scientific Reports*, *3*(1), 2923. https://doi.org/10.1038/srep02923

Lydon-Staley, D. M., Zhou, D., Blevins, A. S., Zurn, P., & Bassett, D. S. (2020). Hunters, busybodies and the knowledge network building associated with deprivation curiosity. *Nature Human Behaviour*, 1–10. https://doi.org/10.1038/s41562-020-00985-7

Marek, S., Tervo-Clemmens, B., Calabro, F. J., Montez, D. F., Kay, B. P., Hatoum, A. S., Donohue, M. R., Foran, W., Miller, R. L., Hendrickson, T. J., Malone, S. M., Kandala, S., Feczko, E., Miranda-Dominguez, O., Graham, A. M., Earl, E. A., Perrone, A. J., Cordova, M., Doyle, O., … Dosenbach, N. U. F. (2022). Reproducible brain-wide association studies require thousands of individuals. *Nature*, *603*(7902), 654–660. https://doi.org/10.1038/s41586-022-04492-9

Mares, M.-L., Oliver, M. B., & Cantor, J. (2008). Age Differences in Adults' Emotional Motivations for Exposure to Films. *Media Psychology*, *11*(4), 488–511. https://doi.org/10.1080/15213260802492026
Moscato, V., Picariello, A., & Sperlí, G. (2021). An Emotional Recommender System for Music. *IEEE Intelligent Systems*, *36*(5), 57–68. https://doi.org/10.1109/MIS.2020.3026000
Muthukrishna, M., & Henrich, J. (2019). A problem in theory. *Nature Human Behaviour*, *3*(3), 221–229. https://doi.org/10.1038/s41562-018-0522-1
Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716. https://doi.org/10.1126/science.aac4716
Ophir, E., Nass, C., & Wagner, A. D. (2009). Cognitive control in media multitaskers. *Proceedings of the National Academy of Sciences*, *106*(37), 15583–15587. https://doi.org/10.1073/pnas.0903620106
Pearl, J., & Mackenzie, D. (2018). *The Book of Why: The New Science of Cause and Effect* (1st). Basic Books, Inc.
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, *12*(85), 2825–2830. http://jmlr.org/papers/v12/pedregosa11a.html
Peterson, J. C., Bourgin, D. D., Agrawal, M., Reichman, D., & Griffiths, T. L. (2021). Using large-scale experiments and machine learning to discover theories of human decision-making. *Science*, *372*(6547), 1209–1214. https://doi.org/10.1126/science.abe2629
Rains, S. A., Levine, T. R., & Weber, R. (2018). Sixty years of quantitative communication research summarized: Lessons from 149 meta-analyses. *Annals of the International Communication Association*, *42*(2), 105–124. https://doi.org/10.1080/23808985.2018.1446350
Rubin, A. (2009). Uses-and-gratifications perspective on media effects. *Media Effects: Advances in Theory and Research*.
Schnauber-Stockmann, A., Scharkow, M., & Breuer, J. (2023). Routines and the predictability of day-to-day web use. *Media Psychology*, *26*(3), 229–251. https://doi.org/10.1080/15213269.2022.2121286
Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, *6*(2), 461–464. https://doi.org/10.1214/aos/1176344136
Shifriss, R., Bodner, E., & Palgi, Y. (2015). When you're down and troubled: Views on the regulatory power of music. *Psychology of Music*, *43*(6), 793–807. https://doi.org/10.1177/0305735614540360
Shmueli, G. (2010). To Explain or to Predict? *Statistical Science*, *25*(3), 289–310. https://doi.org/10.1214/10-STS330
Shrestha, P., Zhang, M., Liu, Y., & Ma, S. (2020). Curiosity-inspired Personalized Recommendation, 33–40. https://doi.org/10.1109/WIIAT50758.2020.00010

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. *Psychological Science*, *22*(11), 1359–1366. https://doi.org/10.1177/0956797611417632

Smaldino, P. E. (2017). Models Are Stupid, and We Need More of Them. In *Computational Social Psychology*. Routledge.

Song, C., Qu, Z., Blumm, N., & Barabási, A.-L. (2010). Limits of Predictability in Human Mobility. *Science*, *327*(5968), 1018–1021. https://doi.org/10.1126/science.1177170

Szollosi, A., & Donkin, C. (2021). Arrested Theory Development: The Misguided Distinction Between Exploratory and Confirmatory Research. *Perspectives on Psychological Science*, *16*(4), 717–724. https://doi.org/10.1177/1745691620966796

Tay, J. K., Narasimhan, B., & Hastie, T. (2023). Elastic net regularization paths for all generalized linear models. *Journal of Statistical Software*, *106*(1), 1–31. https://doi.org/10.18637/jss.v106.i01

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., … van Mulbregt, P. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods*, *17*(3), 261–272. https://doi.org/10.1038/s41592-019-0686-2

Vrieze, S. I. (2012). Model selection and psychological theory: A discussion of the differences between the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). *Psychological Methods*, *17*(2), 228–243. https://doi.org/10.1037/a0027127

Wang, Z., Tchernev, J. M., & Solloway, T. (2012). A dynamic longitudinal examination of social media use, needs, and gratifications among college students. *Computers in Human Behavior*, *28*(5), 1829–1839. https://doi.org/10.1016/j.chb.2012.05.001

Weaver, J. B. (1991). Exploring the links between personality and media preferences. *Personality and Individual Differences*, *12*(12), 1293–1299. https://doi.org/10.1016/0191-8869(91)90203-N

Weber, R., & Popova, L. (2012). Testing equivalence in communication research: Theory and application. *Communication Methods and Measures*, *6*(3), 190–213.

Yarkoni, T., & Westfall, J. (2017). Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspectives on psychological science : a journal of the Association for Psychological Science*, *12*(6), 1100–1122. https://doi.org/10.1177/1745691617693393

Zillmann, D. (1988). Mood management through communication choices. *American Behavioral Scientist*, *31*, 327–340. https://doi.org/10.1177/000276488031003005