Measuring the Impact of Online Personalisation: Past, Present and Future

Markus Zanker

Free University of Bozen-Bolzano, Italy

Laurens Rook

Delft University of Technology, The Netherlands

Dietmar Jannach

University of Klagenfurt, Austria

Abstract

Research on understanding, developing and assessing personalisation systems is spread over multiple disciplines and builds on methodologies and findings from several different research fields and traditions, such as Artificial Intelligence (AI), Machine Learning (ML), Human-Computer Interaction (HCI), and User Modelling based on (applied) social and cognitive psychology. The fields of AI and ML primarily focus on the optimisation of personalisation applications, and concentrate on creating ever more accurate algorithmic decision makers and prediction models. In the fields of HCI and Information Systems, scholars are primarily interested in the phenomena around the use and interaction with personalisation systems, while Cognitive Science (partly) delivers the theoretical underpinnings for the observed effects. The aim and contribution of this work is to put together the pieces about the impact of personalisation and recommendation systems from these different backgrounds in order to formulate a research agenda and provide a perspective on future developments.

Keywords: Web Personalisation, Adaptive Systems, Recommender Systems

1. Introduction

Nowadays, online personalisation encompasses all aspects of individualising the interaction and information content a system exchanges with its users. Different approaches and definitions towards personalisation exist, as summarised by Haiyan and colleagues [1]. The main focus of the present paper is online personalisation, where a system: (1) makes assumptions on an individual's goals, interests and preferences, (2) in order to tailor interaction and content, (3) so as to provide the most relevant user experience. In this perspective, the personalisation process consists of a: (1) learning, (2) matching, and (3) evaluation stage

– as has been proposed, for instance, by [2]. In this paper, we will primarily concentrate on the evaluation stage, and discuss measurement and evaluation approaches to determine the impact of personalisation mechanisms. Specifically, we will describe how personalisation research has evolved over time – from activities aimed at making systems adaptable for better usability to development of ever more accurate prediction mechanisms. Research into personalisation essentially is multidisciplinary in nature – i.e., formed and influenced mainly by academic disciplines such as Artificial Intelligence (AI) and Machine Learning (ML), Human-Computer Interaction (HCI) and Information Systems (IS), and User Modelling based on (applied) social and cognitive psychology. We acknowledge that each of these domains and traditions has made its own respective contribution to personalisation research, and we will elaborate on specific theory contributions, evaluation methodologies and high-impact results in separate subsections.

Also, we will provide an outlook on future opportunities and challenges for research into online personalisation. We will argue that future research endeavours should aim at resolving the disaccord between chasing ever more opportunities for measurement and learning on the one hand, and, on the other hand, raising awareness about the importance of information privacy and regulatory requirements, to improve transparency in the context of a post-GDPR Europe (see Section 4).

2. Dimensions of Personalisation Research

Representative areas, such as adaptive hypertext and hypermedia [3], recommender systems [4, 5], web personalisation [6], information filtering [7], and personalised information retrieval [8], all shaped the concepts of personalising system behaviour and/or its output towards users. As a result, the design space for adaptation and personalisation mechanisms nowadays consists primarily of the following three dimensions:

1. *User Interface*: Different layouts and presentation styles (i.e., adaptation of available control functionality and input elements) are ways to tailor the interaction space between a system and its users. Dating back to the pre-Web era and traditional (Web 1.0) desktop user interfaces (UIs), this research adapts control structures and menu navigation based on system monitoring and assumptions about users' imminent needs [9]. The purpose of this type of adaptivity primarily lies in making users more efficient in using information systems, as has been measured, for instance, through visual search time and required motor movements [10]. Consensus exists on achievable benefits in terms of user satisfaction and performance of adaptive UI elements, but also negative impacts – even of highly accurate adaptation mechanisms – have been identified. Findlater and McGrenere [11], for instance, showed that a user's awareness to new features and the likelihood of using those features later on dramatically reduced performance on new tasks – i.e., personalised adaptation was detrimental to

incidental learning of system features. Interestingly, this finding is called the serendipity problem in content personalisation, and refers to the fact that highly accurate content recommendations may reduce the likelihood to experience unexpected and fortuitous items [12].

2. *Content*: Content traditionally encompasses items or objects such as news articles, products, or media content in the broadest sense of the word; it may however also refer to price tags, service offerings, or highly specific, fine-grained differences in textual wordings. The first research endeavours on selective filtering of information objects for different users date from the late 1950s [13], but the application-oriented research domain on recommender systems (RS) emerged in the early stages of Web 1.0 [14]. Research on RS truly gained momentum in the subsequent Web 2.0 era, when large amounts of (socially) networked data became available [15]. RS produce personalised rankings of large sets of items based on their presumed relevance to recipients. A multitude of evaluation approaches have been proposed for measuring the impact of personalised content, ranging from purely accuracy-driven AI and ML applications, as well as marketing research on customer value and customer churn minimisation, to cognitive and social psychological study of user satisfaction and user engagement.

3. *Interaction process*: The ubiquity of information access opportunities, and the pervasiveness of data collection – also outside of the traditional browser window – in the Web 3.0 era of today marks the potential for novel interaction processes and modalities across devices and environments [15]. Apart from the determination of *which* user interface functionality and *what* content to present, algorithms now also are capable of deciding *when* and *how* to approach users. This additional dimension of personalisation obviously leads to an increase of information privacy concerns, and a discourse on the ethical and societal implications of algorithmic decision making.

As discussed above, personalisation can affect the way software systems are used – as in the case of GUI personalisation – but may also influence the content and information provided and displayed. The recent ubiquity paradigm not only allows personalisation approaches to evolve beyond the content itself, but also to *proactively* determine the point in time plus the situational context, in which personalised information can reach (or more accurately: may be targeted towards) the (unsuspecting) user. It goes without saying that the novel personalisation mechanisms currently being developed fuel the already heated debate on issues and challenges related to invasiveness and user-control.

3. Perspectives from different areas

In this section, we will focus on the particular methodological approaches as well as the theoretical orientations and peculiarities of established research fields such as Machine Learning (ML) / Artificial Intelligence (AI), Human-Computer Interaction (HCI), Information Retrieval (IR) and Information Systems (IS) as

well as Cognitive & Social Psychology that made contributions to online personalisation research. Under the Machine Learning perspective we summarise primarily data-driven and algorithmic research exploiting an offline experimentation methodology, as it is typically applied to develop novel recommendation and retrieval techniques. In contrast, the section on human-centred research methods primarily discusses personalisation research in the fields of HCI and IS as well as on personalised IR that directly investigates the perception of personalisation on users. Next, Section 3.3 takes a complementary viewpoint from industry, while finally Section 3.4 discusses the theoretical underpinnings of the impact of personalisation from the cognitive science perspective.

3.1. Machine Learning perspective on personalisation

Machine Learning (ML) research in the tradition of pragmatic Artificial Intelligence (AI) has developed into a thriving research field. From a highly simplified perspective, the primary focus regarding personalisation applications within ML lies on optimisation – i.e., to create ever more accurate algorithmic decision makers and prediction models [16]. In particular, the recent advancements in learning techniques in neural networks have also triggered research on deep learning-based recommendation systems [17]. Besides such learning techniques, however, the central elements in ML research are data and methodologies to train and learn models or functions from these data to fit as well as possible the portions of unseen data. Recommender systems are generally regarded as the most prominent personalisation applications in the ML field, and all the major online platforms, including Amazon, YouTube, Facebook and Netflix, rely on ML-based RS technology to adapt and personalise presented contents.

The key contributions in ML-based personalisation research include the development of new algorithms, the extension of existing algorithms towards domain-specific aspects, and validation/replication studies based on novel datasets. In order to demonstrate the effectiveness of a newly proposed technique, researchers typically use historical datasets. They either contain (a) explicit ratings provided by a user community to the recommendable items (*explicit feedback*), or (b) a log of recorded user interactions – e.g., purchases, item views or listening events (*implicit feedback*). The subsequent comparison of different algorithms is based on "offline" evaluation procedures and measures common in Machine Learning and Information Retrieval. First, a fraction of the data representing *ground truth* is withheld for later validation (*test set*). Then, a *model* is learned on the remaining data (*training set*), which is used to predict the initially held-out data in the test set. Assessment of the performance of an algorithm, finally, is primarily done by comparing the prediction error and the accuracy of the model with those of the test data. There are basically two different families of measures, depending on whether an algorithm actually predicts a quantitative relevance score (i.e. prediction task) or ranks a set of items based on their presumed relevance (i.e. ranking task). For rating prediction tasks error measures like the Root Mean Squared Error (RMSE) quantify the deviation between predicted and ground truth values. In case of a ranking task, accuracy measures such as Normalised Discounted Cumulative Gain (NDCG),

Mean Average Precision (MAP), or the F-measure, indicate how well a ranked list of items corresponds to an idealised ranking, with all the ground truth content items (i.e., those items that have been actually clicked on, bought or liked) ranked at the highest possible positions. For more details on the methodological underpinnings and the computation of the concrete measures we refer to the following chapters on evaluating recommendation applications [4, 18].

The intuitive assumption underlying the evaluation approach described above is that a more *accurate* prediction of the relevance of an item for the user directly translates into a better user experience. This should be the case, because highly relevant items will be placed at screen positions that will receive more attention from users, whereas less relevant items may be even filtered out. As a consequence, offline experiments in ML tend to approximate the expected *impact* of personalisation by tapping into accuracy measures. However, research on user interface adaptations [11] indicates that the sole reliance on accuracy measures may be insufficient for content personalisation [19]. Accurate prediction (or rather "post-diction" [14]) of historical ratings and/or user actions may not be very helpful for users – as they probably are already aware of the existence of these items, or perhaps regard the resulting list as too similar to items selected previously.

In recent years, researchers have, therefore, come up with a variety of computational measures other than accuracy, so as to capture additional quality aspects, such as list diversity, novelty, and serendipity (surprise) [20, 21] that rely on additional (ground-truth) data such as the content, popularity or freshness of an item. One typical challenge in consideration of such additional measures is that they may merely represent trade-offs. As a point in case, higher levels of serendipity or surprise can typically only be obtained by making a compromise on accuracy. Furthermore, considerable evidence exists that improvements in terms of accuracy either do not translate into a better user experience, or are not perceived as such [22, 23, 24, 25, 26, 27, 28]. These concerns on limited predictive power have also been raised by Netflix researchers [29].

3.2. Human-centred perspectives on personalisation

Particularly scholars in the fields of Human-Computer Interaction (HCI), Information Systems (IS) as well as research on personalised Information Retrieval (IR) share an empirical research methodology and research interest in phenomena concerning computing devices and information systems such as usability and effective system use. Personalisation is seen as a means for minimising a user's costs in accessing information, improving the user experience, and rendering people more efficient in their use of computer devices. However, many of these developments were in fact caused by commercial market forces around end-1990s and beginning of the 2000s. That is, the technical abilities for personalisation developed largely in parallel with the rise of electronic commerce, when web users turned into clients and further transformed into customers. In this context, the concept of customer relationship management (CRM) gained momentum – leading to a new type of relationship-oriented marketing, in which

highly personalised marketing offerings generate high levels of loyalty and engagement among clients [30, 31]. Personalised systems thus provide the technical infrastructure to realise highly targeted one-to-one marketing and individualised CRM strategies at affordable automation costs, capable of turning (online) customers into loyal brand ambassadors [32]. The ultimate goal of personalisation from a commercial standpoint lies in excelling in business metrics so as to build customer value and realise low churn rates.

Tam and Ho [33] did widely recognised work on the impact of (web) personalisation on users' cognitive processes. They measured the attention items or information received as well as users' ability to recall the content later on. The authors manipulated self-referent content (i.e., directly addressing users with their name) and relevant content that matched the task users were up to, and found that the content was perceived as personalised under both conditions – even though only the relevant content was also remembered later on. This clearly showed that personalised communication as such impacts perception, and further indicated that also accuracy in the proposed content is crucial for achieving cognitive processing beyond the initial attention stimulus.

This fundamental issue of a 'placebo effect', in which a message is perceived to be personalised when, in fact, it is not was also researched by Li [34]. They found that users' perception of personalised messages did not necessarily depend on a prior personalisation process of the messages, but on the extent to which the received content matched the receivers' expectations.

Xiao and Benbasat made an inventory of empirical research on the impact of recommendation agents on commercial platforms [35]. They looked into the influence recommendation agents exert on the users' decision making processes, how they influence the outcome of decisions, and the way in which such personalisation mechanisms are perceived and evaluated. The authors discussed how a variety of system characteristics, such as explicit preference elicitation, helps to increase decision quality (or transparency) of a system's reasoning logic. This leads to more trust in the system. Subsequent work on how the impact of different source characteristics can be purposefully exploited so as to create more effective and influential advisory agents (so-called persuasive recommender systems) has, for instance, been summarised here [36, 37]; its theoretical underpinnings will be discussed subsequently in Subsection 3.4.

Interestingly, Knijnenburg and colleagues [38] propose a framework with constructs and questionnaires to guide future user-centred evaluations. Their contribution lies in the identification of an extensive set of measurement constructs grouped into "objective" system aspects, perceptions of the user about the system and the interaction, the situational context as well as personal user characteristics. In addition, the authors offer pointers for relating these aspects to various outcome variables of system usage.

However, controlled laboratory experiments also have their limitations, and need a very careful design and evaluation. Most importantly, the recommendation and decision scenarios are typically artificial in contrast to the field tests performed in industrial environments. Not surprisingly, clear mismatches have been identified between user-centred studies and offline experimentation.

For instance, Ekstrand and colleagues observed that highly accurate algorithms in offline experimentation partly obscured (niche) recommendations in a user study [27]. Likewise, Rosetti and colleagues noticed that users differently ranked the accuracy of algorithms in offline vs. online comparison [22]. Analogously, Ghorab et al. [8] note in their survey on personalised information retrieval techniques the challenges of comparing results derived from different studies. User experiments are therefore crucial for advancing the state-of-the-art, and in order to discover to which extent the computational measures from artificial offline experiments hold in more ecologically valid settings – such as those provided by industry.

3.3. Industry perspective on personalisation

The strong economic interest in the subject of online personalisation beyond the pure *cognitive interest* is the border between academic and industrial evaluation practice. Research labs in industry tend to consent to a three-tiered evaluation approach that includes offline experimentation, exploratory study with beta testers, and so-called A/B testing designs with a representative share of the user base. Corporate research laboratories typically probe the (series of) preliminary results of these approaches before settling on wide-scaled fielding of newly developed algorithm variants. Jannach and Adomavicius [39] proposed a conceptual framework intended to help practitioners (managers and engineers) to align strategic business goals with the most appropriate *metrics* for assessment of the actual impact of RS technology. Furthermore, it should be emphasised that corporate RS service providers hardly ever reveal the true business value of their RS personalisation technologies. Still, some tentative evidence exists of the business impact and common success measures employed in the field:

- *Click-Through Rates:* The Click-Through Rate (CTR) measures the proportion of presented items that actually received a click. The measure is commonly used in online advertising, but also applied in other domains such as news platforms (cf., [40, 41]) with the underlying assumption that more clicks equal more relevant recommendations for the users. Increases of around 35 % in terms of the CTR are not uncommon, when personalised recommendations are field-tested against simple popularity-based techniques [24].

- *Adoption and Conversion Rates:* The CTR is not always the right proxy for assessing the true relevance of recommended and clicked items. Sudden spikes of CTR can also refer to catchy or insufficient link texts that actually misguided and annoyed users. In the online (streaming) media domain, a recommended link is, therefore, only considered a successful hit, if a certain portion of the video or music track was actually played [42]. Auction platforms such as eBay measure how often users make bids on recommended items [43], whereas dating sites measure how many conversations actually followed with a recommended partner [44]. Obviously,

adoption and conversion rates are platform-specific, as a result of which the reported impact of (personalised) recommendations varies across domains. Adoption goes beyond basic CTR, as it considers the use of a service or functionality – like inspecting item details, or watching a video in case of media content. Conversion is typically associated with a commercial or transactional meaning, such as adding to a basket or actually checking out an item on a shopping platform.

- *Sales and Revenues:* Precise revenues and profit margins can be determined for transparent and accessible corporate sales data, such as traditional e-commerce shops. Jannach and Hegelich, for instance, reported a sales increase of around 3-4 % for a mobile commerce platform with game apps [45]. A much more pronounced increase in sales was observed, when a recommender was introduced on an online retail Website and the alternative experimental condition did not involve a recommender at all [46]. Some evidence exists that RS may yield indirect effects on sales, such as when recommendations apparently inspired more purchases in other domains of an online grocery store [47]. The increases in hard business figures are often modest, but represent a considerable return-on-investment compared with the technological investment of the service provider.

- *User Engagement and Behaviour:* User engagement, or more general, the impact on user behaviour, is another commonly used indirect measure. Various studies exist showing that RS leads to more user activity and longer sessions, cf. [24, 48]. In many cases, user engagement measures are application-specific and therefore relate, for instance, to the number of answers provided on a query-answering Website, or to the number of email messages exchanged on a recruiting platform [49, 50].

- *Effects on Sales Distributions:* Some evidence exists that personalisation mechanisms positively impact overall corporate sales distributions. Recommendation agents have been found to improve the quality of matching between items and customers in electronic commerce [35]. Also, several studies observed a smoothing effect on long-tail distributions, which meant that RS actually enabled businesses to create additional business value from niche items [51, 52]. Consistent with this long-tail distribution argument, Netflix measures its "Effective Catalogue Size" [29] to assess to what extent their recommendations help users explore larger portions of their catalogue. The catalogue size is the share of items that actually gets ordered or is viewed by a minimum number of users within a specified time frame. Especially online companies with business models based on flat rate subscription fees exploit the long-tail effect to reduce customer attrition and increase loyalty.

Given the observed effects of personalisation in the previous subsections, we will discuss the impact of personalised content and interaction in the next section from its fundamental behavioural principles.

3.4. Cognitive Science perspective on personalisation

Research into (applied) social and cognitive psychology provides the scientific underpinning of why users may (or may not) appreciate the delivery of personalised services and content in one form or another. Theoretical frameworks such as Similarity Attraction Theory [53] and the classic Belief-Congruence Theory [54] from pre-Web times state that humans are attracted by similar others, and this principle has also been explored in the Web 1.0 era for settings, in which human beings interacted with televisions and personalised computers [55]. Together with the related finding from Cognitive Dissonance Theory [56], that being confronted with facts contradicting personal beliefs and values leads to attitude change, this observation may still hold in the Web 2.0 era that is now rapidly evolving into an ubiquitous Web 3.0.

Over the past decades, *Media Equation Theory* [57] became an influential theoretical perspective among researchers in human-computer interaction to study how people use and interact with computer devices. Derived from a series of explorations of human-computer interaction as a social psychological phenomenon, Media Equation Theory posits that *computers are social actors*, and that social rules from traditional human-to-human interaction also apply to people's interaction with computer devices. Studies reveal that people occasionally praise their desktop computer [58], affiliate with and conform to their computer (when it provides advice) [57], and are sensitive to the desktop computer's praise and flattery [55]. These findings have also been found for other technologies, such as televisions [59, 60].

A significant portion of research in the Media Equation Theory paradigm concerns test of the similarity-attraction effect. Classic research in social psychology [61, 62], such as Belief-Congruence Theory [54, 63, 64], but especially the well-known *attraction paradigm*, provides compelling evidence that people are attracted to others who are similar to them in terms of attitudes, age, personality traits, and many other factors [65, 66]. This so-called similarity-attraction link is "[o]ne of the most robust phenomena in social psychology" [67], (p. 696), and has been replicated across a range of (cross-cultural) situations as well as for a wide range of populations; for review, see [53]; for meta-analysis, see [68, 69]. The attraction paradigm also features prominently in network studies via the principle of *homophily* – i.e., the amount of homogeneity (vs. heterogeneity) of a person's social network in terms of demographic, behavioural and other factors. In a social network, homophily likely yields a larger number of connections with like-minded others, such that "similarity breeds connection" [70], (p. 415).

In Media Equation Theory, the similarity-attraction effect is explored and reported primarily in response to computer-synthesised speech using similarity matching between a user's personality and an extrovert or introvert computer voice, both in spoken and written form [55]. This similarity-attraction effect appears to hold under conditions in which people do not have free choice to select voice-interface modes [71], and in studies on anthropomorphism with regard to (psychophys(iolog)ical) similarity [72].

However, the role of machines has changed dramatically in recent years. Web 2.0 and the currently emerging mobile Web 3.0 era enable the provision

of networked information, accessed through social media [15]. Media Equation Theory seems to apply to human-chatbot interactions on Twitter [73], but scholars increasingly question its viability for present and future times. Advances in ML/AI have made machinery, decision support systems, and mobile devices appearing so (artificially) intelligent that they increasingly outperform humans in problem-solving capabilities – rendering it necessary to reconsider traditional HCI approaches. To deal with the newly arisen *media inequality*, cf. [74], scholars are now starting to turn to state-of-the-art process models in (applied) cognitive and social psychology that are better capable of explaining conditions under which people are implicitly and/or explicitly prone to, feel like, or are motivated to, process information and interact with Web 2.0 and Web 3.0 intelligent machinery, cf. [75]. The implications for future research into personalisation from a cognitive process point of view will be outlined in the next session.

4. What's next? Research Roadmap

In this final section, we will suggest an agenda for future research, and offer an informed discussion on novel measurement approaches – among others involving (neuro)physiological tools and sensors. We will address the necessity and potential of incorporating such measurement instruments in future research endeavours, together with elaborated prospects on other future developments.

4.1. Future perspectives on ML personalisation research

Whether or not small improvements in terms of accuracy – and in terms of other abstract quality measures such as diversity or novelty – actually have a measurable impact in reality is largely under-investigated. Said and Bellogin [76] recently found that implementations of the same algorithms in different recommender system libraries, which were tested against the same measures and according to the same methodology, did not yield the same results. This was due to different data management procedures and slight deviations in the interpretation of the evaluation methodology and algorithmic steps. Similar observations of lacking repeatability, reproducibility and generalisability of experimental results and using weak baselines are also reported from the field of IR [77, 78]. The issue of replication of ML results is therefore a key issue, and should receive much more attention than it currently does in the various RS-related research communities [79].

Generally, we postulate a more holistic view when it comes to assessing the characteristics of personalisation mechanisms. Interestingly, regulatory requirements, such as the General Data Protection Regulation (GDPR) in the European Union, have already begun to put the focus on transparency of algorithmic decision making, on explainability of computed outcomes as well as on ethical considerations in the context of Machine Learning. Therefore, regulatory aspects and questions about the fairness of algorithms will definitely arise more often in the future, and, for instance, more firmly set the research agenda

for improving our understanding on how the recommended information diet of users can become and remain sufficiently diverse [80].

4.2. Future perspectives on Human-centred personalisation research

In recent years, a politically-inspired controversy has materialised on the potentially harmful societal and Web-specific effects of online RS-produced personalisation that facilitates the spreading of misinformation to their target groups [81]. Some people fear that ever more targeted recommendations inevitably put online users in highly specific, profiled grids (or *'filter bubbles'* [82]). The supposed effect of being in a filter bubble would be a dramatic limitation of the diversity of information received and potentially scrutinised, as well as the eventual loss of connection to other social groups in society, cf. [83, 84]. This brings to mind classic pre-Web debates on the societal implications of selective exposure to (news) media, cf. [85], and on group polarisation in (applied) social psychology, cf. [86] – albeit with a twist to the Web 2.0 and Web 3.0 era [87]. Also, it resonates with the observation that higher education systems tend to turn into *'academic tribes'* – i.e., highly inward-oriented disciplines with distinct research practices, conventions, and behaviours, in which actors become resistant to alternative approaches and perspectives [88].

This coincides with criticism from within the Intelligent User Interface community that intelligent systems "dumb down" the user when a portion of the user's cognitive load is offloaded to the system [89]. Intelligent systems may inadvertently harm the development of individual users by offering them an overly selective number of learning opportunities in a specific domain – which may lead to a problematic reduction in the user's breadth of experience. Importantly, first evidence has been generated in the setting of HCI that ongoing online RS personalisation – regardless of the occurrence of a narrowing down in the variety of recommendations a person receives over a longer time period – *does not produce* filter bubbles [90]. Nonetheless, it remains important to conduct further research on the potential 'dark side of online personalisation', and to debunk erroneous opinions and beliefs, whenever necessary. A further line of future research will likely develop at the intersection of understanding the persuasive traits of a system generating recommendations and exploiting individual user differences in order to tailor persuasion strategies to such user traits [91].

4.3. Future industry perspective on personalisation

When it comes to the future industry perspective, the challenges already mentioned in subsection 4.1 will in general also apply to the large-scale industrial application of ML techniques. Thus, practitioners will increasingly struggle to comply with privacy legislation, especially when they seek to provide an enhanced user experience as well. Put differently, what is technically possible to offer to the customer may not be legally allowed, or otherwise be regulatory constrained.

In addition, given the progressing and ageing of currently deployed applications and algorithms, the life-time perspective on personalisation applications

will become increasingly relevant. This may include novel quantitative measurements to understand how metrics on adoption, conversion, user engagement, or the sales distribution itself, actually evolve over time. In particular, industrial data scientists do have perfect opportunities to identify common patterns for different life-cycle stages of personalisation applications over time, should they exist. Nevertheless, it will be hard to perform (let alone, replicate) such research endeavours outside of industrial labs.

4.4. Future perspectives on Cognitive & Psychological Science personalisation research

In pre-Web days, similarity attraction was understood in terms of an affect-based reinforcement model. Similar others were considered to make us feel good (they offered consensual validation for our own viewpoints), whereas dissimilar others made us feel bad (they challenged our attitudes) [92, 53]. Nowadays, similarity attraction is regarded a *cognitive process* determined by information-processing. People judge a person's similarity (or dissimilarity) in attitudes as positive (or negative) information concerning someone else's assumed qualities, cf., [67, 68, 69, 93]. The role of such *cognitive evaluation* (of a person's qualities) on similarity attraction has not yet been studied in human-computer studies, but it will surely have an impact on human interaction with highly personalised and potentially invasive Web 2.0 and Web 3.0 technologies. It stands to reason that scholars in HCI will turn to Cognitive Dissonance Theory [56] to account for unexplained findings on similarity-attraction, cf. [71].

Festinger's [56] famous theory originally stated that people experience cognitive dissonance when displaying behaviours in violation of how they ought to behave, cf. [94, 95], e.g., when contravening normative standards, societal norms and conventions [96, 97]. Attitudes and behaviour would usually be altered in favour of the cognition a person found the hardest to change [98, 99, 56]; see also [100, 101]. The Action-Based Model of Cognitive Dissonance Processes [102] is a recent neuropsychological modification, capable of explaining *why* cognitive dissonance occurs, and *why* it makes people alter initially held beliefs; for review, see [101, 103, 104]. A key insight of the Action-Based Model is that inconsistent action tendencies produce dissonance [101] caused by goal-directed eagerness towards gains, rewards and non-punishment, cf. [105, 106, 107]. Dissonance (or: discrepancy) reduction thus takes place in settings, where prospects of gains, wins, successes, and achievements reign [104]. This insight holds great promise for future research into personalised Web 2.0 and Web 3.0 technologies, as the (partial) removal of competitive persuasion triggers seems to be sufficient in overcoming inconsistent action tendencies due to intrusiveness, infringements of information privacy, and filter bubbles.

4.5. Towards an integrated psychoinformatics / NeuroIS approach

In recent years, scholars in social and cognitive psychology have launched a call for more interdisciplinary research into *psychoinformatics*, aimed at exploration of problems and challenges on the verge of psychological and computer

science [108]. Researchers increasingly turn to mobile technologies, such as the smartphone and the tablet computer, to study human behaviour and mental processes outside of the laboratory [109, 110]. Also an electronically activated recorder (EAR) has been developed to unobtrusively record snippets of natural language, cf. [111, 112, 113]. Likewise, several linguistic inventory and word count (LIWC) software packages have been assembled – and are constantly modified and updated – to accompany speech-sampling methods, and to enable the behavioural analysis of computerised texts derived from Web 2.0 and Web 3.0 social media platforms, cf. [114, 115]. Such endeavours have put the study of 'actual behaviour' more firmly on the research agenda (i.e., they observe real behaviours of real people under real circumstances), and as such move beyond the mere understanding of social and cognitive psychology as "the science of self-reports and finger movements" [116]. Applied to Human-Computer Studies, these and other psychoinformatics tools surely hold promise for ecologically valid future research into the study of actual human behaviour in interaction with Web 2.0 and Web 3.0 technologies, and for the assessment of actual behavioural responses to such technologies in realistic – i.e., real-world – settings.

Interestingly, and in parallel, scholars in the information systems research community have in recent years made comparable endeavours to incorporate behavioural and neuropsychological theorising and measurement techniques to better capture IS phenomena [117] such as system usage or cognitive overload. Like their counterparts in psychological science, IS scholars also are fascinated by the rise of Web 2.0 and Web 3.0, and are fully aware of the necessity to combine the theories, tools and techniques of the behavioural research domain with those of computer science. Some speak of a "computing science social media research" paradigm [118], (p. 26). Others have coined the label NeuroIS [119] instead. In the field of NeuroIS, opportunities especially lie in the development of user interfaces based on fundamental behavioural principles [120]. A cornerstone study using behavioural as well as functional neuroimaging (fMRI) equipment, for instance, found behavioural evidence that high (vs. low) trust (vs. distrust) in the seller on an auction platform impacts price-related buying intentions. In a second stage, also the neural correlates of trust and distrust were mapped, and distinct brain areas associated with trust or distrust were identified [121]. The neural correlates of trust in e-commerce platforms were also explored for gender differences, cf. [122], which seems to suggest that the role of trust in relation to information systems is now well-documented. Apart from fMRI, also less costly neurophysiological tools such as eye-trackers and skin conductance response tools are widely used; cf. [117]. Some have even turned to self-report measures developed within cognitive neuroscience – i.e., the well-known BIS/BAS scales for Behavioural Inhibition and Behavioural Activation [105], which have also been used to validate the Action-Based Model of Cognitive Dissonance Processes discussed above [104]; for an application to online personalisation in the context of recommender systems, see [123]. In spite of its theoretical and methodological diversity [124], however, it is nevertheless intriguing that the field of NeuroIS – in sharp contrast to psychoinformatics in psychological science – seems to move away from real-world applications, and into the laboratory – perhaps to become

the next-generation "science of self-reports and finger movements".

4.6. Putting it all together

Overall, we observe that the vast bulk of personalisation research these days still takes place in the distinct silos of the respective research fields and with separate methodological traditions. Interestingly, this is exploited by the work of Salatino et al. [125], who identify the emergence of novel topics based on intensified collaborations and interactions between such topical silos.

Academic ML research relies largely on offline evaluation measures based on historical data, impact assessments in organisational practice are often based on business value and domain-specific measures. A particularly worrying aspect in the latter case is that offline experiments then may not be very indicative of the success of a recommender in practice. As a future research agenda we, therefore, postulate the need for a more explicit interdisciplinary research approach that would draw a bow from the underpinnings of cognitive and psychological science, to the principled observation and experimentation from HCI, and towards an ecologically valid implementation of results via novel models and efficient machine learning techniques.

In the future, more work is required following common HCI research principles – in particular in the form of user studies – so as to obtain a clearer picture of the impact of content personalisation, and to gain insights for developing algorithms optimising on multiple objectives. The work of [126] could be considered as a first minor step towards this direction. Even though recommendations largely depend on the aggregated opinions, preferences and behaviours of users, this HCI-inspired study showed that the different influential properties of rating summary statistics (such as the total number of ratings, the mean rating value, or the skewness of the distribution) should also be analysed in order to more thoroughly understand how they influence users in their decision making (under a ceteris paribus condition). Drawing from Cognitive Sciences, eye-tracked observations of users' gazes in this interdisciplinary study disclosed that different decision making styles produced significant differences in users' choices. Specifically, users with the tendency to follow maximising behavioural tendencies [127] were more likely to follow compensatory decision making strategies (where different attributes are carefully weighted against each other [128]), and base their decisions on the highest mean rating [126]. The outcome of this interdisciplinary research is currently being exploited as input for developing algorithms, where recommended items are better justified by their rating summaries in the eyes of users. In the end, our understanding of the relation between rating summary statistics and user decision making will thus be deepened thanks to the multiple and integrated perspectives of HCI, User Modelling and AI & ML.

In this context, it is of vital importance to validate the correspondence of abstract computational measures with the users' perceptions, and to find new measures that are more predictive of the adoption of the recommendations. It is surprising to see that the well-established instrument of customer satisfaction surveys from the context of fielded applications does not seem to play a major role in research endeavours in the field of computer science and information

systems. Such surveys about the user experience of fielded personalised services should receive more attention in the future to complement indirect measures of the success of recommenders, and we consider them as a means to obtain impact in practice.

5. Conclusions

In the present paper, we provided an overview of the genesis of research into online personalisation, and described how this led to a large, rich, vibrant, and potentially multidisciplinary body of knowledge spanning Machine Learning (ML) & Artificial Intelligence (AI), Human-Computer Interaction (HCI), Information Retrieval (IR), Information Systems (IS), and User Modelling based on (applied) social and cognitive psychology. Scholars can nowadays choose to study online personalisation from strong accuracy-driven viewpoints and applications, investigate the establishment and maintenance of customer value and customer churn minimisation in electronic commerce settings, or opt for the cognitive and social psychological study of user engagement and satisfaction moderated by explicit and implicit cognitive processes. Future research endeavours in Human-Computer Studies should aim at nurturing the intellectual diversity embodied in these different theoretical and methodological traditions and positions regarding personalisation, while, at the same time, keeping a keen eye for the scientific and societal pitfalls and challenges represented by the ongoing ubiquity of personalisation mechanisms, now and in the future.

References

[1] H. Fan, M. S. Poole, What is personalization? perspectives on the design and implementation of personalization in information systems, Journal of Organizational Computing and Electronic Commerce 16 (3-4) (2006) 179–202.

[2] B. Murthi, S. Sarkar, The role of the management sciences in research on personalization, Management Science 49 (10) (2003) 1344–1362.

[3] P. Brusilovsky, Methods and techniques of adaptive hypermedia, in: Adaptive hypertext and hypermedia, Springer, 1998, pp. 1–43.

[4] D. Jannach, M. Zanker, A. Felfernig, G. Friedrich, Recommender systems: an introduction, Cambridge University Press, 2010.

[5] F. Ricci, L. Rokach, B. Shapira, Recommender systems: introduction and challenges, in: Recommender systems handbook, Springer, 2015, pp. 1–34.

[6] B. Mobasher, R. Cooley, J. Srivastava, Automatic personalization based on web usage mining, Communications of the ACM 43 (8) (2000) 142–151.

[7] P. W. Foltz, S. T. Dumais, Personalized information delivery: An analysis of information filtering methods, Communications of the ACM 35 (12) (1992) 51–60.

[8] M. R. Ghorab, D. Zhou, A. O'connor, V. Wade, Personalised information retrieval: survey and classification, User Modeling and User-Adapted Interaction 23 (4) (2013) 381–443.

[9] S. Greenberg, I. H. Witten, Adaptive personalized interfaces a question of viability, Behaviour & Information Technology 4 (1) (1985) 31–45.

[10] L. Findlater, K. Z. Gajos, Design space and evaluation challenges of adaptive graphical user interfaces, AI Magazine 30 (4) (2009) 68.

[11] L. Findlater, J. McGrenere, Beyond performance: Feature awareness in personalized interfaces, International Journal of Human-Computer Studies 68 (3) (2010) 121–137.

[12] S. M. McNee, J. Riedl, J. A. Konstan, Being accurate is not enough: how accuracy metrics have hurt recommender systems, in: CHI'06 extended abstracts on Human factors in computing systems, ACM, 2006, pp. 1097–1101.

[13] C. Hensley, Selective dissemination of information (sdi): state of the art in may, 1963, in: Proceedings of the May 21-23, 1963, spring joint computer conference, ACM, 1963, pp. 257–262.

[14] D. Jannach, P. Resnick, A. Tuzhilin, M. Zanker, Recommender systems—beyond matrix completion, Communications of the ACM 59 (11) (2016) 94–102.

[15] H. Chen, R. H. Chiang, V. C. Storey, Business intelligence and analytics: From big data to big impact, MIS Quarterly (2012) 1165–1188.

[16] D. Jannach, M. Zanker, M. Ge, M. Gröning, Recommender systems in computer science and information systems–a landscape of research, in: International Conference on Electronic Commerce and Web Technologies, Springer, 2012, pp. 76–87.

[17] S. Zhang, L. Yao, A. Sun, Y. Tay, Deep learning based recommender system: A survey and new perspectives, ACM Computing Surveys (CSUR) 52 (1) (2019) 5.

[18] A. Gunawardana, G. Shani, Evaluating recommender systems, in: Recommender systems handbook, Springer, 2015, pp. 265–308.

[19] S. M. McNee, J. Riedl, J. A. Konstan, Being accurate is not enough: How accuracy metrics have hurt recommender systems, in: CHI '06 Extended Abstracts on Human Factors in Computing Systems, CHI EA '06, 2006, pp. 1097–1101.

[20] C.-N. Ziegler, S. M. McNee, J. A. Konstan, G. Lausen, Improving recommendation lists through topic diversification, in: Proceedings of the 14th international conference on World Wide Web, ACM, 2005, pp. 22–32.

[21] S. Vargas, P. Castells, Rank and relevance in novelty and diversity metrics for recommender systems, in: Proceedings of the fifth ACM conference on Recommender systems, ACM, 2011, pp. 109–116.

[22] M. Rossetti, F. Stella, M. Zanker, Contrasting offline and online results when evaluating recommendation algorithms, in: Proceedings of the 10th ACM conference on recommender systems, ACM, 2016, pp. 31–34.

[23] P. Cremonesi, F. Garzotto, R. Turrin, Investigating the persuasion potential of recommender systems from a quality perspective: An empirical study, ACM Transactions on Interactive Intelligent Systems 2 (2) (2012) 11:1–11:41.

[24] F. Garcin, B. Faltings, O. Donatsch, A. Alazzawi, C. Bruttin, A. Huber, Offline and online evaluation of news recommender systems at swiss-info.ch, in: Proceedings of the 8th ACM Conference on Recommender Systems, RecSys '14, 2014, pp. 169–176.

[25] A. Maksai, F. Garcin, B. Faltings, Predicting online performance of news recommender systems through richer evaluation metrics, in: Proceedings of the 9th ACM Conference on Recommender Systems, ACM, 2015, pp. 179–186.

[26] J. Beel, S. Langer, A comparison of offline evaluations, online evaluations, and user studies in the context of research-paper recommender systems, in: Proc. of the 22nd International Conference on Theory and Practice of Digital Libraries, TPDL '15, 2015, pp. 153–168.

[27] M. D. Ekstrand, F. M. Harper, M. C. Willemsen, J. A. Konstan, User perception of differences in recommender algorithms, in: Proceedings of the 8th ACM Conference on Recommender Systems, RecSys '14, 2014, pp. 161–168.

[28] S. M. McNee, I. Albert, D. Cosley, P. Gopalkrishnan, S. K. Lam, A. M. Rashid, J. A. Konstan, J. Riedl, On the recommending of citations for research papers, in: Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work, CSCW '02, 2002, pp. 116–125.

[29] C. A. Gomez-Uribe, N. Hunt, The Netflix recommender system: Algorithms, business value, and innovation, Transactions on Management Information Systems 6 (4) (2015) 13:1–13:19.

[30] W. Reinartz, M. Krafft, W. D. Hoyer, The customer relationship management process: Its measurement and impact on performance, Journal of marketing research 41 (3) (2004) 293–305.

[31] P. C. Verhoef, W. J. Reinartz, M. Krafft, Customer engagement as a new perspective in customer management, Journal of service research 13 (3) (2010) 247–252.

[32] S. Godin, Permission marketing: Turning strangers into friends and friends into customers, Simon and Schuster, 1999.

[33] K. Y. Tam, S. Y. Ho, Understanding the impact of web personalization on user information processing and decision outcomes, MIS Quarterly 30 (4) (2006) 865–890.

[34] C. Li, When does web-based personalization really work? the distinction between actual personalization and perceived personalization, Computers in human behavior 54 (2016) 25–33.

[35] B. Xiao, I. Benbasat, E-commerce product recommendation agents: use, characteristics, and impact, MIS quarterly 31 (1) (2007) 137–209.

[36] K.-H. Yoo, U. Gretzel, M. Zanker, Persuasive recommender systems: conceptual background and implications, Springer Science & Business Media, 2012.

[37] K.-H. Yoo, U. Gretzel, M. Zanker, Source factors in recommender system credibility evaluation, in: Recommender Systems Handbook, Springer, 2015, pp. 689–714.

[38] B. P. Knijnenburg, M. C. Willemsen, Evaluating recommender systems with user experiments, in: F. Ricci, L. Rokach, B. Shapira (Eds.), Recommender Systems Handbook, Springer US, 2015, pp. 309–352.

[39] D. Jannach, G. Adomavicius, Recommendations with a purpose, in: Proceedings of the 10th ACM Conference on Recommender Systems (RecSys 2016), Boston, Massachusetts, USA, 2016, pp. 7–10. `doi:10.1145/2959100.2959186`.

[40] A. S. Das, M. Datar, A. Garg, S. Rajaram, Google news personalization: Scalable online collaborative filtering, in: Proceedings of the 16th International Conference on World Wide Web, WWW '07, 2007, pp. 271–280.

[41] E. Kirshenbaum, G. Forman, M. Dugan, A live comparison of methods for personalized article recommendation at forbes.com, in: Machine Learning and Knowledge Discovery in Databases, 2012, pp. 51–66.

[42] J. Davidson, B. Liebald, J. Liu, P. Nandy, T. Van Vleet, U. Gargi, S. Gupta, Y. He, M. Lambert, B. Livingston, D. Sampath, The youtube video recommendation system, in: Proceedings of the Fourth ACM Conference on Recommender Systems, RecSys '10, 2010, pp. 293–296.

[43] Y. Chen, J. F. Canny, Recommending ephemeral items at web scale, in: Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '11, 2011, pp. 1013–1022.

[44] W. Wobcke, A. Krzywicki, Y. Sok, X. Cai, M. Bain, P. Compton, A. Mahidadia, A deployed people-to-people recommender system in online dating, AI Magazine 36 (3) (2015) 5–18.

[45] D. Jannach, K. Hegelich, A case study on the effectiveness of recommendations in the mobile internet, in: Proceedings of the 10th ACM Conference on Recommender Systems, RecSys '09, 2009, pp. 205–208.

[46] D. Lee, K. Hosanagar, How Do Recommender Systems Affect Sales Diversity? A Cross-Category Investigation via Randomized Field Experiment, Information Systems Research.

[47] M. B. Dias, D. Locher, M. Li, W. El-Deredy, P. J. Lisboa, The value of personalised recommender systems to e-business: A case study, in: Proceedings of the 2008 ACM Conference on Recommender Systems, RecSys '08, 2008, pp. 291–294.

[48] M. A. Domingues, F. Gouyon, A. M. Jorge, J. P. Leal, J. Vinagre, L. Lemos, M. Sordo, Combining usage and content in an online recommendation system for music in the long tail, International Journal of Multimedia Information Retrieval 2 (1) (2013) 3–13.

[49] Y. Xu, Z. Li, A. Gupta, A. Bugdayci, A. Bhasin, Modeling professional similarity by mining professional career trajectories, in: Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '14, 2014, pp. 1945–1954.

[50] I. Szpektor, Y. Maarek, D. Pelleg, When relevance is not enough: Promoting diversity and freshness in personalized question recommendation, in: Proceedings of the 22nd International Conference on World Wide Web, WWW '13, 2013, pp. 1249–1260.

[51] M. Zanker, M. Bricman, S. Gordea, D. Jannach, M. Jessenitschnig, Persuasive online-selling in quality and taste domains, in: E-Commerce and Web Technologies, EC-Web '06, 2006, pp. 51–60.

[52] R. Lawrence, G. Almasi, V. Kotlyar, M. Viveros, S. Duri, Personalization of supermarket product recommendations, Data Mining and Knowledge Discovery 5 (1) (2001) 11–32.

[53] D. Byrne, An overview (and underview) of research and theory within the attraction paradigm, Journal of Social and Personal Relationships 14 (3) (1997) 417–431.

[54] M. Rokeach, A theory of organization and change within value-attitude systems 1, Journal of Social Issues 24 (1) (1968) 13–33.

[55] C. Nass, K. M. Lee, Does computer-synthesized speech manifest personality? experimental tests of recognition, similarity-attraction, and consistency-attraction., Journal of Experimental Psychology: Applied 7 (3) (2001) 171.

[56] L. Festinger, A theory of cognitive dissonance, Vol. 2, Stanford university press, 1962.

[57] C. Nass, B. Fogg, Y. Moon, Can computers be teammates?, International Journal of Human-Computer Studies 45 (6) (1996) 669–678.

[58] C. Nass, J. Steuer, L. Henriksen, D. C. Dryer, Machines, social attributions, and ethopoeia: Performance assessments of computers subsequent to" self-" or" other-" evaluations, International Journal of Human-Computer Studies 40 (3) (1994) 543–559.

[59] B. Reeves, C. I. Nass, The media equation: How people treat computers, television, and new media like real people and places., Cambridge university press, 1996.

[60] C. Nass, Y. Moon, Machines and mindlessness: Social responses to computers, Journal of social issues 56 (1) (2000) 81–103.

[61] L. Festinger, A theory of social comparison processes, Human relations 7 (2) (1954) 117–140.

[62] T. M. Newcomb, The acquaintance process., Holt, Rinehart & Winston, 1961.

[63] C. A. Insko, R. W. Nacoste, J. L. Moe, Belief congruence and racial discrimination: Review of the evidence and critical evaluation, European Journal of Social Psychology 13 (2) (1983) 153–174.

[64] S. Roccas, S. H. Schwartz, Effects of intergroup similarity on intergroup relations, European Journal of Social Psychology 23 (6) (1993) 581–595.

[65] D. Byrne, Interpersonal attraction and attitude similarity., The Journal of Abnormal and Social Psychology 62 (3) (1961) 713.

[66] D. Byrne, Response to attitude similarity-dissimilarity as a function of affiliation need 1, Journal of Personality 30 (2) (1962) 164–177.

[67] R. M. Montoya, R. S. Horton, On the importance of cognitive evaluation as a determinant of interpersonal attraction., Journal of personality and social psychology 86 (5) (2004) 696.

[68] R. M. Montoya, R. S. Horton, J. Kirchner, Is actual similarity necessary for attraction? a meta-analysis of actual and perceived similarity, Journal of Social and Personal Relationships 25 (6) (2008) 889–922.

[69] R. M. Montoya, R. S. Horton, A meta-analytic investigation of the processes underlying the similarity-attraction effect, Journal of Social and Personal Relationships 30 (1) (2013) 64–94.

[70] M. McPherson, L. Smith-Lovin, J. M. Cook, Birds of a feather: Homophily in social networks, Annual Review of Sociology 27 (1) (2001) 415–444.

[71] K. M. Lee, Y. Jung, C. Nass, Can user choice alter experimental findings in human–computer interaction?: Similarity attraction versus cognitive dissonance in social responses to synthetic speech, Intl. Journal of Human–Computer Interaction 27 (4) (2011) 307–322.

[72] E. J. de Visser, S. S. Monfort, R. McKendrick, M. A. Smith, P. E. McKnight, F. Krueger, R. Parasuraman, Almost human: Anthropomorphism increases trust resilience in cognitive agents., Journal of Experimental Psychology: Applied 22 (3) (2016) 331.

[73] C. Edwards, A. Edwards, P. R. Spence, A. K. Shelton, Is that a bot running the social media feed? testing the differences in perceptions of communication quality for a human agent and a bot agent on twitter, Computers in Human Behavior 33 (2014) 372–376.

[74] Y. Mou, K. Xu, The media inequality: Comparing the initial human-human and human-ai social interactions, Computers in Human Behavior 72 (2017) 432–440.

[75] K. E. Culley, P. Madhavan, A note of caution regarding anthropomorphism in hci agents, Computers in Human Behavior 29 (3) (2013) 577–579.

[76] A. Said, A. Bellogín, Comparative recommender system evaluation: benchmarking recommendation frameworks, in: Proceedings of the 8th ACM Conference on Recommender systems, ACM, 2014, pp. 129–136.

[77] J. Arguello, F. Diaz, J. Lin, A. Trotman, Sigir 2015 workshop on reproducibility, inexplicability, and generalizability of results (rigor), in: Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '15, ACM, New York, NY, USA, 2015, pp. 1147–1148.

[78] N. Ferro, N. Fuhr, A. Rauber, Introduction to the special issue on reproducibility in information retrieval: Evaluation campaigns, collections, and analyses, J. Data and Information Quality 10 (3) (2018) 9:1–9:4.

[79] J. Beel, C. Breitinger, S. Langer, A. Lommatzsch, B. Gipp, Towards reproducibility in recommender-systems research, User Modeling and User-Adapted Interaction 26 (1) (2016) 69–101.

[80] N. Helberger, K. Karppinen, L. D'Acunto, Exposure diversity as a design principle for recommender systems, Information, Communication & Society 21 (2) (2018) 191–207.

[81] M. Fernandez, H. Alani, Online misinformation: Challenges and future directions, in: Companion of the The Web Conference 2018 on The Web Conference 2018, International World Wide Web Conferences Steering Committee, 2018, pp. 595–602.

[82] E. Pariser, The filter bubble: What the Internet is hiding from you, Penguin UK, 2011.

[83] E. Bozdag, J. van den Hoven, Breaking the filter bubble: democracy and design, Ethics and Information Technology 17 (4) (2015) 249–265.

[84] C. R. Sunstein, # Republic: Divided democracy in the age of social media, Princeton University Press, 2018.

[85] D. O. Sears, J. L. Freedman, Selective exposure to information: A critical review, Public Opinion Quarterly 31 (2) (1967) 194–213.

[86] D. G. Myers, H. Lamm, The group polarization phenomenon., Psychological bulletin 83 (4) (1976) 602.

[87] S. Messing, S. J. Westwood, Selective exposure in the age of social media: Endorsements trump partisan source affiliation when selecting news online, Communication Research 41 (8) (2014) 1042–1063.

[88] T. Becher, P. R. Trowler, Academic tribes and territories, SHRE, UK, 2001.

[89] J. Lanier, Agents of alienation, Journal of Consciousness Studies 2 (1) (1995) 76–81.

[90] T. T. Nguyen, P.-M. Hui, F. M. Harper, L. Terveen, J. A. Konstan, Exploring the filter bubble: the effect of using recommender systems on content diversity, in: Proceedings of the 23rd international conference on World wide web, ACM, 2014, pp. 677–686.

[91] M. Kaptein, P. Markopoulos, B. De Ruyter, E. Aarts, Personalizing persuasive technologies: Explicit and implicit personalization using persuasion profiles, International Journal of Human-Computer Studies 77 (2015) 38–51.

[92] D. Byrne, G. L. Clore Jr, Effectance arousal and attraction., Journal of Personality and Social Psychology 6 (4p2) (1967) 1.

[93] R. Singh, R. Ng, E. L. Ong, P. K. Lin, Different mediators for the age, sex, and attitude similarity effects in interpersonal attraction, Basic and Applied Social Psychology 30 (1) (2008) 1–17.

[94] E. Aronson, J. M. Carlsmith, Performance expectancy as a determinant of actual performance., The Journal of Abnormal and Social Psychology 65 (3) (1962) 178.

[95] J. Aronson, G. Cohen, P. R. Nail, Self-affirmation theory: An update and appraisal, in: E. Harmon-Jones, J. Mills (Eds.), Cognitive dissonance theory: Progress on a pivotal theory in social psychology, American Psychological Association, Washington, DC, 1999, pp. 127–147.

[96] J. Cooper, R. H. Fazio, A new look at dissonance, Advances in experimental social psychology 17 (1984) 229–268.

[97] J. Stone, J. Cooper, A self-standards model of cognitive dissonance, Journal of Experimental Social Psychology 37 (3) (2001) 228–243.

[98] L. Festinger, H. Riecken, S. Schachter, When prophecy fails. minneapolis, mn, us (1956).

[99] L. Festinger, J. M. Carlsmith, Cognitive consequences of forced compliance., The Journal of Abnormal and Social Psychology 58 (2) (1959) 203.

[100] J.-L. Beauvois, R.-V. Joule, R. Joule, A radical dissonance theory, Taylor & Francis, 1996.

[101] E. Harmon-Jones, C. Harmon-Jones, Cognitive dissonance theory after 50 years of development, Zeitschrift für Sozialpsychologie 38 (1) (2007) 7–16.

[102] E. Harmon-Jones, Toward an understanding of the motivation underlying dissonance effects: Is the production of aversive consequences necessary?, in: E. Harmon-Jones, J. Mills (Eds.), Cognitive dissonance theory: Progress on a pivotal theory in social psychology, American Psychological Association, Washington, DC, 1999, pp. 71–99.

[103] E. Harmon-Jones, D. M. Amodio, C. Harmon-Jones, Action-based model of dissonance: A review, integration, and expansion of conceptions of cognitive conflict, Advances in Experimental Social Psychology 41 (2009) 119–166.

[104] E. Harmon-Jones, C. Harmon-Jones, N. Levy, An action-based model of cognitive-dissonance processes, Current Directions in Psychological Science 24 (3) (2015) 184–189.

[105] C. S. Carver, T. L. White, Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the bis/bas scales., Journal of personality and social psychology 67 (2) (1994) 319.

[106] P. M. Gollwitzer, Implementation intentions: strong effects of simple plans., American psychologist 54 (7) (1999) 493.

[107] P. M. Gollwitzer, P. Sheeran, Implementation intentions and goal achievement: A meta-analysis of effects and processes, Advances in experimental social psychology 38 (2006) 69–119.

[108] T. Yarkoni, Psychoinformatics: New horizons at the interface of the psychological and computing sciences, Current Directions in Psychological Science 21 (6) (2012) 391–397.

[109] S. Dufau, J. A. Duñabeitia, C. Moret-Tatay, A. McGonigal, D. Peeters, F.-X. Alario, D. A. Balota, M. Brysbaert, M. Carreiras, L. Ferrand, et al., Smart phone, smart science: how the use of smartphones can revolutionize research in cognitive science, PloS one 6 (9) (2011) e24974.

[110] G. Miller, The smartphone psychology manifesto, Perspectives on psychological science 7 (3) (2012) 221–237.

[111] M. R. Mehl, J. W. Pennebaker, D. M. Crow, J. Dabbs, J. H. Price, The electronically activated recorder (ear): A device for sampling naturalistic daily activities and conversations, Behavior Research Methods, Instruments, & Computers 33 (4) (2001) 517–523.

[112] M. R. Mehl, S. Vazire, N. Ramírez-Esparza, R. B. Slatcher, J. W. Pennebaker, Are women really more talkative than men?, Science 317 (5834) (2007) 82–82.

[113] M. R. Mehl, S. Vazire, S. E. Holleran, C. S. Clark, Eavesdropping on happiness: Well-being is related to having less small talk and more substantive conversations, Psychological science 21 (4) (2010) 539–541.

[114] J. W. Pennebaker, M. R. Mehl, K. G. Niederhoffer, Psychological aspects of natural language use: Our words, our selves, Annual review of psychology 54 (1) (2003) 547–577.

[115] Y. R. Tausczik, J. W. Pennebaker, The psychological meaning of words: Liwc and computerized text analysis methods, Journal of language and social psychology 29 (1) (2010) 24–54.

[116] R. F. Baumeister, K. D. Vohs, D. C. Funder, Psychology as the science of self-reports and finger movements: Whatever happened to actual behavior?, Perspectives on Psychological Science 2 (4) (2007) 396–403.

[117] A. Dimoka, F. D. Davis, A. Gupta, P. A. Pavlou, R. D. Banker, A. R. Dennis, A. Ischebeck, G. Müller-Putz, I. Benbasat, D. Gefen, et al., On the use of neurophysiological tools in is research: Developing a research agenda for neurois, MIS quarterly (2012) 679–702.

[118] B. Shneiderman, J. Preece, P. Pirolli, Realizing the value of social media requires innovative computing research, Communications of the ACM 54 (9) (2011) 34–37.

[119] A. Dimoka, P. A. Pavlou, F. D. Davis, Research commentary—neurois: The potential of cognitive neuroscience for information systems research, Information Systems Research 22 (4) (2011) 687–702.

[120] K. Reinecke, A. Bernstein, Knowing what a user likes: A design science approach to interfaces that automatically adapt to culture., Mis Quarterly 37 (2).

[121] A. Dimoka, What does the brain tell us about trust and distrust? evidence from a functional neuroimaging study, Mis Quarterly (2010) 373–396.

[122] R. Riedl, M. Hubert, P. Kenning, Are there neural gender differences in online trust? an fmri study on the perceived trustworthiness of ebay offers, Mis Quarterly 34 (2) (2010) 397–428.

[123] L. Rook, A. Sabic, M. Zanker, Reinforcement sensitivity and engagement in proactive recommendations: Experimental evidence, in: Information Systems and Neuroscience, Springer, 2018, pp. 9–15.

[124] R. Riedl, F. D. Davis, A. R. Hevner, Towards a neurois research methodology: intensifying the discussion on methods, tools, and measurement, Journal of the Association for Information Systems 15 (10) (2014) I.

[125] A. A. Salatino, F. Osborne, E. Motta, How are topics born? understanding the research dynamics preceding the emergence of new areas, PeerJ Computer Science 3 (2017) e119.

[126] L. Coba, L. Rook, M. Zanker, P. Symeonidis, Decision making strategies differ in the presence of collaborative explanations: Two conjoint studies, in: Proceedings of the 24th International Conference on Intelligent User Interfaces, IUI '19, ACM, New York, NY, USA, 2019, pp. 291–302.

[127] B. Schwartz, et al., On the meaning and measurement of maximization, Judgment and Decision making 11 (2) (2016) 126.

[128] J. W. Payne, Task complexity and contingent processing in decision making: An information search and protocol analysis, Organizational behavior and human performance 16 (2) (1976) 366–387.