Proposed chapter

Mark P. Graus* and Bruce Ferwerda

Theory-Grounded User Modeling for Personalized HCI

Abstract: Personalized systems are systems that adapt themselves to meet the inferred needs of individual users. The majority of personalized systems mainly rely on data describing how users interacted with these systems. A common approach is to use historical data to predict users’ future needs, preferences, and behavior to subsequently adapt the system to cater to these predictions. However, this adaptation is often done without leveraging the theoretical understanding between behavior and user traits that can be used to characterize individual users or the relationship between user traits and needs that can be used to adapt the system. Adopting a more theoretical perspective can benefit personalization in two ways: (i) letting systems rely on theory can reduce the need for extensive data-driven analysis, and (ii) interpreting the outcomes of data-driven analysis (such as predictive models) from a theoretical perspective can expand our knowledge about users. However, incorporating theoretical knowledge in personalization brings forth a number of challenges. In this chapter, we review literature that taps into aspects of (i) psychological models from traditional psychological theory that can be used in personalization, (ii) relationships between psychological models and online behavior, (iii) automated inference of psychological models from data, and (iv) how to incorporate psychological models in personalized systems. Finally, we propose a step-by-step approach on how to design personalized systems that take users’ traits into account.

Keywords: personalization, psychological models, cognitive models, psychology, user modeling, theory-driven

*Corresponding author: Mark P. Graus, Department of Human-Technology Interaction, Eindhoven University of Technology, the Netherlands (currently employed at Department of Marketing and Supply Chain Management, School of Business and Economics, Maastricht University, the Netherlands)

Bruce Ferwerda, Department of Computer Science and Informatics, School of Engineering, Jönköping University, Sweden
1 Introduction

Personalization is performed by adapting aspects of systems to match individual users’ needs in order to improve efficiency, effectiveness, and satisfaction. Current personalization strategies are mainly data-driven in the sense that they are based on the way users have been and are interacting with a system, after which the system is dynamically adapted to match inferred user needs. The more theory-driven counterparts of personalization are often designed based on general knowledge about how user traits influence user needs, and how these needs influence the requirements of a system. Systems are adapted to individual users based on a set of rules. Although both strategies are used separately, combining the knowledge gained from both strategies could be used to achieve greater personalization possibilities.

To mitigate personalization challenges, current research has primarily focused on using historical data that describes interaction behavior. Using this data, personalization strategies are developed that predict users’ future interactions. The prediction of these future interactions is often done without leveraging the understanding of the relationship between user behavior and user traits. In other words, predictions are made without considering the root cause of certain behavior that users are showing. A prominent direction using this approach is the field of recommender systems in which historical behavioral data is used to alter the order of items in a catalog (from highest predicted relevance to lowest predicted relevance), with the goal of making users consume more items or helping them to find relevant items more easily [69].

By adopting a more theoretical perspective (often based on psychological literature), the root cause of behavior can be identified and thereby benefit personalization opportunities. Using a theoretical perspective can benefit personalization in two ways: (i) a large body of theoretical work can be used to inform personalized systems without the need of extensive data-driven analysis. For example, research has shown that it might be beneficial to adapt the way course material is presented to match students’ working memory capacity [42]. And (ii) including theory can help to interpret the results gained from the data-driven perspective and thereby meaningfully expand our knowledge about users. For example, research on music players has demonstrated that different types of people base their decisions on what to listen to on different sources of information [29].

Although adopting a more theoretical perspective by considering the relationship between user behavior and user traits has been shown to benefit personalization, this theoretical perspective comes with theoretical and methodological challenges. A first challenge is to identify and measure user traits that
play a role in the needs for personalization (e.g., cognitive style [77], personality [14], susceptibility to persuasive strategies [19]) and to capture these traits in a formal user model. A second challenge is to infer the relevant user traits from interaction behavior (e.g., inferring user preferences from historical ratings, or inferring a person’s personality from the content they share on social media). A third challenge is to identify the aspects of a system that can be altered based on these user traits. In certain cases, this is straightforward (e.g. altering the order of a list of items based on predicted relevance), while in other cases the required alterations can be more intricate and require more thought to implement (e.g. altering the way in which information is presented visually to match a user’s cognitive style).

While the aforementioned challenges are interconnected, they are often addressed in isolation. The current chapter provides an overview of work that relied on user traits for several (system) aspects:

- introduction of psychological models that are currently used in personalization
- psychological models that have been linked to online behavior
- automatic inference of psychological models from behavioral data
- incorporating psychological models in personalized systems or systems for personalization

The literature discussed throughout the chapter can serve as starting points for theory-grounded personalization in certain applications (e.g., e-learning, recommendations) and content domains (e.g., movies, music). Finally, the chapter concludes with a blueprint for designing personalized systems that take user traits into consideration.

2 Psychological Models in Personalization

Psychological models serve to explain how aspects of the environment influence human behavior and cognition. Since these models provide information on how people react to their surroundings, they can also be used to anticipate how people will react to aspects of technological systems and can thus provide insight in people’s needs in technological contexts. The proposition to use psychological models for personalization is not a new concept. Rich [89] already proposed in 1979 the use of psychological stereotypes for personalizing digital systems. However, the current abundance of available user data have made personalization strategies adopt more data-driven approaches and move away from incorporating theoretical knowledge. While the availability of user data obviously benefits
data-driven approaches, there are opportunities for theory-driven approaches as well to exploit the available data (e.g., the implicit acquisition of user traits). In the following section we will lay out different models that are currently used in personalization. We will then continue with providing an overview of prior research that has focused on the relationship between psychological models and online behaviors, then continue with work that has looked at the automated inference of psychological models, and the discuss work that have been personalizing systems based on psychological models.

2.1 Personality

Personality is a long lasting research area in psychology [2]. Personality is considered to reflect behavior through the coherent patterning of affect, cognition, and desires. Aside from this patterning, personality has shown to be a stable construct over time [63]. Through the construct of personality, research has aimed to capture observable individual behavioral differences [21]. Traditional personality psychology has established numerous associations between personality and concepts such as happiness, physical and psychological health, spirituality, and identity at an individual level; the quality of relationships with peers, family, and romantic others at an interpersonal level; and occupational choice, satisfaction, and performance, as well as community involvement, criminal activity, and political ideology at a social institutional level (for an overview see [84]).

Different models have been developed to express personality of people. The most commonly used model is the five-factor model (FFM; mostly used in academic research). The FFM found its roots in the lexical hypothesis, which proposes that personality traits and differences that are the most important and relevant to people eventually become a part of their language. Thus the lexical hypothesis relies on the analysis of language to derive personality traits [2]. The notion of the lexical hypothesis was used by Cattell [14] to lay out the foundation of the FFM by identifying 16 distinct factors. Based on the identified 16 factors, Tupes and Christal [107] found recurrences among the factors that resulted in clusters that represent the five personality traits that make up the FFM (see Table 1). The FFM thus describes personality in five factors (also called the big five personality traits): openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism. Different measurements were created to assess the five personality factors of which the big five inventory (BFI: 44-item) [63]

1 Also called the OCEAN model due to the acronym of the five factors
and the ten item personality inventory (TIPI : 10-item) [46] are two commonly used surveys.

<table>
<thead>
<tr>
<th>General Dimensions</th>
<th>Primary Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Openness to Experience</td>
<td>Artistic, curious, imaginative, insightful, original, wide interest</td>
</tr>
<tr>
<td>Conscientiousness</td>
<td>Efficient, organized, planful, reliable, responsible, thorough</td>
</tr>
<tr>
<td>Extraversion</td>
<td>Active, assertive, energetic, enthusiastic, outgoing, talkative</td>
</tr>
<tr>
<td>Agreeableness</td>
<td>Appreciative, forgiving, generous, kind, sympathetic, trusting</td>
</tr>
<tr>
<td>Neuroticism</td>
<td>Anxious, self-pitying, tense, touchy, unstable, worrying</td>
</tr>
</tbody>
</table>

Tab. 1: Five-factor model adopted from John, Donahue, and Kentle [63]

2.2 Cognitive Styles

Cognitive styles refer to the psychological dimensions that determines individuals’ modes of perceiving, remembering, thinking, and problem-solving [56, 77]. Different cognitive styles have been identified to indicate individual processes, such as analytic-holistic and verbal-visual. In an attempt to make sense out of the fuzziness of different kinds of cognitive styles, Miller [78] proposed a hierarchal framework for systematizing cognitive styles (particularly the analytical-holistic dimension) by connecting them to different stages of cognitive processes ² (see Figure 1).

² We acknowledge the existence of basic, higher, and complex cognitive processes. However, since this chapter focuses on personalization strategies, the focus will be put on individual differences (i.e., cognitive styles) rather than the generic cognitive processes (e.g., perception, memory, thought).
To measure general cognitive styles multiple questionnaires have been developed. The cognitive style index (CSI; Hayes and Allinson [56]) is often used. The CSI consists of 38 items and assigns a score along a single dimension ranging from intuitive to analytical. The items are based on the cognitive style dimensions: active-passive, analytic-holistic, and intuitive-systematic. An alternative measurement to the CSI is the cognitive style analysis (CSA; [90]), which assigns scores to the analytic-holistic and verbal-visual dimensions.

Although it is recognized that individual differences exist in general cognitive functioning, their effects are often diluted by overlapping characteristics of humans. Cognitive styles have been shown to be a better predictor of influence for particular situations and tasks rather than general functioning [70].

For example, cognitive styles have been shown to be related to academic achievements among students (see for an overview Coffield et al. [20]). Despite the domain-dependent variations of cognitive styles that adhere to their own measurements (e.g., learning style questionnaire [57] to assess learning styles), studies have shown that correlations between learning styles and cognitive styles exist (see for an overview Allinson and Hayes [1]). In the following sections, we discuss the domain-dependent cognitive styles that are currently used for the purpose of personalization.
2.2.1 Learning Styles

The educational field has given much attention to identifying individual differences based on a subset of cognitive styles, namely: learning styles. Messick [77] discussed the merit of using cognitive styles to characterize people in an educational setting. Related to people having different preferences regarding processing information as described by cognitive styles, people have different preferences for acquiring knowledge, which are captured in learning styles. In applications that have as goal to assist people in learning, learning styles are a logical candidate to base personalization on. Coffield et al. [20] provide an extensive overview of these learning styles, comprising a selection 350 papers from over 3000 references, in which they identify 13 key models of learning styles. Aside from this overview they provide references to surveys to measure learning styles, together with a list of studies in which these surveys are validated. Two notable models are the learning style inventory (LSI) by Kolb [68] and the learning styles questionnaire (LSQ) by Honey and Mumford [57]. The LSI assesses learning styles through a 100-item self-report questionnaire indicating preferences for environment (e.g., temperature), emotional (e.g., persistence), sociological (e.g., working alone or with peers), physical (e.g., modality preferences), and psychological factors (e.g., global-analytical). The LSQ uses an 80-item checklist to assess learning styles on four dimensions: activist, reflector, theorist, and pragmatist.

2.2.2 Personal Styles

Whereas the previously described FFM of personality (see Section 2.1) found its ground in the lexical hypothesis, the Myers-Briggs model is based on cognitive styles. The Myers-Briggs model of personality is commonly used in the consultancy and training world. People’s scores on the Myers-Briggs model is measured by the Myers-Briggs Type Indicator (MBTI) [83], consisting of 50 questions to measure personality types. The MBTI describes a person’s personality across four dimensions:

1. Extraversion-introversion (E vs. I): how a person gets energized
2. Sensing-intuition (S vs. N): how a person takes in information
3. Thinking-feeling (T vs. F): the means a person uses to make decisions
4. Judging-perceiving (J vs. P): the speed with which a person makes decisions

The combinations of these four dimensions results in one of 16 personality types that are based on Jung’s personality theory from the early 1920s [64] (see Figure 2).
2.3 Relationship Between Models

Psychological models all have their foundation in either behavioral or cognitive assessments of people. As behavior and cognition are so interconnected, the different models are expected to be related as well. Busato et al. [10] found distinct correlations between several personality traits and the type of learning styles people adhere to. Zhang [115] showed relationships between personality traits and cognitive styles and found that creativity-generating and more complex
cognitive styles were related to extraversion and openness. Also other models from traditional psychology that have not been used yet for personalization purposes have shown to correlate with personality traits. For example, Greenberg et al. [53] found correlations between personality traits and people’s ‘music sophistication’ (Gold-MSI; [82]); a measurement to indicate the musical expertise of people.

An extensive literature review that summarizes the findings of how cognitive styles correlate with other psychological models is given in Allinson and Hayes [1]. In multiple studies cognitive styles have been shown to correlate with other measures of learning, thinking or teaching, such as the Learning Style Questionnaire [57]. This indicates that models are related and user traits according to one model can be indicative of their traits in another model.

3 The Relationship Between Psychological Traits and Online Behavior

Aside from using psychological traits defined by traditional psychology to explain the root cause of online behavior, there is an uncharted terrain evolving with the advancement of technologies. Especially as technologies are becoming increasingly ubiquitous and pervasive, new ways of interaction become available between technologies and users. These new ways of interaction may reduce how straightforward the relationship between online behavior and traditional psychological traits is. Hence, there is an increased importance to verify to what extent results from traditional psychology still hold in computer-mediated scenarios before implementation. Aside from that, there is a need for critical reflection on the findings of new relationships between psychological traits and online behavior (e.g., the differentiation between correlation and causation) and the implications for implementation of such findings in personalized systems. In this section we lay out the related work that has focused on verifying relationships between behaviors and traits based on results from traditional psychology, and work that has focused on identifying new relationships between psychological traits and online behaviors. Subsequently Section 5 discusses work that focused on incorporating psychological traits in personalized systems.

3.1 Personality

The way we communicate with others is becoming increasingly mediated through technology in the form of social networking sites (SNSs), such as Facebook,
Instagram, and Twitter [24, 108]. Just like personalities are found to be related to many behaviors in the social or physical world, the digital footprint that people leave behind on these SNSs can be a reflection of people’s personalities as well. Factors such as images (e.g., profile pictures), expressions of thoughts (e.g., content postings), and content preferences (e.g., reactions on content) are in general the information that people leave behind digitally. Similar factors in the real world have already shown to consist of information that people use to generate impressions about others [40]. Back et al. [6] showed that the personalities we express online have resemblances with the personalities that we express in the real world. In other words, it seems that how people express themselves online is an extension of their personality-based behavior, preferences, and needs in the real world.

The notion of extended personality has led to different studies that investigated the digital footprint of users in relation to their personality traits. In particular Facebook has received a lot of attention in the search for personality related behaviors. To exemplify the abundance of research on online personalities, we highlight some of the work that has been done. Some of the results that are found indicate a direct interpretation of personality characteristics to certain online behaviors. For example, one of the findings of Ross et al. [93] showed that extroverts on average belong to more Facebook groups, which is argued to relate to the social nature of extroverts, that leverage Facebook as a social tool. Neurotics (less emotionally stable) were found to spend more time on Facebook, allegedly in an attempt to try to make themselves look as attractive as possible [81]. Conscientiousness has been shown to be related to an increased usage of Twitter although this is not the case for Facebook. Hughes et al. [61] explains this finding to the limitation of the characters that can be used in a Tweet, which causes conscientious people to still able to partake in social networking without it becoming a temporal distraction.

Although most research in online behavior has been done in the context of SNSs in which relationships are tried to be found between personality traits and online behavior, research has also been done in other areas focusing on the transferability of personality judgment. For example, Biel, Aran, and Gatica-Perez [9] found that personality impressions can be transferred via video logs (vlogs). An overview of current research on personality related to online behaviors can be found in Table 2.
<table>
<thead>
<tr>
<th>Study</th>
<th>Domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ellison et al. [24]</td>
<td>Facebook</td>
</tr>
<tr>
<td>Valkenburg and Peter [108]</td>
<td>Facebook</td>
</tr>
<tr>
<td>Moore and McElroy [81]</td>
<td>Facebook</td>
</tr>
<tr>
<td>Ross et al. [93]</td>
<td>Facebook</td>
</tr>
<tr>
<td>Back et al. [5]</td>
<td>Facebook</td>
</tr>
<tr>
<td>Seidman [99]</td>
<td>Facebook</td>
</tr>
<tr>
<td>Gosling, Gaddis, Vazire, et al. [48]</td>
<td>Facebook</td>
</tr>
<tr>
<td>Gosling et al. [47]</td>
<td>Facebook</td>
</tr>
<tr>
<td>Rosenberg and Egbert [92]</td>
<td>Facebook</td>
</tr>
<tr>
<td>Bachrach et al. [4]</td>
<td>Facebook</td>
</tr>
<tr>
<td>Carpenter [12]</td>
<td>Facebook</td>
</tr>
<tr>
<td>Skues, Williams, and Wise [101]</td>
<td>Facebook</td>
</tr>
<tr>
<td>Stieger et al. [102]</td>
<td>Facebook</td>
</tr>
<tr>
<td>Lee, Ahn, and Kim [73]</td>
<td>Facebook</td>
</tr>
<tr>
<td>Ljepava et al. [75]</td>
<td>Facebook</td>
</tr>
<tr>
<td>Eftekhar, Fullwood, and Morris [23]</td>
<td>Facebook</td>
</tr>
<tr>
<td>Winter et al. [111]</td>
<td>Facebook</td>
</tr>
<tr>
<td>Chen and Marcus [17]</td>
<td>Facebook</td>
</tr>
<tr>
<td>Jenkins-Guarnieri, Wright, and Hudiburgh [62]</td>
<td>Facebook</td>
</tr>
<tr>
<td>Wu, Chang, and Yuan [113]</td>
<td>Facebook</td>
</tr>
<tr>
<td>Ryan and Xenos [94]</td>
<td>Facebook</td>
</tr>
<tr>
<td>Quercia et al. [88]</td>
<td>Facebook</td>
</tr>
<tr>
<td>Davenport et al. [22]</td>
<td>Facebook &amp; Twitter</td>
</tr>
<tr>
<td>Hughes et al. [61]</td>
<td>Facebook &amp; Twitter</td>
</tr>
<tr>
<td>Qiu et al. [87]</td>
<td>Twitter</td>
</tr>
<tr>
<td>Ferwerda and Tkalcic [34]</td>
<td>Instagram</td>
</tr>
<tr>
<td>Lay and Ferwerda [72]</td>
<td>Instagram</td>
</tr>
<tr>
<td>Schrammeln, Köffel, and Tscheligi [97]</td>
<td>Online communities</td>
</tr>
<tr>
<td>Ferwerda, Tkalcic, and Schedl [35, 36]</td>
<td>Online music listening</td>
</tr>
<tr>
<td>Ferwerda, Schedl, and Tkalcic [29]</td>
<td>Online music listening</td>
</tr>
<tr>
<td>Ferwerda et al. [37]</td>
<td>Online music listening</td>
</tr>
<tr>
<td>Tkalčič et al. [104]</td>
<td>Online music listening</td>
</tr>
<tr>
<td>Marcus, Machilek, and Schütz [76]</td>
<td>Personal website</td>
</tr>
<tr>
<td>Ferwerda et al. [38]</td>
<td>Recommender system</td>
</tr>
<tr>
<td>Chen, Wu, and He [18]</td>
<td>Recommender system</td>
</tr>
<tr>
<td>Hu and Pu [58]</td>
<td>Recommender system</td>
</tr>
<tr>
<td>Golbeck and Norris [43]</td>
<td>Recommender system</td>
</tr>
<tr>
<td>Biel, Aran, and Gatica-Perez [9]</td>
<td>Video logs</td>
</tr>
</tbody>
</table>

**Tab. 2:** An overview of current research on the relationship between personality traits and online behavior.
3.2 Cognitive Styles: Learning Styles

Whereas personality research has mainly focused on the relationship in SNSs, the research on cognitive styles has influenced other application areas, such as online learning environments. The research in this area has mainly been focusing on identifying whether the cognitive styles as known from traditional psychology have the same effects in an online setting. The findings that can be drawn from the research that has been done on cognitive styles and online learning environments are inconclusive. The disposition of the various results sketches the importance to validate the effect of psychological traits in relation to online behavior. For example, Zacharis [114] looked at 161 students in which the difference of learning styles was investigated between online and offline participation of students, but no differences between the two different groups were found. Huang, Lin, and Huang [60] found differences between online and offline students, such as sensing learners (i.e., those who were patient with details and were good at practical work) engaged online more frequently and for a longer duration. Similarly, Wang et al. [110] showed that online achievements are influenced by the learning styles of students.

As mentioned previously, the majority of research on cognitive styles has focused on investigating to what extent understanding of offline learning translates to online learning environments. Research has been conducted that looked at adapting content delivery based on cognitive styles in these environments. Also, the results are mainly inconclusive in whether cognitive styles can explain individual behavior. For example, Graf and Liu [50] identified different navigational strategies based on learning styles, information that can potentially be used to create adaptive interfaces. However, Mitchell, Chen, and Macredie [80] showed that adaptive interfaces based on cognitive styles do not have an advantageous effect on student performance.

One of few studies that investigated differences based on cognitive styles in a different domain than a learning environment is Belk et al. [7]. They found that people with different cognitive styles differ in preferences with regards to CAPTCHA (acronym for “Completely Automated Public Turing test to tell Computers and Humans Apart”) 3. They found that the cognitive processing style (i.e., verbal or visual) plays a role in the speed of the CAPTCHA completion: those that possess a more verbal cognitive style showed a faster completion of textual CAPTCHAs (e.g., text-recognition: deciphering a scrambled text),

---

3 Commonly used as a security defense mechanism to determine whether an entity interacting with a system is a human and not, for example, an automated bot.
whereas those adhering to a visual cognitive style showed a faster completion of visual CAPTCHAs (e.g., image-recognition: finding a matching set of pictures).

4 Automated Inference of Psychological Traits or Styles

By investigating how psychological knowledge from the offline world transfers to online environments, research has turned to personalization opportunities based on psychological traits and/or styles. The research on personalization has not only focused on implementing psychological traits and styles in systems, but also on how to implicitly infer these traits and styles. By being able to implicitly infer relevant psychological traits and styles, personalization strategies based on these traits and styles can be implemented without the use of extensive questionnaires that are normally used to assess psychological models. Although the use of questionnaires has its advantages (e.g., increased validity and reliability), it also has apparent drawbacks (e.g., time consuming and interrupting the flow between user and system). Moreover, the data for implicit inference does not necessarily need to come from the system directly. Thus, inference and implementation of psychological traits/styles can be achieved across different (connected) platforms [11].

4.1 Personality

As personality has shown to be related to online behavior, attempts have been made to infer personality traits from online behavior. Although alls kind of data can be exploited for personality prediction, research has primarily focused on data retrieved from SNSs. Especially Facebook, Twitter, and Instagram have received a lot of attention in attempts to infer personality from users’ information and generated content (see Table 3 for an overview).

Golbeck, Robles, and Turner [44] looked at how language features expressed in Facebook profiles can be used to infer personality traits. They were able to create a personality predictor with mean absolute errors (MAE) between 0.099 and 0.138 (on a normalized 0-1 scale) across the five personality traits. Similarly, Celli, Bruni, and Lepri [15] and Segalin et al. [98] showed that compositions of Facebook profiles can be used to infer users’ personalities. However, these approaches rely on content that people share on their Facebook page. With extensive privacy options available, users may limit the content they share on
their profile. Ferwerda, Schedl, and Tkalcic [30] showed that by the decisions users make with regards to what sections of their Facebook profile they disclose is indicative of their personality information as well. They were able to create a personality predictor with a root-mean-square error (RMSE) between 0.73 and 0.99 for each personality trait (on a 1-5 scale).

Other attempts to infer personality from online behavioral data have been done on Twitter. Quercia et al. [88] looked at traits of Twitter profiles (e.g., number of followers and following) and found that these characteristics could be used to infer personality. Their personality predictor achieved an RSME between 0.69 and 0.88 for each personality trait (on 1-5 scales) with openness to experience as the trait that could be predicted most accurately and extraversion as the least accurately predictable one. Golbeck et al. [45] analyzed language features (e.g., use of punctuations, sentiment) of Twitter feeds and that they could predict personality with mean absolute errors ranging from 0.119 to 0.182 (on a normalized 0-1 scale).

Ferwerda, Schedl, and Tkalcic [31, 32] analyzed the picture sharing social network Instagram. More specifically they investigated how users manipulate the pictures they upload with filters. They found that personality could be inferred from hue saturation valence (HSV) properties of the uploaded pictures.

Skowron et al. [100] combined information of Twitter and Instagram. By using linguistic and meta data from Twitter and linguistic and image data from Instagram, personality prediction could be significantly improved. They were able to achieve an RMSE personality scores between 0.50 and 0.73.

<table>
<thead>
<tr>
<th>Study</th>
<th>Domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Golbeck, Robles, and Turner [44]</td>
<td>Facebook</td>
</tr>
<tr>
<td>Celli, Bruni, and Lepri [15]</td>
<td>Facebook</td>
</tr>
<tr>
<td>Segalin et al. [98]</td>
<td>Facebook</td>
</tr>
<tr>
<td>Ferwerda, Schedl, and Tkalcic [30]</td>
<td>Facebook</td>
</tr>
<tr>
<td>Quercia et al. [88]</td>
<td>Twitter</td>
</tr>
<tr>
<td>Golbeck et al. [45]</td>
<td>Twitter</td>
</tr>
<tr>
<td>Ferwerda, Schedl, and Tkalcic [31, 32]</td>
<td>Instagram</td>
</tr>
<tr>
<td>Ferwerda and Tkalcic [33]</td>
<td>Instagram</td>
</tr>
<tr>
<td>Skowron et al. [100]</td>
<td>Twitter &amp; Instagram</td>
</tr>
</tbody>
</table>

Tab. 3: An overview of current personality predictors
4.2 Cognitive Styles

Although inference of cognitive styles (e.g., learning styles) has mostly been done within a learning environment (see Section 4.2.1), there is a limited amount of work that has focused on general cognitive styles in other domains. One type of applications in which this has been studied is digital libraries. Frias-Martinez, Chen, and Liu [39] investigated to what extent behavior in digital libraries can be used to make inferences about users’ cognitive styles and how these inferences can be used to personalize these digital libraries. They outline the steps to construct a predictive model that infers cognitive styles from clickstreams and show how this can be done successfully. Similarly, Hauser et al. [54] inferred cognitive styles from the way users interacted with an advice tool for mobile phone contracts and showed that incorporating these cognitive styles improved the buying propensity of users. Belk et al. [8] argued that the cognitive style of users (i.e., the way people organize and perceive information) influences their navigational behaviors in Web 2.0 interactive systems. To investigate the effects of cognitive styles, their study consisted of three steps: (i) investigating the relationship between cognitive styles and navigational behaviors, (ii) investigating whether clustering techniques can group users based on their cognitive style, and (iii) investigating which navigational metrics can be used to predict users’ cognitive style.

4.2.1 Learning Styles

Learning styles are thought to be reflected in the way students acquire knowledge. Specifically online learning environments provide students with a variety of ways to learn and the possibility to log the way students behave within the system. As such, they provide large amounts of data that enables the possibility to build models that can infer learning styles. On average the algorithms that are developed to are able to achieve a 66% to 80% precision. For example, Sanders and Bergasa-Suso [95] developed a system that allowed teachers to help their students study. They collected information related to how students used this system and subsequently conducted a study to investigate how well this information could be used to infer students’ learning styles expressed in the Felder-Silverman Learning Styles Model (FSLSM [26]) and measured through surveys. The FSLSM defines learning styles by four dimensions: active/reflective (A/R), sensing/intuitive (S/I), verbal/visual (V/V), and sequential/global (S/G). Sanders and Bergasa-Suso [95] were able to make significantly better predictions over naive best guesses, which indicates that the way students interact with a learning environment can indeed be used to infer their learning styles.
Similarly, García et al. [41] used a Bayesian Network to infer students’ learning styles as expressed by the FSLSM from how intensively the students interacted with the different elements (e.g., chat, mail, revising exam questions) in the learning system. They found that they could predict students’ learning styles with around 77% accuracy, permitted that the students had prior experience with online learning systems. An overview of other current research on inference of the FSLSM dimensions can be found in Table 4.

<table>
<thead>
<tr>
<th>Study</th>
<th>Algorithm</th>
<th>A/R</th>
<th>S/I</th>
<th>V/V</th>
<th>S/G</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cha et al. [16]</td>
<td>Decision tree</td>
<td>66.70%</td>
<td>77.80%</td>
<td>100%</td>
<td>71.40%</td>
</tr>
<tr>
<td>García et al. [41]</td>
<td>Bayesian network</td>
<td>58%</td>
<td>77%</td>
<td>N/A</td>
<td>63%</td>
</tr>
<tr>
<td>Graf and Liu [51]</td>
<td>Rule-based</td>
<td>79%</td>
<td>77%</td>
<td>77%</td>
<td>73%</td>
</tr>
<tr>
<td>Latham et al. [71]</td>
<td>Rule-based</td>
<td>86%</td>
<td>75%</td>
<td>83%</td>
<td>72%</td>
</tr>
<tr>
<td>Özpolat and Akar [85]</td>
<td>NB tree classification</td>
<td>70%</td>
<td>73.30%</td>
<td>73.30%</td>
<td>53.30%</td>
</tr>
<tr>
<td>Villaverde, Godoy, and Amandi [109]</td>
<td>Artificial neural network</td>
<td>69.30%</td>
<td>69.30%</td>
<td>N/A</td>
<td>69.30%</td>
</tr>
<tr>
<td></td>
<td>Hidden Markov model</td>
<td>66.70%</td>
<td>77.80%</td>
<td>85.70%</td>
<td>85.70%</td>
</tr>
</tbody>
</table>

Tab. 4: An overview of learning style inferences based on the FSLSM by García et al. [41]. Percentages are represent reported precision measures for the four dimensions of the FSLSM: active/reflective (A/R), sensing/intuitive (S/I), verbal/visual (V/V), and sequential/global (S/G).

5 Incorporating Psychological Models in Personalized Systems

In Section 3 we discussed work that focused on identifying differences/similarities between offline and online behaviors as well as work identifying new relationships between psychological traits in online environments. Section 4 presented prior work on the implicit inference of psychological traits and styles from online behavior. Whereas the majority of previous work has focused on identifying relationships between behavior and psychological traits and the implicit inference of said psychological traits, limited work has incorporated psychological traits and styles in personalized systems. In this section we illustrate work that has incorporated psychological traits in systems to create personalized experiences to users.
5.1 Personality

Several studies have shown how incorporating users’ personalities can improve prediction accuracy in the domain of recommender systems. Hu and Pu [59] demonstrated that personality can be used to overcome the new-user cold start problem that arises when no or insufficient information about a user is available to make predictions on. By relying on new users’ personality scores expressed in the FFM, predictions could be made without the need for any additional rating data. Similarly, Fernández-Tobías et al. [27] showed that incorporating personality data in a recommender algorithm allowed for more easily recommending across domains. By considering users’ personality scores in conjunction with ratings across domains (i.e., books, movies and music), they found that they were better able to predict users’ ratings in one domain based on another if they considered personality on top of rating information. Similarly Ferwerda and Schedl [28] and Ferwerda, Schedl, and Tkalcic [30] showed how personality information can be integrated and exploited to improve music recommender systems, whereas Tkalcic, Delic, and Felfernig [103] proposes to incorporate personality to better serve individual needs in group recommendations by taking into account different personality types.

5.2 Cognitive Styles

Personalization strategies based on cognitive styles have been primarily been applied in a learning context (Section 5.2.1 discusses work in a learning context that specifically used cognitive learning styles instead of the general cognitive styles). Karampiperis et al. [65] adapted the cognitive trait model of Lin [74], focusing on the inductive reasoning ability of participants (next to the basic cognitive trait: working memory) to personalize learning environments in which they dynamically adapted the content presentation based on different learners’ navigational behavior. Triantafillou et al. [105] proposed the Adaptive Educational System based on Cognitive Styles system (AES-CS); a system based on the cognitive field dependency and independency style of Witkin et al. [112]. Different navigational organization, amount of control, and navigational support tools were adapted based on this cognitive style. The dynamic adaptation of the interaction elements based on field dependency and independency showed a significant increase in performance then when a static version of the system was presented. In a similar fashion, Tsianos et al. [106] used Riding’s Cognitive Style Analysis [91] to categorize users: imager/verbal and wholist/analyst. Based on the measured categorization they provided users with an adaptive content
presentation and navigational organization. An overview of the studies that applied cognitive styles for adaptation can be found in Table 5.

<table>
<thead>
<tr>
<th>Study</th>
<th>Personalization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Karampiperis et al. [65]</td>
<td>• Adaptive presentation of content</td>
</tr>
<tr>
<td>Triantafillou et al. [105]</td>
<td>• Navigation organization</td>
</tr>
<tr>
<td></td>
<td>• Amount of user control</td>
</tr>
<tr>
<td></td>
<td>• Navigation support tools</td>
</tr>
<tr>
<td>Tsianos et al. [106]</td>
<td>• Adaptive presentation of content</td>
</tr>
<tr>
<td></td>
<td>• Navigation organization</td>
</tr>
</tbody>
</table>

Tab. 5: An overview of studies in which cognitive styles were adopted for personalization

5.2.1 Learning Styles

Knowing a student’s learning styles can be used to alter online learning environments to provide students with information in line with the way they prefer to process information. This is shown to lead to improved learning. Although there is debate about the merit of matching a learning environment to learning styles, there are indications that personalizing learning environments improves the learning effectiveness. Related to learning styles, the working memory capacity has been shown to be a trait that influences what the best learning environments [42, ch. 4]. Students for whom the instructional style matched their learning style scored higher in tests and expressed lower levels of anxiety.

Graf et al. [49] proposed a framework using the FSLSM [26]. The framework consist of adaptation strategies using the sequence as well as the amount of examples and exercises. The same FSLSM [26] was used by Papanikolaou et al. [86] in which interaction preferences were investigated in two educational systems (i.e., FLexi-OLM and INSPIRE) to provide personalized learner support.

Carver, Howard, and Lane [13] used the B.S. Solomon’s Inventory of Learning Styles [25] to create an interface with adaptive content presentation based on the learning style. Milosevic et al. [79] used the LSI [68] in addition to capture preferences, knowledge, goals, and navigational histories of users to adapt the learning environment. An overview of the studies that applied learning styles for adaptation can be found in Table 6.


6 Combining Trait Inference and Trait-Based Personalization

The examples in the previous sections (Sections 4 and 5) all address two subproblems related to using psychological traits in personalization. They either aim to infer user traits or styles from online behavioral data, or they aim to use already measured user traits or styles to improve personalization approaches. These two problems however can be addressed together. Relevant, domain-dependent psychological traits can be identified from psychological theory, measured through surveys to serve as ground-truth and incorporated in user models in personalized system. The current section presents two studies in which this is done and describes the steps in the approach.

6.1 Adapting Comparison Tool Based on Cognitive Styles

Hauser et al. [54] hypothesized that cognitive styles would be traits that influence how users of an online comparison tool would best be helped. They developed and tested a tool to compare cell phone contracts, that relied on two subsystems for personalization. The first subsystem was a Bayesian inference loop, which was used to infer users’ cognitive styles based on what elements of the system they interacted with. The second subsystem was an automatic Gittins loop, which was used to learn how to adapt the content and form of the system to match the cognitive style. There were immediate feedback loops, where the Bayesian inference loop was updated with each click users made and the Gittins loop updated after each time a user finished using the system. If the user exhibited the desired behavior, the systems’ predictions were reinforced. Similarly if the system did not manage to convince the user to make a purchase, the prediction

<table>
<thead>
<tr>
<th>Study</th>
<th>Personalization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carver, Howard, and Lane</td>
<td>• Adaptive presentation of content</td>
</tr>
<tr>
<td>Germanakos and Belk</td>
<td>• Instructional style</td>
</tr>
<tr>
<td>Graf et al.</td>
<td>• Sequence of content</td>
</tr>
<tr>
<td>Milosevic et al.</td>
<td>• Amount of content</td>
</tr>
<tr>
<td>Papanikolaou et al.</td>
<td>• Different course material sequencing</td>
</tr>
<tr>
<td></td>
<td>• Adaptive navigation support (i.e., sequential, global)</td>
</tr>
<tr>
<td></td>
<td>• Adaptive content presentation (i.e., visual/verbal)</td>
</tr>
</tbody>
</table>

Tab. 6: An overview of studies in which learning styles were adopted for personalization
parameters were changed. The study showed that people expressed a higher propensity to buy, which indicates that incorporating the cognitive styles for personalization indeed improved the system. The system was initially not tested in an actual field study but the system was not tested in an actual field study so no conclusions in terms of behavior could be made. In a follow-up study they conducted a field study and found that their approach improves the purchase likelihood in a more natural setting [55].

6.2 Library Personalization Based on Parenting Styles

Graus, Willemsen, and Snijders [52] personalized a digital library for new parents, or parents with children younger than 2 years old. When designing personalized systems for young parents an additional challenge arises from the fact that not only can the users be new to a system, the users themselves are most likely new to the domain of parenting. As such they may not be aware of what type of content is relevant to them and a mismatch may exist between interests and interaction behavior. Information about how to get babies to sleep appears relevant to everyone, but in practice is more relevant for people that raise their kids according to a schedule than it is for parents allow their children to decide when they go to bed. Parents however might beforehand not know whether they want to raise their children according to a schedule or more flexibly and thus incorrectly judge the relevance of content on getting kids to sleep.

Graus, Willemsen, and Snijders [52] compared user experience and behavior in a library personalized based on reading behavior against a library personalized based on survey responses to measure parenting styles. Their study consisted of an initial data collection, that served to gather interaction and survey data to measure parenting styles. They used this data to create personalized relevance predictions which were used to reorder the articles in the library for each individual user. In a second session the same users were reinvited to the now personalized library and data regarding their behavior and user experience was collected. The data showed that personalizing the order of articles based on the survey responses resulted in a higher user experience than relying on reading behavior, despite the fact that the former had lower objective prediction accuracy.
7 Conclusion

Over the years, personalization strategies have been applying different methods. Whereas in the past personalization took a more theoretical approach (e.g., by developing systems that require explicit authoring [67, 89] or by leveraging different psychological models as discussed in this chapter), the abundance of behavioral data and computational power nowadays have caused a shift to a more data-driven perspective (e.g., collaborative filtering [69]). Although, these two different perspectives on personalization are often used in isolation, they could be used together to maximize each other’s potential and mitigate each other’s limitations.

The way users interact with systems can in part be explained through psychological models of the users and incorporating these same psychological models in personalized systems may be a way to improve these systems in terms of effectiveness, efficiency or user satisfaction over considering interaction behavior alone. The current chapter illustrates the possible advantages of combining psychological theory with more naive, data-driven methods for personalization. Doing so leverages the potential of data describing interaction behavior, with the benefits of having interpretable, meaningful user models.

The current chapter presented a number of psychological models that are used in personalization and how they can complement approaches that rely on behavioral data only. Furthermore, the chapter presents a number of ways in which user traits in terms of these models can be inferred from their interaction behavior, and presents ways in which users’ inferred traits can be used to improve how systems are personalized. The benefit of this approach is illustrated with two studies that created a full system by incorporating the inference of psychological models as well as implementing them to personalize systems.

The current chapter considers mainly stable user traits, but more dynamic user characteristics can be considered as well. In principle, any latent trait or characteristic that is related to how users interact with systems and their needs of a system can be used. For example, expertise or experience with a system has been shown to have an effect on how people prefer interacting with a system [66]. Similarly in adaptive hypermedia the inferred level of knowledge dictates what information the system presents [67]. These characteristics are more prone to change, even as a result of using the system, which brings additional challenges with it. As they are related to both the way people interact with a system and what they need from a system, they are logical candidates for being incorporating in user models.
In summary, the chapter demonstrates that adopting a more theoretical perspective by incorporating user traits into personalized systems can lead to improvements of existing systems [27, 59], and that this approach can be used to build new systems [52, 54]. The presented findings warrant future research to focus on incorporating theoretical knowledge about users in personalized systems, instead of solely relying on behavioral data. Apart from providing directions for future research, the literature can be used to generate a blueprint that captures the idea of combining the theory and data-driven perspective on personalization (See Section 7.1).

7.1 Theory-Driven Personalization Blueprint

The approach of incorporating psychological traits into personalization approaches can be formulated into a blueprint. The proposed multidisciplinary approach comprises both theoretical and methodological challenges.

Designing a theory-driven personalized system involves four steps. The first step (Section 7.1.1) is the identification of the right user traits and the right model to measure those. There is a virtually unlimited freedom of choice and making the right choice can be daunting. A model’s suitability depends on the application, domain and the users. After identifying the right model, the second step (Section 7.1.2) consists of collecting data regarding the users’ traits through surveys or through existing inference methods. After collecting this data, the third step (Section 7.1.3) is to find methods to infer the user traits measured in the previous step from natural interaction behavior with the target system. This third step is optional, as in some cases user traits might readily be available. The fourth step (Section 7.1.4) is incorporating the user traits in the personalized systems through formal user models. This section serves to explain these four steps in more detail.

7.1.1 Step 1: Identifying the Right Psychological Model

The first step consists of identifying the right user traits that can be used to improve systems through personalization. Two aspects play a role, firstly the level of generality or specificity. The more specific, the more likely it is that the user traits can be inferred reliably and the more likely that they can be used to improve the system.

Another challenge is the availability of measurement instruments. Regardless of how the user traits will be measured, collecting ground-truth to incorporate
in the personalization is essential. If validated measurement instruments are available, the chance of success is much higher as there is no need to develop and validate a new measurement instrument.

A drawback of generic models such as personality is that they are not necessarily strongly related to what a user needs from a system. More specific models are more likely to be related to users’ needs. If no specific model is available, another possibility is that of developing an instrument to measure relevant user traits, either by designing it from scratch, or by combining existing instruments that measure relevant aspects. This step however then requires designing and validating a survey.

7.1.2 Step 2: Collecting Data Regarding Individual User Traits

After identifying the right model, the second step is collecting data regarding the user traits of the individual users of a system. This can be done in two ways. On the one hand this can be done through surveys as part of the system. Measurement instruments already exist for most psychological models. Collecting data then becomes a matter of administering surveys to users of the system. However, using surveys can be time consuming and interrupts the user from interacting with the system. An alternative way to acquire user data can be done by using inference methods through the use of external data sources (e.g., through the connectedness of single sign-on mechanisms). If traits can be inferred from external data, collecting this data suffices to start personalizing the system without the need to interrupt the interaction flow between the user and the system.

7.1.3 Step 3: Inferring Individuals’ User Traits from Interaction Behavior

The data collected on the user traits in the second step can be used as ground truth to build models that can infer user traits from natural behavior with the system. Section 4 describes for different models how user traits can be inferred.

Hauser et al. [54] performed this step in what they called a priming study. This priming study served to create a baseline model that inferred cognitive styles from clickstream behavior. Later on they relied on a Bayesian inference loop to

---

4 Buttons that allow users to register with or login to a system with accounts from other applications. For example, social networking services: “Login with your Facebook account.”
relate certain aspects of behavior (e.g., what elements users interacted with) to cognitive styles. Similarly, Frias-Martinez, Chen, and Liu [39] trained a neural network to infer cognitive styles from navigation behavior in a digital library.

As mentioned in Step 2, this inference from interaction behavior is in some cases not needed. When using for example single sign-on mechanisms, data from an external system can be used to make inferences as the interconnectedness can indefinitely provide information regarding the systems’ users. A problem that occurs is that not all data from the external system is necessarily readily useful for personalization. By relying on psychological models, this data can be exploited through methods as described by Golbeck et al. [45] and Ferwerda, Schedl, and Tkalcic [31], resulting in information useful for personalization. The use of psychological models thus allows for maximum usage of data as even data that is not directly related to the system of interest can be exploited for the inference. Acquiring data from external sources can mitigate the cold start problem that occurs when users use a system for the first time and no historical interaction behavior is available to base predictions on [96].

### 7.1.4 Step 4: Incorporating User Traits in Personalization Models

The third and final step consists of incorporating these traits in the personalization models. Most straightforwardly this can be done in business rules (similar to as described in Rich [89]. If we know for example that a user has a visual cognitive style, a system might start putting more emphasis on visual information. In a more data-driven way this can be done following Hauser et al. [54], who used a Gittins loop to decide what way of presenting content resulted in the desired behavior (purchasing). In such a way, the system learned how to adapt the content to the users’ cognitive styles.

An additional advantage of incorporating user traits is that the user cold start problem can be (partially) alleviated. If we rely on external data or surveys to infer user traits, the system can be personalized even for users for whom no interaction behavior is available. Hu and Pu [59] did this by using personality information for calculating predictions during the user cold start stage. Personality information made it possible to calculate rating predictions even for users for whom no rating information was available.
References


[34] Bruce Ferwerda and Marko Tkalcic. “You Are What You Post: What the Content of Instagram Pictures Tells About Users’ Personality.” In: Companion Proceedings
REFERENCES


[38] Bruce Ferwerda et al. “The Influence of Users’ Personality Traits on Satisfaction and Attractiveness of Diversified Recommendation Lists.” In: EMPIRE@ RecSys. 2016, pp. 43–47.


[97] Johann Schrammel, Christina Köffel, and Manfred Tscheligi. “Personality traits, usage patterns and information disclosure in online communities.” In: Proceedings of


