UNIVERSITY of CALIFORNIA PRESS | Collabra: Psychology

García Alanis, J. C., Strelow, A. E., Dort, M., Christiansen, H., Pinquart, M., & Panitz, C. (2023). Expectation Violations, Expectation Change, and Expectation Persistence: The Scientific Landscape as Revealed by Bibliometric Network Analyses. *Collabra: Psychology*, *9*(1). https://doi.org/10.1525/collabra.73830

Cognitive Psychology

Expectation Violations, Expectation Change, and Expectation Persistence: The Scientific Landscape as Revealed by Bibliometric Network Analyses

José C. García Alanis[1,2][a], Anna E. Strelow[1], Martina Dort[1], Hanna Christiansen[1], Martin Pinquart[1], Christian Panitz[1,3,4]

[1] Department of Psychology, Philipps-Universität Marburg, Marburg, Germany, [2] Department of Psychology, Johannes Gutenberg-Universität, Mainz, Germany, [3] Department of Psychology, Universität Leipzig, Leipzig, Germany, [4] Center for the Study of Emotion and Attention, University of Florida, Gainesville, FL, USA

Keywords: expectation violations, expectation change, expectation persistence, prediction error, bibliometric analysis

https://doi.org/10.1525/collabra.73830

Collabra: Psychology

Vol. 9, Issue 1, 2023

Expectation violations occur when there is a discrepancy between expected and perceived events or experiences. Often, however, expectations persist despite disconfirming evidence. Therefore, research on expectation violations, expectation change, and expectation persistence has been conducted in several fields of psychology with wide-ranging theoretical assumptions and empirical considerations. In the present review, we analysed how these research fields relate to each other using bibliometric network analyses. For this purpose, we conducted a systematic literature search to identify scientific publications on expectation violations, expectation change, and expectation persistence. The literature corpus was then quantitatively analysed using similarity measures that allow a data-driven classification of publications into groups, revealing their conceptual, theoretical, and empirical commonalities. Our results indicate that many influential publications have focused on finding reactivity measures (e.g., brain activation) to the discrepancy experienced between expectations and outcomes. Furthermore, these measures have been used to assess when and to which degree learning and behavioural adaptation (i.e., expectation change) take place. We discuss the potential application of these measures for understanding expectation violations, expectation change, and expectation persistence in more complex settings (e.g., social interaction). The goal of this review was to foster interdisciplinarity in psychology, enabling scientists and practitioners to identify new topics, promising empirical approaches, and previously neglected variables.

Introduction

Expectations are conditional beliefs about the probabilities of future events, experiences, or information (cf. Hoorens, 2012; Panitz et al., 2021). Expectations vary from (more or less) possible to certain (Roese & Sherman, 2007), they can be more or less explicit (i.e., conscious; Proulx & Inzlicht, 2012), and they may focus on heterogeneous contents (cf. Laferton et al., 2017). For instance, individuals can have expectations regarding their behaviour and experience in a potentially demanding situation. However, they may also have expectations regarding the characteristics of the situation itself or others' reactions to one's behaviour. As such, expectations relate to a wide spectrum of anticipatory mechanisms, ranging from basic perceptual and motor functions (de Lange et al., 2018; Yon et al., 2019) to highly elaborated and sometimes stereotypical beliefs about others (Dort, Strelow, Schwinger, et al., 2020; Kotzur & Wagner, 2021). Due to their probabilistic nature, expectations often prove inaccurate, or even entirely erroneous, in the face of disconfirming information. Consequently, expectation violations occur in the case of discrepancy between expected and perceived situational outcomes. It is widely accepted that the processing of expectation violations is necessary for learning and effective interaction with the environment (cf. Gollwitzer et al., 2018; Rief et al., 2015). Furthermore, research indicates that expectation violations are of central relevance for many domains of psychology (Pinquart, Endres, et al., 2021; Pinquart, Rothers, et al., 2021), as they very often enable individuals to detect ongo-

a Correspondence regarding this article should be addressed to José C. García Alanis, Department of Psychology, University of Marburg, Gutenbergstr. 18, 35032, Marburg, Germany. e-mail: jose.alanis@staff.uni-marburg.de

ing deviations from internally represented goals (Roese & Sherman, 2007), adapt goal-directed behaviour (Holroyd & Coles, 2002), and regulate experience (Clark, 2013; Panitz et al., 2021; Pinquart, Endres, et al., 2021) in accordance to situational characteristics.

Different theoretical models have been proposed to explain how organisms cope with expectation violations (cf. Pinquart, Endres, et al., 2021 for review). Typically, the experienced discrepancy between expectation and disconfirming evidence is referred to as *prediction error* (Friston, 2009; Rescorla & Wagner, 1972; Schultz et al., 1997). A large body of research indicates that the degree to which expectation violations trigger compensatory organismic, cognitive, emotional, and motivational responses scales proportionally to the magnitude of the prediction error (Den Ouden et al., 2012; Hajcak, 2012; Holroyd & Coles, 2002; Sambrook & Goslin, 2015). In addition, it has been proposed that one of humans' core motives is to reduce the probability of future expectation violations (Clark, 2013; Proulx & Inzlicht, 2012), as they often evoke aversive emotional states (Proulx et al., 2012), enhance control demands (Shenhav et al., 2013) and might threaten individuals' self-concept (Brandtstädter & Greve, 1994; Korn et al., 2012; Pinquart & Block, 2020). Previous research (Panitz et al., 2021; Rief et al., 2015) proposes that minimising future expectation violations can be achieved by adaptively adjusting expectations as a function of the prediction error, leading to updated versions of those expectations that integrate previous and current expectation-disconfirming information. Evidence for this process is provided in situations in which individuals use performance feedback to proactively recruit control resources and prepare for comparable situations in the future (Bejjani et al., 2020; Cavanagh et al., 2010), but also when the goal-directedness of unfavourable choices is questioned in order to prevent the formation and maintenance of maladaptive habits (Gillan et al., 2015; McKim et al., 2016). In this line of research, particular attention has been paid to event-related transient fluctuations (i.e., Event-Related Potentials - ERPs) in the scalp-recorded electroencephalogram (EEG) following expectation violations, such as the Feedback-Related Negativity (FRN; Miltner et al., 1997) and the Error-Related Negativity (ERN; Gehring et al., 1993)[1]. It is believed that the FRN and the ERN reflect endogenous alarm signals that allow individuals to identify deviations from internally represented goals states (i.e., expectations) and adjust behaviour and/or expectations accordingly (Holroyd & Coles, 2002; Sambrook & Goslin, 2015), thus providing physiological correlates for the processing of expectation violations across different psychological domains (cf. Weinberg et al., 2015) and situational embeddings (cf. García Alanis et al., 2019; Mueller et al., 2014).

Often, however, individuals do not update their expectations after these were violated by situational outcomes (Panitz et al., 2021; Proulx et al., 2012; Roese & Sherman, 2007). Such expectation maintenance can be advantageous when the obtained outcomes could be disregarded as probable noise (Hohwy, 2017), when avoiding or attenuating negative affect after worse-than-expected experiences (Proulx et al., 2012), or when expectations need to be robust because they relate to values and positive beliefs that individuals hold about themselves or the world (Pinquart & Block, 2020). Conversely, maintaining expectations despite disconfirming evidence can also have negative consequences for oneself (e.g., psychopathology; Kube et al., 2017, 2019) or others (e.g., stereotypes; Kotzur & Wagner, 2021). For instance, rigid negative expectations have been identified as one of the core factors involved in the development and chronicity of mental disorders (Craske et al., 2014). However, the precise cognitive (Pinquart, Rothers, et al., 2021) and neural (D'Astolfo & Rief, 2017) mechanisms that underlie coping with expectation violations are still a matter of discussion. Research on coping with expectation violations has been conducted in several fields of psychology (e.g., general, social, or consumer psychology; Pinquart, Endres, et al., 2021) and with different theoretical and empirical foci (e.g., neurophysiological processes: D'Astolfo & Rief, 2017; expectation persistence versus change: Panitz et al., 2021; coping with associated negative emotions: Proulx & Inzlicht, 2012). As there is considerable heterogeneity and often little exchange between these different research domains, there is a need to analyse how different approaches relate to each other and which studies have been most influential in the field. Therefore, a comprehensive overview of the conceptual structure and theoretical commonalities of a growing body of research on expectation violation, change, and persistence is needed to help inform future studies and identify new areas of possible application.

To address these issues, the present study pursued a science mapping approach by implementing bibliometric analysis techniques (Diodato & Gellatly, 2013) to quantitatively analyse the conceptual patterns that emerge in scientific publications on expectation change and maintenance in the context of expectation violations. More precisely, we analysed manuscript keywords of scientific articles as well as their reference lists to compute similarity and network analysis measures, provide a data-driven classification of bibliographic objects into groups, and reveal commonalities and relationships within and between the identified clusters of research (cf. Zupic & Čater, 2015). A bibliometric approach has methodological implications and advantages complementary to other types of review studies (cf. Andersen, 2019; Dort, Strelow, French, et al., 2020). First, a bibliometric review allows a broader perspective because researchers can analyse larger bodies of literature than in purely qualitative reviews, with the latter being often focused or limited in scope and typically based on a few tens of articles. Second, a bibliometric review provides the opportunity to identify higher-order patterns that offer con-

1 or fronto-central negativity ("NE") as in Falkenstein and colleagues (1991).

text and location for more focused and detailed reviews in the broader field. Third, as the concept of expectation violation is relevant for many fields of research, it is likely that a traditional narrative review would miss important, but perhaps underrepresented, clusters of literature. Furthermore, as traditional reviews are grounded in the author's experience and expertise in the field, their assessment of the relationships between clusters of literature might be more subjective. This also makes traditional reviews susceptible to multiple sources of bias, such as the author's degree of specialisation in the field and, more generally, citation bias toward work that is central in the author's field (Mullen & Ramírez, 2006). In contrast, bibliometric analyses follow fully replicable and explicit data gathering and analysis plans, making bibliometric reviews less susceptible to researcher degrees of freedom and leading to more robust objective estimates for evaluating the relationships within and between clusters of literature (cf. Archambault et al., 2009).

The aim of this study was to provide a structural overview of the scientific avenues present in the field of expectation violation research. To the best of our knowledge, this is the first bibliometric review of expectation violation literature. It builds upon existing systematic and meta-analytical reviews focussing on different theoretical models of expectation violation (Pinquart, Endres, et al., 2021), as well as the brain structures (D'Astolfo & Rief, 2017) and personal or situational variables (Pinquart, Rothers, et al., 2021) that support expectation persistence or change. The present study intends to help researchers navigate the field and identify promising and perhaps underdeveloped research domains. By finding links and commonalities in the field of expectation violation research, the present study is also meant to facilitate interdisciplinarity, enabling scientists and practitioners to look outside their field-specific box. Our goal is to provide guidance for the discovery of alternative approaches, helping researchers identify important topics and variables that might have been neglected otherwise.

Method

Literature search

The Social Sciences Citation Index (SSCI), which is available online through the Web of Science (WoS), was selected as the primary data source. It features more than 500 journals related to psychological science. In addition, it covers a wide spectrum of sub-disciplines and interdisciplinary research fields related to human behaviour, cognition, and affect. We included publications focused on the biological and neuroscientific bases of behaviour and research focused on social, personality, emotional, and motivational phenomena, along with possible implications for clinical, developmental, and educational fields of application.

The SSCI was searched on December 16, 2020, using the Boolean search term presented in Table 1. The search was limited to documents published between 1990 and 2020 (December 16) within the WoS category of psychology. The search term was designed to look for the term expectation violation (or variations of the same, such as "violation of expectations" or "violated expectations") in the title, abstract, author keywords, and Keywords Plus ® fields of each document. The search was constrained to documents focussing on the impact of expectation violation on behaviour and experience, i.e., whether the violation of expectations induced changes in expectations or whether behavioural, cognitive, or emotional patterns persisted afterwards. This was achieved by searching for combinations of words with the stems "expect" and "violat" in scientific publications using wildcards (i.e., "*") combined with the NEAR/0 and NEAR/1 Boolean operators in WoS. Thus, the Boolean search term TS=((expect*) NEAR/0 (violat*)) leads to hits in publications in which "expectation violation" and "expectancy violation" were provided as keywords, as they share the same word stems and are not separated by any word (i.e., are NEAR/0). In contrast, ((violat*) NEAR/1 (expect*)) leads to hits in publications in which "violated expectations" was provided as a keyword, but also publications with "violation of expectations" as a keyword, as they share the same word stems and are separated by a maximum of one word (i.e., are NEAR/1). Topic terms closely related to expectation violation were also included in the search (e.g., prediction error). To identify terms closely related to expectation violation, expectation persistence, and expectation change, we first screened literature on theoretical models of expectations and expectation violations for potentially relevant terms (cf. D'Astolfo & Rief, 2017; Gollwitzer et al., 2018; Rief et al., 2015). Second, we analysed project reports from 16 PhD students enrolled in the local Research Training Group on expectation violations (DFG Research Training Group 2271 "Breaking Expectations") who had provided information about their projects (e.g., theoretical background, hypotheses). The final list of search terms (see **Table 1**) was created based on the collected terms and discussion among the authors of the present study.

Data preparation and analysis

Bibliometric analyses allow for the computation of quantitative measures of relatedness between bibliographic objects (e.g., journal articles) based on their bibliographic characteristics, such as reference lists, authors, or keywords (cf. Zupic & Čater, 2015). The aim of the present study was to provide an overview of the seminal work, psychological phenomena, and concepts that have influenced expectation violation research. Data was therefore screened to identify different spellings and misspellings in the document identifiers (i.e., author names, document titles, journal names) prior to analysis. This was done to filter out potential duplicates and merge them into a single entry, thus minimising the danger of artefacts caused by the repeated inclusion of single documents in the analysis (i.e., false positives). In addition, documents with less than ten citations were excluded from analyses as they would increase the complexity of the results without representing highly influential nodes in the networks. A second analysis, in which all documents were included, produced similar results (see supplemental materials).

Table 1. Search string used to identify records of interest in the Social Sciences Citation Index (SSCI).

TS = (							
	(						
		(	"expect*"	***OR***	"predict*"	***OR***	"anticipat*"
		)					
		NEAR/0					
		(	"violat*"	***OR***	"falsif*"	***OR***	error*
		OR	"disconfirm*"	***OR***	"devia*"	***OR***	"discrep*"
		OR	inconsist*	***OR***	"incongruen*"	***OR***	"diverg*"
		OR	"dissonan*"				
		)					
	OR	(	"violat*"	***OR***	"falsif*"	***OR***	error*
		OR	"disconfirm*"	***OR***	"devia*"	***OR***	"discrep*"
		OR	inconsist*	***OR***	"incongruen*"	***OR***	"diverg*"
		OR	"dissonan*"				
		)					
		NEAR/1					
		(	"expect*"	***OR***	"predict*"	***OR***	"anticipat*"
		)					
	)						
)							
AND							
TS = (							
	(		"maint*"	***OR***	"stabil*"	***OR***	"persist*"
		OR	"adapt*"	***OR***	"adjust*"	***OR***	"updat*"
		OR	"learn*"	***OR***	"consolidat*"	***OR***	"chang*"
		OR	"modif*"	***OR***	"accomodat*"	***OR***	"compensat*"
	)						
)							
AND							
PY = ((1990-2020))							
AND							
WC= (PSYCHOLOGY)							

TS = Web of Science search field "Topic" (title, abstract, keywords).
* = wild card for word ending (e.g., *expect** includes expectation, expectations, expectancy, etc.).
NEAR/0 = Key terms need to be located by located next to each other (e.g., expectation violation).
NEAR/1 = Key terms can be separated by no more than 1 word (e.g., violation of expectations).
PY = Publication year range.
WC = Web of Science category.

Analysis of keywords and trend topics

Keywords resulting from the SSCI search were preprocessed as follows. First, all keywords were transformed to lower case for easier identification of duplicates. Generic terms (e.g., "method", "construct"), words that do not denote a concept, such as names of countries or nationalities (e.g., "American"), and too broad research methods or disciplines (e.g., "behavioural experiment", "cognitive neuroscience") were excluded. Common single terms were converted into abbreviated synonyms commonly used in the field (e.g., "Cognitive Behavioural Therapy" as "cbt", "Event-Related potential" as "erp") to avoid special characters and long single-term overlap in the visualisations. Plural words were converted into their singulars (e.g., "expectations" to "expectation") to avoid separate nodes for the plural and singular forms of the same underlying term. Finally, the words "expectancy" and "expectation" were merged into the single term "expectation". In practice, the use of the terms "expectation" and "expectancy" overlap in the psychological literature, suggesting that quantitative effects arising from the distinction of these terms would be spurious in nature (cf. Andersen, 2019). The same was done with the terms "expectation violation", "expectancy violation", "violation of expectations", and "violated expectations" (i.e., all merged into "expectation violation"). A complete thesaurus of the raw and cleaned keywords is provided as supplementary data. To visualise important keywords (i.e., important topics of research in the literature corpus), we computed the frequency of each keyword in the Keywords Plus ® field of each document for each year in the time range 2010-2020. The time range was constrained to 2010-2020 to show emerging and influential topics in the literature. Trend topics were computed by calculating the median year of appearance for each keyword based on these frequency measures. Further bibliometric analyses were structured along two major analytical steps: Co-citation and bibliographic coupling.

Co-citation analyses

We measured the relevance of bibliographic objects (e.g., journal articles) for expectation violation research based on how often these objects appeared together in the reference lists of other bibliographic objects (i.e., were co-cited) in the full time range (i.e., 1990-2020) of the literature search (McCain, 1990; Small, 1973). Thus, the units of analysis here are the journal articles (or books) that are typically cited by articles identified through our literature search. It is important to note that these documents are not necessarily part of the literature search results but are rather found in the reference lists of that body of research. The rationale of this method is that co-cited documents have related claims, either by means of agreement or disagreement, about a shared subject of research and that the frequency of their co-citation scales proportionally to the relevance of their (combined) claims for a specific field of research (cf. Andersen, 2019; Dort, Strelow, French, et al., 2020). The importance of a document in the network was measured as a function of the importance of the documents with which it is often cited. To achieve this, we calculated the eigenvector centrality metric (Bonacich, 1987, 2007) for each document in the network. Eigenvector centrality is a measure of prestige for a node (e.g., a document) in a network (e.g., a co-citation network). The eigenvector centrality of a node was computed as a weighted average of the eigenvector centralities of all the other nodes connected to it:

$$e_i = \frac{1}{\lambda} \sum_{i<j} a_{ij}\, e_j$$

Here a_{ij} refers to the similarity matrix (see below) describing the strength of the relationship between pairs of documents. This equation can be reformulated in vector notation as the eigenvector equation ($\lambda\, e = A\, e$). In mathematical terms, the eigenvector centrality of the nodes in a network is equivalent to the left-land eigenvector with the largest eigenvalue of the similarity matrix A = a_{ij}. Thus, eigenvector centrality considers not only the number of connections a node has (i.e., its degree) but also the centrality of those nodes, with high-scoring nodes counting more than low-scoring nodes (cf. Golbeck, 2013, pp. 25–44). Thus, this method helps identify seminal work in a field of research, providing an approximation of the intellectual (or ideological) bases of the same (Dort, Strelow, French, et al., 2020; Pasadeos et al., 1998).

Bibliographic coupling analyses

We estimated the relatedness of bibliographic objects based on their reference lists' similarity. In contrast to co-citation analysis, this method shows how related two documents are by looking at how many references they have in common. The logic here is that if two documents cite the same references, it is likely that their subject of research or approach to a certain research question is similar (Biscaro & Giupponi, 2014). In contrast, documents with fewer commonalities in their reference lists are assumed to be influenced by less interconnected bodies of research (Dort, Strelow, French, et al., 2020). Documents with high eigenvector centrality in the bibliographic coupling network provide representative reference lists for numerous connected publications, as they share substantial overlap with many related documents. In contrast to co-citation analysis, bibliographic coupling analysis was limited to documents published between 2010 and 2020. This was done to achieve some basic level of similarity in the reference lists of the scientific documents and reduce systematic effects of the publication year, as current literature cites more recent publications and also has a wider range of available resources to cite (cf. Andersen et al., 2021).

Data clustering and visualisation

Exploratory data analysis was carried out using the R package *bibliometrix* (Aria & Cuccurullo, 2017) and customs scripts written in the R programming environment (R Core Team, 2021). Clustering and visualisation were carried out using a visualisation of similarities approach (VOS; van Eck & Waltman, 2010) as implemented in *VOS viewer* software 1.6.18. This method takes a co-occurrence matrix as an input. In the case of co-citation analyses, this matrix is a non-negative symmetric co-citation matrix of order *n x n*. Objects *1, …, n* are scientific publications that appear on the reference lists of the document corpus identified by the literature search (note that, in co-citation analysis, the objects *1, …., n* themselves need not necessarily be part of the document corpus itself). Two publications are co-cited if there is a third publication that cites both publications. The entries in the co-citation matrix correspond to the strength of the co-citation relationship linking two objects (Small, 1973). The higher the number of publications that cite two publications, the stronger the co-citation relationship between these publications. In the case of bibliographic coupling analyses, the *n x n* matrix is the opposite of a co-citation matrix. In a bibliographic coupling matrix, two publications are coupled if there is a third publication that both publications cite (van Eck & Waltman, 2014). The larger the number of publications two documents have in common, the stronger the bibliographic coupling relationship between these publications. The input matrix is normalised by correcting the matrix for differences in the total number of occurrences of the items in question. This is achieved by calculating the association strength between two nodes *i* and *j* as $s_{ij} = 2ma_{ij} / k_i\, k_j$, where a_{ij} denotes the weight of the edge between the two nodes (e.g., the number of times those two publications were co-cited by a third publication). k_i and k_j denote the total weight of all edges of node i and node j (i.e., the total number of times the publications were co-cited with other publications). m is the total weight of all edges in the network (i.e., the cumulated number of co-citations in the network; see van Eck & Waltman, 2009, 2014 for a more detailed description and empirical validation of this method). Next, a two-dimensional map is constructed by positioning nodes in a two-dimensional space in a way that strongly related nodes are close to each other, and weakly connected nodes are far away from each other. This is achieved via an optimisation algorithm (Borg & Groenen, 2005; van Eck & Waltman, 2009) that minimises the weighted sum of the squared Euclidean distances between all nodes. Finally, the assignment of nodes to clusters is achieved by maximising the modularity function:

$$V(c_1, \ldots, c_n) = \sum_{i<j} \delta(c_i, c_j)(s_{ij}, \gamma)$$

Here, c_i denotes the cluster to which i is assigned. $\delta(c_i, c_j)$ is 1 if $c_i = c_j$. γ is a resolution parameter that determines the level of detail of the resulting clustering solution (i.e., the number of clusters). In our analyses, we used a resolution parameter of 1.0 and constrained the solution to have at least 20 items in each cluster. We refer to van Eck and Waltman (van Eck & Waltman, 2014) and Newman and Girvan (Newman & Girvan, 2004) for an in-depth detailed discussion of this method.

To compare clusters to other clusters, we computed two graph measures of internal linkage, cluster density and average path length. Cluster density was defined as the ratio between successful connections between nodes within a cluster and the number of possible connections within a cluster. Density values closer to 1 indicate a higher degree of internal linkage. Average path length refers to the average graph distance between all pairs of nodes within a cluster. Connected nodes have a graph distance equal to 1.0. Higher values in average path length indicate that the nodes of a cluster are located farther away from each other (i.e., information must pass through more intermediate nodes to reach its destination). Graph measures were computed using Gephi 0.9.7 (Bastian et al., 2009).

Results

Our search for literature on expectation change vs expectations persistence following expectation violations identified 1445 bibliographic objects from 353 different sources (e.g., journals, books). After removing meeting abstracts, news items, and documents with missing data (e.g., no reference lists), 1407 documents from 349 sources remained for analysis. Journal articles (i.e., research reports) constituted the biggest group of documents (N = 1212), followed by reviews (N = 102), proceedings papers (N = 51), editorial material (22), and book chapters (N = 20). As depicted in Figure 1, the number of scientific publications related to expectation violations has increased exponentially in the time range 1990 – 2020 ($F(3,27) = 181.2$, $p < 2.2e\text{-}16$, $R^2 = 0.953$), with little variation in the number of documents published between the years 1990 and 2006 ($b_{change} = 1.37$, SE = 0.43) and a sustained increment in scientific production between 2007 and 2020 ($b_{change} = 8.72$, SE = 0.57).

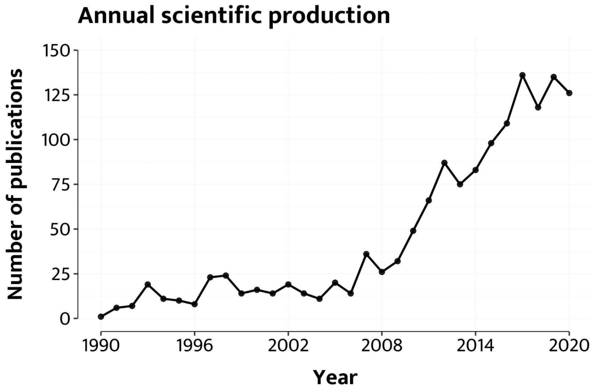

Figure 1. The number of publications per year for the identified literature corpus.

Note. Solid dots depict the absolute number of publications.

Figure 2 shows the references with the highest number of local citations (i.e., citations from within the identified literature corpus). Many of these documents are focused on how mesencephalic dopamine neurons encode reward prediction errors in the central nervous system (Montague et al., 1996; Schultz, 1998; Schultz et al., 1997; Schultz & Dickinson, 2000), providing a biological proxy for the magnitude of the experienced expectation violation, and how higher brain areas use this information to respond in accordance to situational demands (Behrens et al., 2007; O'Doherty et al., 2003) and adjust behaviour when situational outcomes are worse than expected (Botvinick et al., 2001; Gehring et al., 1993; Gehring & Willoughby, 2002; Holroyd & Coles, 2002; Miltner et al., 1997). Conversely, other documents focused on explaining variation in the degree to which prediction errors lead to changes in behaviour or experience (Pearce & Hall, 1980; Rescorla & Wagner, 1972) as well as on developing statistical (Friston, 2005) and computational (Sutton & Barto, 1998) frameworks and methods (Delorme & Makeig, 2004) to examine these phenomena. As depicted in Figure 3, analyses revealed that research focused on dopamine and frontal brain structures, such as the Anterior Cingulate Cortex (ACC), performance measures (e.g., behavioural responses, errors) and outcomes (e.g., rewards), have moved centre stage. In addition, research focused on the processing of expectation violations in psychological disorders (e.g., obsessive-compulsive disorder – OCD, substance abuse, depression) and at different stages of the human lifespan (e.g., childhood) have gained traction since 2010.

Co-citation network

Co-citation analysis revealed a total of 512 documents organised into five clusters. As depicted in Figure 4, cluster 1 (orange) was the biggest cluster of research (containing 31.9% of the data), followed by cluster 2 (blue; 23.3%), cluster 3 (magenta; 21.3 %), cluster 4 (green; 15.5 %), and clus-

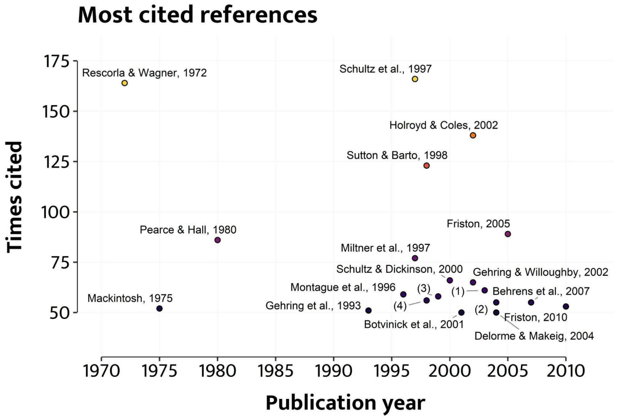

Figure 2. References with the highest number of citations from within the identified literature corpus.

Note. (1) = O'Doherty et al., 2003, (2) O'Doherty et al., 2003, (3) Rao & Ballard, 1999, (4) = Schultz, 1998

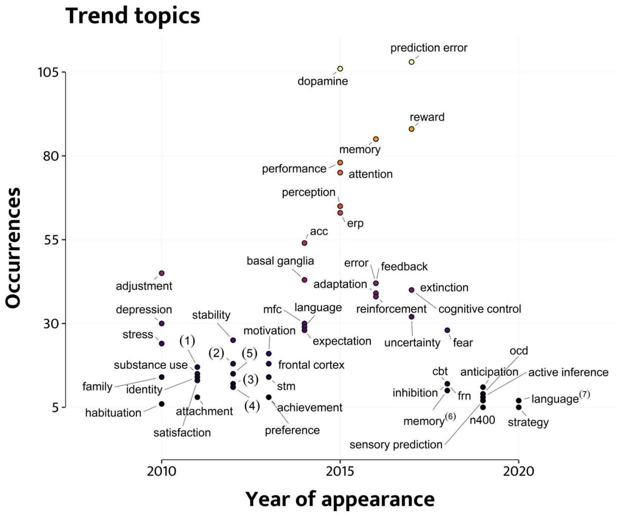

Figure 3. Trend topics for each year in the time range 2000-2020.

Note. Depicted are keywords with more than 5 occurrences and only the top 5 ranking keywords for each year. Abbreviations are "acc" = "anterior cingulate cortex", "cbt" = "cognitive behavioural therapy", "ern" = "error-related negativity", "ltm" = "long-term memory", "stm" = "short-term memory", and "ocd" = "obsessive-compulsive disorder. (1) = judgment, (2) = identification, (3) = attribution, (4) = awareness, (5) communication, (6) = memory refers to "episodic memory" rather than general memory processes, (7) = language refers to "language comprehension" rather than general language-related processes.

ter 5 (8 %). Here, we describe these clusters ordered by the degree of their interconnections and relevance for the network (i.e., as opposed to ordering them by cluster size). In addition, we refer to a subset of the most central publications in each of the clusters (a more in-depth discussion of the clusters and their relations is performed in individual sections in the discussion).

Co-citation cluster 1 (orange) was located centrally in the network. It was characterised by nodes with high eigenvector centrality, providing crucial links between other clusters of literature, most prominently, co-citation clusters 2 (blue) and 4 (green). In other words, literature in co-citation cluster 1 was heavily cited with literature from co-citation clusters 2 (blue) and 4 (green). However, publications from clusters 2 and 4 were not cited together as frequently. Literature assigned to co-citation cluster 1 was dominated by research showing that phasic dopaminergic activity tracks the discrepancy between expected and obtained outcomes (e.g., Schultz et al., 1997) in accordance with a temporal-difference learning rule (cf. Sutton & Barto, 1998). A hallmark of this process is that the expected value of an action (or event) is updated dynamically as a function of available outcome information, such as its (expected) outcome history (Daw et al., 2005). Co-citation cluster 2 (blue) was characterised by a high density (see Table 2), as depicted by the high degree and closeness of its internal linkage (i.e., high number of interconnections within the cluster; Golbeck, 2013, pp. 25–44). This means that the cluster contains a highly interrelated set of literature that is typically cited by other literature from within the cluster. Literature assigned to co-citation cluster 2 (blue) was dominated by research on electrophysiological measures of brain activity evoked by different kinds of expectation violations, such as ERPs following behavioural errors (Miltner et al., 1997) or performance feedback (Alexander & Brown, 2011), which indicate that the ACC is the main hub in a neural network that supports action outcome monitoring (Holroyd & Coles, 2002). In contrast, the distribution of co-citation cluster 4 (green) was more widespread. More precisely, there appeared to be two sub-clusters, of which one was more strongly associated with co-citation cluster 1 (orange). The most central documents in co-citation cluster 4 (green) were focused on the modelling of how prediction errors drive learning by triggering expectation updating (e.g., Schultz & Dickinson, 2000) either by directly modulating the associative value of events (Rescorla & Wagner, 1972) or, more indirectly, by affecting the degree of attention that is allocated to events associated with prediction errors (Pearce & Hall, 1980).

Co-citation cluster 3 (magenta) and co-citation cluster 5 (yellow) were located on the outer regions of the network and appeared to share some degree of overlap. Co-citation cluster 3 (magenta) was dominated by literature indicating that the brain processes the statistical regularities of events in accordance with a Bayesian scheme (Rao & Ballard, 1999; Yu & Dayan, 2005). The main idea here is that higher cortical structures guide lower-level processing areas by formulating expectations about bottom-up input. When these expectations are violated, lower cortical structures generate prediction error signals that can be used by higher brain structures to update expectations and reduce uncertainty in the future (cf. predictive coding; Friston, 2005). Conversely, the most influential documents in co-citation cluster 5 (yellow) concerned emotional and psychological reactions to expectation violations, such as strategies for compensating inconsistency between expectations and expectation-disconfirming evidence (Proulx et al., 2012), but also individuals' propensity to positive and negative affect (Watson et al., 1988) or sensitivity to rewards and punishments (Carver & White, 1994). In addition, this Co-citation Cluster 5 (yellow) included research on statistical procedures for evalu-

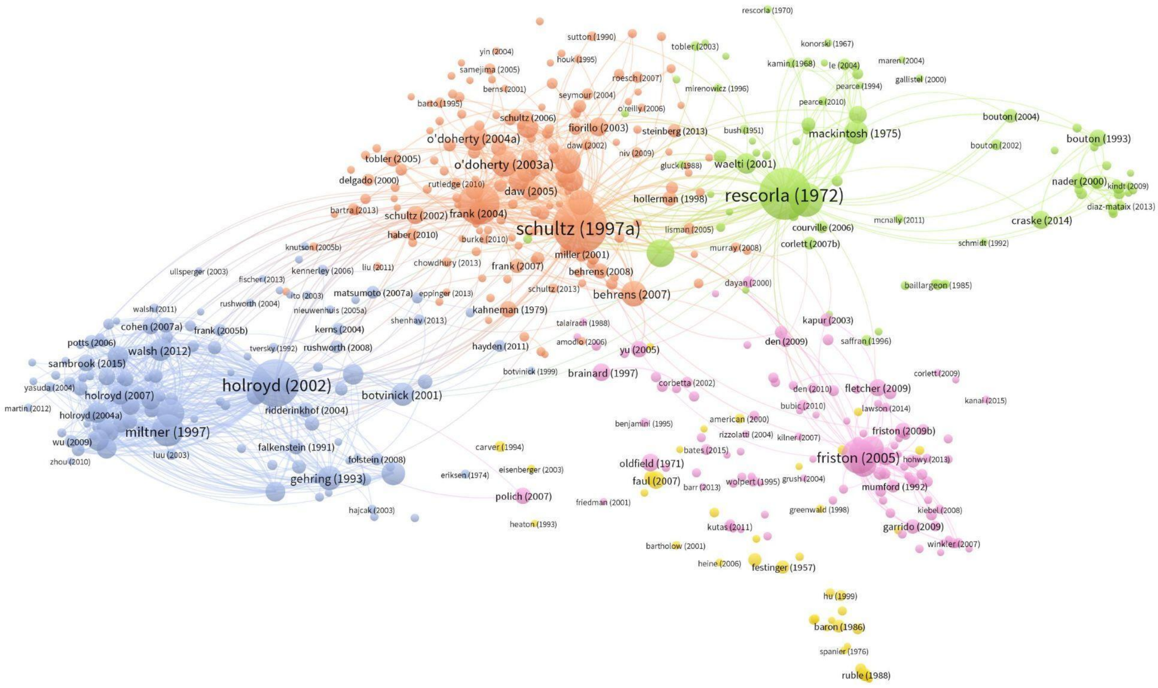

Figure 4. Visualisation of the co-citation network.

Note. Depicted are journal articles, books, and book chapters cited by the documents identified through the literature search. The size of the circles shows the number of citations received by a reference (i.e., bigger nodes received more citations that smaller nodes). The proximity of the network nodes shows their relatedness, with nodes close to each other often appearing together in the reference lists of other documents. Cluster topics (clockwise from left): Blue cluster = "electrophysiological measures of brain activity and their relation to different kinds of expectation violations", orange cluster = "dopaminergic coding of prediction errors in reward learning", green cluster = "The role of prediction errors in (associative) learning", yellow cluster = "emotional and psychological reactions to expectation violations", magenta cluster = "Integrative theoretical accounts of brain function based on hierarchical prediction error coding". Node names were abbreviated to avoid overlap as much as possible. Digital object identifiers (DOIs) for the top-ranking nodes in each cluster are provided in Tables S1-S5. In addition, the files needed to create an interactive version of the network can be found the project's online repository (https://osf.io/ymk8q/).

Table 2. Co-citation network, cluster statistics.

Cluster	Focus of research	Density	Av. path length
1 (orange)	Dopaminergic coding of prediction errors in reward learning	0.625	1.375
2 (blue)	Electrophysiology of expectation violation	0.791	1.209
3 (magenta)	Integrative theoretical accounts of brain function based on hierarchical prediction error coding	0.504	1.498
4 (green)	The role of prediction errors in (associative) learning	0.409	1.515
5 (yellow)	Emotional and psychological reactions to expectation violations	0.266	1.869

Note. Av. path length = Average path length.

ating and planning empirical studies, such as the computation and evaluation of statistical effect sizes and statistical power (Cohen, 1988; Faul et al., 2007).

Bibliographic coupling network

Bibliographic coupling analysis of documents published between 2010 and 2020 revealed a total of 454 interconnected nodes organised into six clusters (see **Figure 5** and **Table 3**). Cluster 1 (orange) was the biggest, containing 29.3% of the data, followed by cluster 2 (blue, 21.4%), cluster 3 (magenta, 18.5 %), cluster 4 (green, 14.3 %), cluster 5 (yellow, 11.5 %), and cluster 6 (brown, 5 %). Here, clusters are described by their relevance within the network.

Bibliographic coupling cluster 3 (magenta) was the most central in the network. Literature assigned to coupling cluster 3 (magenta) was characterised by high eigencentrality, providing links for other clusters of literature, most prominently documents assigned to coupling cluster 4 (green) and coupling cluster 5 (yellow). The most central documents in coupling cluster 3 (magenta) were literature reviews indicating that prediction errors constitute the primary mechanism for information sharing within the human brain (e.g., Den Ouden et al., 2012; Kim, 2013; Walsh & Anderson, 2014). A central claim in the studies is that prediction errors trigger information processing along different paths in the neural hierarchy (e.g., striatum, ventromedial prefrontal cortex; Walsh & Anderson, 2014) to support different types of learning (cf. model-free vs model-based

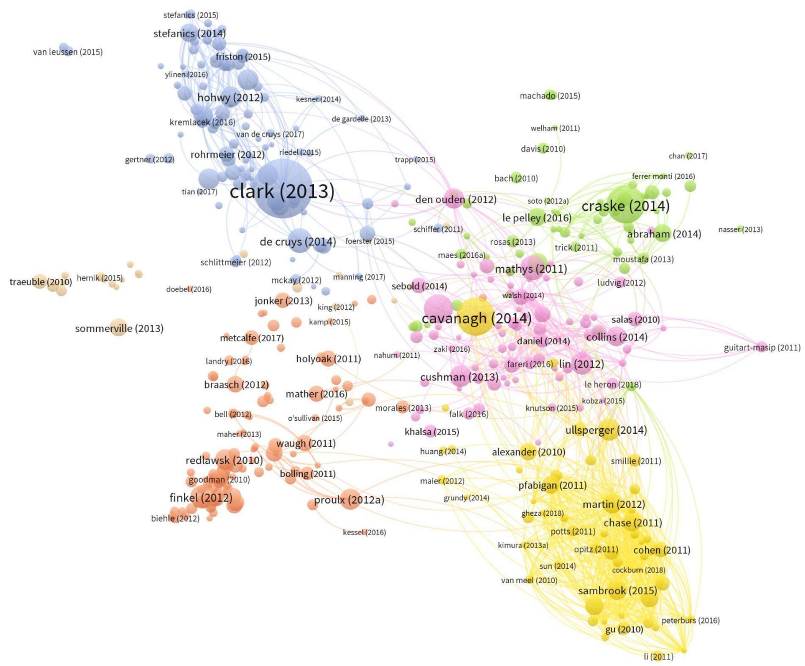

Figure 5. Visualisation of the bibliographic coupling network.

Note. Depicted are journal articles, books, and book chapters identified by the literature search. The size of the circles shows the number of citations received by a document (i.e., bigger nodes received more citations than smaller nodes). The proximity of the network nodes shows their relatedness, with nodes close to each other having more references in common in their reference lists (i.e., more similar reference lists). Cluster topics (clockwise from left): Brown cluster: "Developmental aspects of expectation violations", orange cluster = "Expectation violations in interpersonal contexts", blue cluster = "Expectation violations as core processes in psychopathology", magenta cluster = "Prediction errors trigger information processing in the neural hierarchy", green cluster = "Role of expectation violations in psychotherapy", yellow cluster = "biological markers for expectation violation". Node names were abbreviated to avoid overlap as much as possible. Digital object identifiers (DOIs) for the top-ranking nodes in each cluster are provided in Tables S6-S11. In addition, the files needed to create an interactive version of the network can be found the project's online repository (https://osf.io/ymk8q/).

Table 3. Bibliographic coupling network, cluster statistics.

Cluster	Focus of research	Density	Av. path length
1 (orange)	Expectation violations in interpersonal contexts	0.066	2.729
2 (blue)	Expectation violations as core processes in psychopathology	0.358	1.763
3 (magenta)	Prediction errors trigger information processing in the neural hierarchy	0.702	1.306
4 (green)	Role of expectation violations in psychotherapy	0.526	1.645
5 (yellow)	Biological markers for expectation violations	0.980	1.020
6 (brown)	Developmental aspects of expectation violations	0.229	2.233

Note. Av. path length = Average path length.

learning; Walsh & Anderson, 2014), giving rise to individual differences in preferential choice behaviours and cognitive control recruitment (Kim, 2013). Conversely, literature assigned to coupling cluster 4 (green) was dominated by research analysing how different aspects of expectation violation (and expectation violation processing in the human brain) lead to better or worse learning. For instance, some studies (Greve et al., 2017) showed that larger and more precise prediction errors lead to better memory formation, while others (Moustafa et al., 2013) described dissociable brain circuits for the processing of value and contextual information of aversive and appetitive stimuli. Coupling cluster 4 (green) also provided context for applying these findings in clinical practice (e.g., role of expectation violations in fear extinction and exposure therapy; Craske et al., 2014). In contrast, research assigned to bibliographic coupling cluster 5 (yellow) focused on identifying candidate biological markers for assessing the impact of expectation violation on adaptive behaviour (Cavanagh & Frank, 2014; San Martín, 2012; Ullsperger, Fischer, et al., 2014). Here, transient changes in the theta band of the EEG (Cavanagh & Frank, 2014), and ERP components, such as the FRN (Pfabigan et al., 2011) and the P300 (San Martín, 2012), emerged as proxy measures for ongoing outcome processing (e.g., assessment of value and predictability) and outcome dependent behavioural adjustments (see Ullsperger, Fischer, et al., 2014 for review).

Visual inspection of the bibliographic coupling network map suggested that there was less overlap between nodes from clusters 1 (orange), 2 (blue) and 6 (brown). The most central documents in coupling cluster 1 (orange) shared a strong focus on the impact of expectation violations on social (i.e., interpersonal) processes (Harris & Fiske, 2010; Proulx et al., 2012; Proulx & Inzlicht, 2012). In particular, one set of articles focused on how inconsistencies between expected and experienced demands in social situations can evoke aversive cognitive and emotional states (Proulx et al., 2012; Proulx & Inzlicht, 2012), motivating compensatory mechanisms in multiples domains (e.g., behaviour and experience; Randles et al., 2015). Results revealed strong connections between coupling cluster 1 (orange) and clusters 4 (magenta) and 5 (yellow), mainly mediated by studies focused on finding corresponding correlates of physiological activity that could help better characterise the effect of expectation violations in social contexts (e.g., Bell et al., 2016). In contrast, literature assigned to bibliographic cluster 2 (blue) focused on adaptations of the predictive coding framework (cf. Clark, 2013) for various fields of (clinical) application (e.g., treatment of depression: Chekroud, 2015; physical movement: Schiffer & Schubotz, 2011; selective deficits in autism: Van de Cruys et al., 2014). One main claim in coupling cluster 2 (blue) is that psychological disorders arise from maladaptive strategies for minimizing the impact of expectation violations, such as developing overly pessimistic expectations in the first place (e.g., negativistic beliefs in depression; Chekroud, 2015) or due to lack of flexibility when coping with prediction errors (e.g., stereotyped behaviours in autism; Van de Cruys et al., 2014). Finally, bibliographic coupling cluster 6 (brown) appeared less consistently developed, with a subset of documents sparsely distributed across the network and a subset of documents with more interrelations located in the outer regions of the network. Research in this sub-cluster of documents was centred on the effects of prior experience on interpersonal expectation violations in early childhood (He et al., 2011; Sommerville et al., 2013; Träuble et al., 2010).

Discussion

Expectation violations have motivated a wide variety of research avenues within different fields of psychology. The present study describes how these research domains are related and provides a scientific map that quantifies their theoretical and empirical commonalities. Our goal was to inform future research projects by providing an overview of the field's most influential work, helping researchers identify links between different research domains, such as empirical approaches, topics, and variables.

The literature search identified 1445 documents focussed on expectation change vs persistence following expectation violations. While expectations are a very prominent topic in psychology (Rief et al., 2015), a smaller group of studies has addressed expectation violations and whether (and to what degree) expectation violations lead to expectation updates. The bibliometric analysis revealed that the number of scientific publications related to expectations violations increased exponentially since the early 2000s, underlining the need for a comprehensive review. Indeed, many of the most cited publications in the literature corpus were published in this time period (e.g., Friston, 2005; Gehring & Willoughby, 2002; Holroyd & Coles, 2002; Rao & Ballard, 1999; e.g., Schultz et al., 1997; e.g., Schultz & Dickinson, 2000). In addition, analyses of trend topics indicated that expectation violations arouse scientific interest in multiple research areas, such as cognitive, biological, developmental, and clinical psychology. Furthermore, co-citation analyses revealed that many of the most cited articles in the literature corpus constituted a central node in the co-citation network and that each of these nodes was located in a different cluster (e.g., co-citation cluster 3 (magenta): Friston, 2005; co-citation cluster 2 (blue): Holroyd & Coles, 2002; co-citation cluster 1 (red): Schultz et al., 1997; co-citation cluster 4 (green): Schultz & Dickinson, 2000), providing the theoretical and empirical basis for separable areas of research. Many of these highly cited scientific publications focused on different measures of neural activation (e.g., fMRI: Behrens et al., 2007; EEG/ERPs: Gehring & Willoughby, 2002; firing of dopamine neurons: Schultz & Dickinson, 2000) along with methods to analyse them (Delorme & Makeig, 2004), indicating that the ability to map the discrepancy between expected and observed outcomes to measurable indices of brain function has been one of the motors driving advancements in expectation violation research.

Expectation updating as a function of prediction errors

Co-citation analyses revealed that expectation violation research is grounded on the assumption that experiencing discrepancies between expectations and disconfirming evidence (i.e., prediction errors) is necessary for expectation updating, thus constituting the core mechanism of learning in general. However, the co-citation network also revealed that prediction errors do not necessarily serve the same function in different theoretical models and across different stages of the information processing stream. For instance, co-citation cluster 4 (green) was dominated by the Rescorla-Wagner model (Rescorla & Wagner, 1972), where prediction errors drive changes in the predictive value of cues/events in a direct manner. Following Rescorla and Wagner (1972), large discrepancies between expected and observed outcomes (i.e., large prediction errors) will result in large changes in the associative value of the cues predicting the outcome (e.g., specific situational characteristics). Within cognitive psychological frameworks, associative values have been used synonymously with expectations (Den Ouden et al., 2012; Gershman, 2015; Riels et al., 2022). Another aspect is that the sign of the prediction error (negative means the outcome is worse than expected, and positive means the outcome is better than expected) determines the type of learning (e.g., the predictiveness of cues is reduced after negative prediction errors, thus leading to expectation change). Conversely, when the prediction error is zero, the outcome has been fully predicted, and therefore expectations do not change. In contrast, in Pearce and Hall's model (1980), also a prominent node in co-citation cluster 4 (green), the absolute value of the prediction error determines the degree of attention paid to the preceding cues when encountered in future situations rather than changing their associative value directly. This *"attentional gain"* is believed to facilitate subsequent learning when the prediction error is large and to inhibit learning when the prediction error is small. While the predictions of the Rescorla-Wagner and Pearce and Halls models might seem mutually exclusive, a growing body of research indicates that they might in fact serve complementary roles (cf. Roesch et al., 2012). For instance, in the Rescorla-Wagner model, the degree to which prediction error leads to changes in expectations is scaled by a learning parameter that could be explained by a mechanism like Pearce and Hall's attentional gain. In the co-citation network, literature addressing this issue was assigned to co-citation cluster 1 (orange), demonstrating that research focussing on how different learning rules explain dynamic expectation updating (e.g., by integrating predictions from the Rescorla-Wagner and Pearce-Hall models; cf. Sutton & Barto, 1998) provide essential links between different research domains in the expectation violation literature. The tight connections between co-citation cluster 4 (green) nodes and co-citation cluster 1 (orange) reflect the importance of classical computational models of learned behaviour for neurobiological investigations of prediction error coding as a potential underlying mechanism - and vice versa. Meanwhile, a second, less central group of co-citation cluster 4 (green) studies showed fewer associations with co-citation cluster 1 (orange). While these studies were part of an associative learning framework and built upon classical theories like Rescorla-Wagner and Pearce-Hall, their focus was on learning processes beyond the initial acquisition of conditioned responses, such as retrieval and reconsolidation (Bouton, 1993; Díaz-Mataix et al., 2013; Nader et al., 2000), or extinction and return of previously conditioned responses (Bouton, 2002; Craske et al., 2014; Kindt et al., 2009), partly with special focus on aversive learning. Although the dopaminergic system has been linked to extinction and consolidation processes (Abraham et al., 2014; Haaker et al., 2013, 2015; Panitz et al., 2018), the co-citation network suggests that there has been less exchange between fields compared to the investigation of de novo acquired expectations.

Neural markers of expectation violation processing

Interestingly, numerous studies assigned to co-citation cluster 1 (orange) indicate that the activity of the midbrain dopamine system provides an accurate measure of how signed prediction errors are processed by organisms (Frank et al., 2004; Pessiglione et al., 2006) in a dynamic (i.e., temporal difference) manner (O'Doherty et al., 2003). Here, dopaminergic activity has been shown to increase (i.e., neural activity bursts) in response to rewards. However, when an organism learns that a reward is predicted by a cue, dopamine bursts are increasingly elicited by the cue rather than by the rewarding outcome itself. In contrast, dopaminergic activity decreases (i.e., neural activity dips) when the predicted reward is omitted, leading to a devaluation of the preceding cue (Schultz et al., 1997). Co-citation cluster 1 (orange) was dominated by studies implementing instrumental or classical (i.e., Pavlovian) learning paradigms (O'Doherty et al., 2003) and fMRI (Daw et al., 2011) to analyse how these processes influence the activity of higher cortical structures in the human brain. These studies show that, in addition to the midbrain dopamine system, the medial prefrontal cortex plays a crucial role in action-outcome prediction and valuation (Behrens et al., 2007). However, fMRI is limited by its poor temporal resolution (Carlson et al., 2011), in particular with regard to the valuation of outcomes, which single-cell recordings in dopaminergic midbrain areas indicate is highly temporally constrained (Fiorillo et al., 2003; Montague et al., 1996). Therefore, the ERP technique, which allows for the analysis of outcome and prediction error processing at a much higher temporal resolution, has emerged as a prominent imaging alternative in the expectation violation literature. EEG/ERP literature was mostly assigned to co-citation cluster 2 (blue) and bibliographic cluster 5 (yellow). In these studies, salient and surprising events, such as unexpected stimuli or outcomes, have been shown to elicit amplitude modulation of the N2 and P3 ERP components (Baker & Holroyd, 2011; also see Ullsperger, Danielmeier, et al., 2014 for review). However, expectation violation research has been particularly interested in two other ERP components, the ERN (Gehring et al., 1993) and the FRN (Miltner et al., 1997). ERN and FRN are believed to reflect the activity of the Anterior Cingulate Cortex (ACC), a structure in the Medial Prefrontal Cortex that is highly innervated by the midbrain dopamine system (Walsh & Anderson, 2014). The ERN is a transient negative deflection of the ERP that reaches its peak over frontal-central regions of the scalp within 100 ms from error commission in speeded response tasks. Conversely, the FRN is a negative deflection of the ERP that shows a scalp topography similar to the ERN and is elicited 200 – 400 ms after performance feedback, in particular when feedback is needed to estimate the value of an action (cf. Hajcak et al., 2007). Research assigned to co-citation cluster 2 (blue) indicates that phasic dips in dopaminergic activity disinhibit neurons in the ACC, leading to negative shifts (i.e., increases) in FRN amplitude (Walsh & Anderson, 2014). In contrast, dopamine bursts inhibit the ACC, inducing positive shifts (i.e., decreases) in the FRN (Holroyd et al., 2008; Pfabigan et al., 2011). Moreover, drug-induced stimulation of the midbrain dopamine system has been shown to selectively affect the magnitude of these ERPs (de Bruijn et al., 2004; Santesso et al., 2009), indicating that fluctuations in ACC activity (indexed by ERN and FRN) act as endogenous monitoring mechanism that uses predictions errors from lower brain structures to signal the need for adaptation in case of deviations from internally represented goals (Holroyd & Coles, 2002; Shenhav et al., 2013). Indeed, spectral analyses of ERP components like the N2, ERN, and FRN indicate that they are supported by synchronisation in the theta frequency band of the EEG following unexpected or motivationally salient events (Cavanagh & Frank, 2014) and that coupling in the theta frequency band facilitates communication across brain areas (Cavanagh & Frank, 2014). These findings converge in the assumption that prediction errors at higher cortical structures are influenced by prediction error signals elicited by lower cortical structures (Alexander & Brown, 2015). Furthermore, it has been suggested that higher-order prediction errors act as teaching signals for the system, as it aims to gradually fine-tune the action-selection process to achieve better outcomes (Luque et al., 2012). In the expectation violation literature, this process has often been referred to as predictive coding (Clark, 2013; Friston, 2005; Rao & Ballard, 1999). Predictive coding proposes that higher cortical structures use the prediction error generated by lower cortical areas to predict under which circumstances (e.g., situational cues) that prediction error is elicited. These higher-order representations (i.e., more abstract and generalised rules) are then used to modify lower-order representations (i.e., less abstract anticipation of sensory information) to reduce the error margin of future predictions (Alexander & Brown, 2018). The updated higher-order representation of the environment is believed to integrate temporal and spatial information about immediate and distant influences on the prediction error, giving the organism the ability to update expectations dynamically and predict the consequences of future events (Clark, 2013). In the co-citation network, research focused on predictive coding was mostly assigned to co-citation cluster 3 (magenta).

Psychological responses to expectation violations

ERPs like the FRN have been shown to be sensitive to expectation violations in more complex contexts, such as interpersonal cooperation (Bell et al., 2016). In addition, research shows that personal factors that influence the experienced significance of these contexts lead to significant changes in performance monitoring ERPs (García Alanis et al., 2019; Mueller et al., 2014). Signals like the ERN and FRN, therefore, provide a valuable tool to assess the degree of arousal induced by expectation violations in a wide variety of contexts. Indeed, a growing body of literature concerns how individual-level factors interact with situational characteristics to influence the psychological consequences of expectations violations (e.g., Rapp et al., 2021; Weinberg et al., 2016). While many studies have focused on expectation change following expectation violations, bibliographic coupling analysis revealed a separable subset of studies focusing on the persistence of expectations even when individuals encounter disconfirming evidence. For instance, Proulx and Inzlicht (2012) propose that the aversive arousal evoked by expectation violation might be the key to understanding why and when expectations change. Proulx and colleagues (2017) showed that aversive arousal is particularly evoked by expectation violations by measuring variations in pupillary dilation in response to expectation-violating faces (i.e., presented upside-down and distorted). These stimuli produced pupillary dilation earlier than other faces (e.g., neutral or threatening faces). In the Meaning Maintenance Model (MMM), Proulx and Inzlicht (2012) outline five different types of responses that either reduce or eliminate the source of the aversive arousal or reduce arousal without changing the original source (cf. Proulx et al., 2012). The first two responses are intended to reduce the discrepancy between an expectation and a disconfirming experience: Individuals may reinterpret the experience in a way that is consistent with their beliefs (*"assimilation"*) or change the expectation post-hoc so that it becomes consistent with the discrepant experience (*"accommodation"*). If individuals do not have resources available to resolve the expectation violation (e.g., if accommodation or assimilation cannot be applied or do not work), the model postulates that the negative arousal elicited by expectation violations can be dispelled in one of three indirect ways (a process known as *"fluid compensation"*): Individuals either engage in heightened commitment to other (existing) expectations that have not been disconfirmed (*"affirmation"*), or they find (*"abstraction"*) or construct (*"assembly"*) expectations that have not been disconfirmed. In some cases, newly assembled expectations may also give meaning to the expectation violation (e.g., someone might think: "this happened for a reason"), reducing the significance of the original violated expectation. Most of the research on the psychological responses to expectation violations focused on fluid compensation processes rather than on processes that directly reduce the discrepancy between an expectation and a disconfirming experience (cf. Randles et al., 2015).

Limitations

The present bibliometric review provides an overview of research focused on expectation violation, expectation change, and expectation persistence. As in any data created by humans, bibliographic data may possess systematic biases that warrant caution when interpreting results. For example, it has been shown that papers are cited disproportionally more often when they are first-authored and/or last-authored by men (Chatterjee & Werner, 2021; Dworkin et al., 2020), White persons (Bertolero et al., 2020), and researchers from highly research-active countries (e.g., North America, Western Europe, East Asia; Gomez et al., 2022). While it can be argued that bibliometric reviews - based on citation counts - allow one to draw conclusions about the influence and reach of publications, they may not necessarily provide a perfectly objective picture of research efforts, importance, or quality. Moreover, bibliographic reviews are not necessarily a substitute for extensive literature surveys. Instead, the present analysis provides complementary information to other narrative and systematic reviews focused on smaller, more specialised research areas. More targeted reviews are necessary to assess whether and to what degree the here-discussed psychological, and neural mechanisms interact with individual-level and situational factors to influence the processing of expectation-violating information (cf. Pinquart, Endres, et al., 2021). One further limitation of the present study is that we did not analyse and control for the quality of the discussed scientific publications (i.e., other than the controls required by the respective journals in which the studies were published). A challenge for future, more specialised reviews could be to analyse how the here discussed effects on neural and psychological measures are consistent across a wide variety of studies (e.g., using meta-analytic approaches). Lastly, the present study is limited by its focus on a subset of data, as we only analysed research published from 1990 to 2020 in the field of psychology. Other, more in-depth reviews may include the most recent publications and literature from adjacent fields of research. However, the former might represent a challenge, as it takes time for scientific publications to be cited.

Regarding possible fields of application, translational clinical research has helped elucidate central mechanisms in psychopathology (e.g., anxiety disorders; Craske et al., 2018; Pittig & van den Berg, 2016). Cluster 4 (green) in the co-citation network captured recent studies that have successfully built on basic research knowledge to inform clinical applications. For example, Bach and colleagues (2010) used classical learning theories as well as Dynamic Causal Modelling (a popular method within the predictive coding framework) to model fear-conditioned skin conductance responses – a marker of expectation strength in learned prediction processes (Hamm & Vaitl, 1996). Moreover, Abraham and colleagues (Abraham et al., 2014, 2016) showed that dopaminergic mechanisms of associative learning (as discussed in predictive coding and reinforcement learning research) play a crucial role in the development, generalisation, and extinction of fear. Beyond highlighting the role of expectations and expectation change for the development and persistence of psychological symptoms and disorders (Rief et al., 2015), research in clinical psychology has moved towards applying these findings in the context of psychotherapy. As expectations are among the strongest predictors of treatment outcomes in the treatment of various medical conditions, promising interventions focus on changing patients' dysfunctional expectations about their disorders and potential treatment effects (Craske et al., 2014; Doering et al., 2018; Rief & Glombiewski, 2016). Knowledge of the mechanisms that influence the persistence or change of expectations has also led to the development of new intervention methods in other fields, such as correcting stereotypes about social groups (Dort, Strelow, Schwinger, et al., 2020; Zingora et al., 2020), a much more underrepresented topic in expectation violation research, as captured by the present bibliometric review.

Conclusion

The present study is the first bibliometric review of expectation violation research. To the best of our knowledge, no other study has sought to quantitatively assess and map the structure of this growing body of research, providing an overview of similarities and links between different subdomains. Our results show that, historically and currently, much expectation violation research is grounded on associative learning theories that explain how organisms process deviations from internally represented goal states and what characteristics of expectation-violating information make expectation change more probable. In addition, our results firmly demonstrate that biological markers, such as fluctuations in the activity of the midbrain dopamine system, in the activity of the ACC, and ERPs (e.g., ERN, FRN, and N2) can be seen as indices for expectation violation processing in the human brain, providing promising non-invasive methods to analyse expectation violations across a wide variety of experimental settings. The present study thus allows an overarching perspective for expectation violation research, providing informative networks and clusters of literature for scientists and practitioners who work across different research areas, thus facilitating interdisciplinary advancements in psychology.

Author Contributions

Original idea and conceptualization: Hanna Christiansen, Christian Panitz, Martin Pinquart, José C. García Alanis, Anna Strelow, Martina Dort; Literature search: José C. García Alanis; Data pre-processing: José C. García Alanis, Christian Panitz, Hanna Christiansen, Martin Pinquart, Anna Strelow, Martina Dort; Formal analysis: José C. García Alanis; Visualization: José C. García Alanis; Writing – original draft preparation: José C. García Alanis; Writing – review, editing, and approval of final draft: José C. García Alanis, Christian Panitz, Martin Pinquart, Hanna Christiansen, Anna Strelow, Martina Dort

Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

Acknowledgements and Funding

This study was conducted in the context of the Research Training Group "Expectation Maintenance vs. Change in the Context of Expectation Violations: Connecting Different Approaches" funded by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG Ref. no.: 290878970-GRK 2271).

Open Access funding was provided by the Open Access Publishing Fund of Philipps-Universität Marburg with support of the German Research Foundation (DFG, Deutsche Forschungsgemeinschaft).

Data Accessibility Statement

All analysis scripts and data needed to reproduce the reported results are available at:
https://osf.io/ymk8q/

Submitted: January 31, 2022 PDT, Accepted: March 17, 2023 PDT

References

Abraham, A. D., Neve, K. A., & Lattal, K. M. (2014). Dopamine and extinction: a convergence of theory with fear and reward circuitry. *Neurobiology of Learning and Memory*, *108*, 65–77. https://doi.org/10.1016/j.nlm.2013.11.007

Abraham, A. D., Neve, K. A., & Lattal, K. M. (2016). Effects of D1 receptor knockout on fear and reward learning. *Neurobiology of Learning and Memory*, *133*, 265–273. https://doi.org/10.1016/j.nlm.2016.07.010

Alexander, W. H., & Brown, J. W. (2011). Medial prefrontal cortex as an action-outcome predictor. *Nature Neuroscience*, *14*(10), 1338–1344. https://doi.org/10.1038/nn.2921

Alexander, W. H., & Brown, J. W. (2015). Hierarchical Error Representation: A Computational Model of Anterior Cingulate and Dorsolateral Prefrontal Cortex. *Neural Computation*, *27*(11), 2354–2410. https://doi.org/10.1162/neco_a_00779

Alexander, W. H., & Brown, J. W. (2018). Frontal cortex function as derived from hierarchical predictive coding. *Scientific Reports*, *8*(1), 3843. https://doi.org/10.1038/s41598-018-21407-9

Andersen, N. (2019). Mapping the expatriate literature: a bibliometric review of the field from 1998 to 2017 and identification of current research fronts. *The International Journal of Human Resource Management*, *32*(22), 4687–4724. https://doi.org/10.1080/09585192.2019.1661267

Andersen, N., Corr, P. J., & Furnham, A. (2021). A bibliometric analysis of H. J. Eysenck's research output: Clarifying controversy. *Personality and Individual Differences*, *169*, 109935. https://doi.org/10.1016/j.paid.2020.109935

Archambault, É., Campbell, D., Gingras, Y., & Larivière, V. (2009). Comparing bibliometric statistics obtained from the Web of Science and Scopus. *Journal of the American Society for Information Science and Technology*, *60*(7), 1320–1326. https://doi.org/10.1002/asi.21062

Aria, M., & Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, *11*(4), 959–975. https://doi.org/10.1016/j.joi.2017.08.007

Bach, D. R., Daunizeau, J., Friston, K. J., & Dolan, R. J. (2010). Dynamic causal modelling of anticipatory skin conductance responses. *Biological Psychology*, *85*(1), 163–170. https://doi.org/10.1016/j.biopsycho.2010.06.007

Baker, T. E., & Holroyd, C. B. (2011). Dissociated roles of the anterior cingulate cortex in reward and conflict processing as revealed by the feedback error-related negativity and N200. *Biological Psychology*, *87*(1), 25–34. https://doi.org/10.1016/j.biopsycho.2011.01.010

Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: An Open Source Software for Exploring and Manipulating Networks. *Proceedings of the International AAAI Conference on Web and Social Media*, *3*(1), 361–362. https://doi.org/10.1609/icwsm.v3i1.13937

Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, *10*(9), 1214–1221. https://doi.org/10.1038/nn1954

Bejjani, C., Tan, S., & Egner, T. (2020). Performance feedback promotes proactive but not reactive adaptation of conflict-control. *Journal of Experimental Psychology: Human Perception and Performance*, *46*(4), 369–387. https://doi.org/10.1037/xhp0000720

Bell, R., Sasse, J., Möller, M., Czernochowski, D., Mayr, S., & Buchner, A. (2016). Event-related potentials in response to cheating and cooperation in a social dilemma game. *Psychophysiology*, *53*(2), 216–228. https://doi.org/10.1111/psyp.12561

Bertolero, M. A., Dworkin, J. D., David, S. U., Lloreda, C. L., Srivastava, P., Stiso, J., Zhou, D., Dzirasa, K., Fair, D. A., Kaczkurkin, A. N., Marlin, B. J., Shohamy, D., Uddin, L. Q., Zurn, P., & Bassett, D. S. (2020). *Racial and ethnic imbalance in neuroscience reference lists and intersections with gender* (p. 2020.10.12.336230). bioRxiv. https://doi.org/10.1101/2020.10.12.336230

Biscaro, C., & Giupponi, C. (2014). Co-authorship and bibliographic coupling network effects on citations. *PloS One*, *9*(6), e99502. https://doi.org/10.1371/journal.pone.0099502

Bonacich, P. (1987). Power and Centrality: A Family of Measures. *American Journal of Sociology*, *92*(5), 1170–1182. https://doi.org/10.1086/228631

Bonacich, P. (2007). Some unique properties of eigenvector centrality. *Social Networks*, *29*(4), 555–564. https://doi.org/10.1016/j.socnet.2007.04.002

Borg, I., & Groenen, P. J. F. (2005). *Modern Multidimensional Scaling: Theory and Applications*. Springer.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*(3), 624–652. https://doi.org/10.1037/0033-295x.108.3.624

Bouton, M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin*, *114*(1), 80–99. https://doi.org/10.1037/0033-2909.114.1.80

Bouton, M. E. (2002). Context, ambiguity, and unlearning: sources of relapse after behavioral extinction. *Biological Psychiatry*, *52*(10), 976–986. https://doi.org/10.1016/s0006-3223(02)01546-9

Brandtstädter, J., & Greve, W. (1994). The Aging Self: Stabilizing and Protective Processes. *Developmental Review*, *14*(1), 52–80. https://doi.org/10.1006/drev.1994.1003

Carlson, J. M., Foti, D., Mujica-Parodi, L. R., Harmon-Jones, E., & Hajcak, G. (2011). Ventral striatal and medial prefrontal BOLD activation is correlated with reward-related electrocortical activity: a combined ERP and fMRI study. *NeuroImage*, *57*(4), 1608–1616. https://doi.org/10.1016/j.neuroimage.2011.05.037

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS Scales. *Journal of Personality and Social Psychology*, *67*(2), 319–333. https://doi.org/10.1037/0022-3514.67.2.319

Cavanagh, J. F., & Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends in Cognitive Sciences*, *18*(8), 414–421. https://doi.org/10.1016/j.tics.2014.04.012

Cavanagh, J. F., Frank, M. J., Klein, T. J., & Allen, J. J. B. (2010). Frontal theta links prediction errors to behavioral adaptation in reinforcement learning. *NeuroImage*, *49*(4), 3198–3209. https://doi.org/10.1016/j.neuroimage.2009.11.080

Chatterjee, P., & Werner, R. M. (2021). Gender Disparity in Citations in High-Impact Journal Articles. *JAMA Network Open*, *4*(7), e2114509. https://doi.org/10.1001/jamanetworkopen.2021.14509

Chekroud, A. M. (2015). Unifying treatments for depression: an application of the Free Energy Principle. *Frontiers in Psychology*, *6*, 153. https://doi.org/10.3389/fpsyg.2015.00153

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181–204. https://doi.org/10.1017/s0140525x12000477

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Lawrence Erlbaum Associates.

Craske, M. G., Hermans, D., & Vervliet, B. (2018). State-of-the-art and future directions for extinction as a translational model for fear and anxiety. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1742), 20170025. https://doi.org/10.1098/rstb.2017.0025

Craske, M. G., Treanor, M., Conway, C. C., Zbozinek, T., & Vervliet, B. (2014). Maximizing exposure therapy: an inhibitory learning approach. *Behaviour Research and Therapy*, *58*, 10–23. https://doi.org/10.1016/j.brat.2014.04.006

D'Astolfo, L., & Rief, W. (2017). Learning about expectation violation from prediction error paradigms - A meta-analysis on brain processes following a prediction error. *Frontiers in Psychology*, *8*, 1253. https://doi.org/10.3389/fpsyg.2017.01253

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, *69*(6), 1204–1215. https://doi.org/10.1016/j.neuron.2011.02.027

Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, *8*(12), 1704–1711. https://doi.org/10.1038/nn1560

de Bruijn, E. R. A., Hulstijn, W., Verkes, R. J., Ruigt, G. S. F., & Sabbe, B. G. C. (2004). Drug-induced stimulation and suppression of action monitoring in healthy volunteers. *Psychopharmacology*, *177*(1–2), 151–160. https://doi.org/10.1007/s00213-004-1915-6

de Lange, F. P., Heilbron, M., & Kok, P. (2018). How Do Expectations Shape Perception? *Trends in Cognitive Sciences*, *22*(9), 764–779. https://doi.org/10.1016/j.tics.2018.06.002

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Den Ouden, H. E. M., Kok, P., & De Lange, F. P. (2012). How Prediction Errors Shape Perception, Attention, and Motivation. *Frontiers in Psychology*, *3*, 548. https://doi.org/10.3389/fpsyg.2012.00548

Díaz-Mataix, L., Ruiz Martinez, R. C., Schafe, G. E., LeDoux, J. E., & Doyère, V. (2013). Detection of a temporal error triggers reconsolidation of amygdala-dependent memories. *Current Biology*, *23*(6), 467–472. https://doi.org/10.1016/j.cub.2013.01.053

Diodato, V. P., & Gellatly, P. (2013). *Dictionary of Bibliometrics*. Routledge. https://doi.org/10.4324/9780203714133

Doering, B. K., Glombiewski, J. A., & Rief, W. (2018). Expectation-Focused Psychotherapy to Improve Clinical Outcomes. *International Review of Neurobiology*, *138*, 257–270. https://doi.org/10.1016/bs.irn.2018.02.004

Dort, M., Strelow, A. E., French, B., Groom, M., Luman, M., Thorell, L. B., Biele, G., & Christiansen, H. (2020). Bibliometric Review: Classroom Management in ADHD—Is There a Communication Gap Concerning Knowledge Between the Scientific Fields Psychiatry/Psychology and Education? *Sustainability: Science Practice and Policy*, *12*(17), 6826. https://doi.org/10.3390/su12176826

Dort, M., Strelow, A. E., Schwinger, M., & Christiansen, H. (2020). Working with children with ADHD—A latent profile analysis of teachers' and psychotherapists' attitudes. *Sustainability: Science Practice and Policy*, *12*(22), 9691. https://doi.org/10.3390/su12229691

Dworkin, J. D., Linn, K. A., Teich, E. G., Zurn, P., Shinohara, R. T., & Bassett, D. S. (2020). The extent and drivers of gender imbalance in neuroscience reference lists. *Nature Neuroscience*, *23*(8), 918–926. https://doi.org/10.1038/s41593-020-0658-y

Falkenstein, M., Hohnsbein, J., Hoormann, J., & Blanke, L. (1991). Effects of crossmodal divided attention on late ERP components. II. Error processing in choice reaction tasks. *Electroencephalography and Clinical Neurophysiology*, *78*(6), 447–455. https://doi.org/10.1016/0013-4694(91)90062-9

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/bf03193146

Fiorillo, C. D., Tobler, P. N., & Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science*, *299*(5614), 1898–1902. https://doi.org/10.1126/science.1077349

Frank, M. J., Seeberger, L. C., & O'Reilly, R. C. (2004). By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science*, *306*(5703), 1940–1943. https://doi.org/10.1126/science.1102941

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *360*(1456), 815–836. https://doi.org/10.1098/rstb.2005.1622

Friston, K. (2009). The free-energy principle: a rough guide to the brain? *Trends in Cognitive Sciences*, *13*(7), 293–301. https://doi.org/10.1016/j.tics.2009.04.005

García Alanis, J. C., Baker, T. E., Peper, M., & Chavanon, M.-L. (2019). Social context effects on error-related brain activity are dependent on interpersonal and achievement-related traits. *Scientific Reports*, *9*(1), 1728. https://doi.org/10.1038/s41598-018-38417-2

Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E., & Donchin, E. (1993). A Neural System for Error Detection and Compensation. *Psychological Science*, *4*(6), 385–390. https://doi.org/10.1111/j.1467-9280.1993.tb00586.x

Gehring, W. J., & Willoughby, A. R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science*, *295*(5563), 2279–2282. https://doi.org/10.1126/science.1066893

Gershman, S. J. (2015). A Unifying Probabilistic View of Associative Learning. *PLoS Computational Biology*, *11*(11), e1004567. https://doi.org/10.1371/journal.pcbi.1004567

Gillan, C. M., Otto, A. R., Phelps, E. A., & Daw, N. D. (2015). Model-based learning protects against forming habits. *Cognitive, Affective, & Behavioral Neuroscience*, *15*(3), 523–536. https://doi.org/10.3758/s13415-015-0347-6

Golbeck, J. (2013). Network Structure and Measures. In J. Golbeck (Ed.), *Analyzing the Social Web* (pp. 25–44). Morgan Kaufmann. https://doi.org/10.1016/b978-0-12-405531-5.00003-1

Gollwitzer, M., Thorwart, A., & Meissner, K. (2018). Editorial: Psychological Responses to Violations of Expectations. *Frontiers in Psychology*, *8*, 2357. https://doi.org/10.3389/fpsyg.2017.02357

Gomez, C. J., Herman, A. C., & Parigi, P. (2022). Leading countries in global science increasingly receive more citations than other countries doing similar research. *Nature Human Behaviour*, *6*(7), 919–929. https://doi.org/10.1038/s41562-022-01351-5

Greve, A., Cooper, E., Kaula, A., Anderson, M. C., & Henson, R. (2017). Does prediction error drive one-shot declarative learning? *Journal of Memory and Language*, *94*, 149–165. https://doi.org/10.1016/j.jml.2016.11.001

Haaker, J., Gaburro, S., Sah, A., Gartmann, N., Lonsdorf, T. B., Meier, K., Singewald, N., Pape, H.-C., Morellini, F., & Kalisch, R. (2013). Single dose of L-dopa makes extinction memories context-independent and prevents the return of fear. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(26), E2428–E2436. https://doi.org/10.1073/pnas.1303061110

Haaker, J., Lonsdorf, T. B., & Kalisch, R. (2015). Effects of post-extinction l-DOPA administration on the spontaneous recovery and reinstatement of fear in a human fMRI study. *European Neuropsychopharmacology*, *25*(10), 1544–1555. https://doi.org/10.1016/j.euroneuro.2015.07.016

Hajcak, G. (2012). What we've learned from mistakes. *Current Directions in Psychological Science*, *21*(2), 101–106. https://doi.org/10.1177/0963721412436809

Hajcak, G., Moser, J. S., Holroyd, C. B., & Simons, R. F. (2007). It's worse than you thought: The feedback negativity and violations of reward prediction in gambling tasks. *Psychophysiology*, *44*(6), 905–912. https://doi.org/10.1111/j.1469-8986.2007.00567.x

Hamm, A. O., & Vaitl, D. (1996). Affective learning: Awareness and aversion. *Psychophysiology*, *33*(6), 698–710. https://doi.org/10.1111/j.1469-8986.1996.tb02366.x

Harris, L. T., & Fiske, S. T. (2010). Neural regions that underlie reinforcement learning are also active for social expectancy violations. *Social Neuroscience*, *5*(1), 76–91. https://doi.org/10.1080/17470910903135825

He, Z., Bolz, M., & Baillargeon, R. (2011). False-belief understanding in 2.5-year-olds: evidence for violation-of-expectation change-of-location and unexpected-contents tasks. *Developmental Science*, *15*(2), 181–193. https://doi.org/10.1111/j.1467-7687.2010.00980.x

Hohwy, J. (2017). Priors in perception: Top-down modulation, Bayesian perceptual learning rate, and prediction error minimization. *Consciousness and Cognition*, *47*, 75–85. https://doi.org/10.1016/j.concog.2016.09.004

Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*(4), 679–709. https://doi.org/10.1037/0033-295x.109.4.679

Holroyd, C. B., Pakzad-Vaezi, K. L., & Krigolson, O. E. (2008). The feedback correct-related positivity: sensitivity of the event-related brain potential to unexpected positive feedback. *Psychophysiology*, *45*(5), 688–697. https://doi.org/10.1111/j.1469-8986.2008.00668.x

Hoorens, V. (2012). Expectation. In V. S. Ramachandran (Ed.), *Encyclopedia of Human Behavior* (2nd ed., pp. 142–149). Elsevier Science. https://doi.org/10.1016/b978-0-12-375000-6.00163-4

Kim, S.-I. (2013). Neuroscientific model of motivational process. *Frontiers in Psychology*, *4*, 98. https://doi.org/10.3389/fpsyg.2013.00098

Kindt, M., Soeter, M., & Vervliet, B. (2009). Beyond extinction: erasing human fear responses and preventing the return of fear. *Nature Neuroscience*, *12*(3), 256–258. https://doi.org/10.1038/nn.2271

Korn, C. W., Prehn, K., Park, S. Q., Walter, H., & Heekeren, H. R. (2012). Positively biased processing of self-relevant social feedback. *The Journal of Neuroscience*, *32*(47), 16832–16844. https://doi.org/10.1523/jneurosci.3016-12.2012

Kotzur, P. F., & Wagner, U. (2021). The dynamic relationship between contact opportunities, positive and negative intergroup contact, and prejudice: A longitudinal investigation. *Journal of Personality and Social Psychology*, *120*(2), 418–442. https://doi.org/10.1037/pspi0000258

Kube, T., Rief, W., & Glombiewski, J. A. (2017). On the maintenance of expectations in major depression - investigating a neglected phenomenon. *Frontiers in Psychology*, *8*, 9. https://doi.org/10.3389/fpsyg.2017.00009

Kube, T., Rief, W., Gollwitzer, M., Gärtner, T., & Glombiewski, J. A. (2019). Why dysfunctional expectations in depression persist – Results from two experimental studies investigating cognitive immunization. *Psychological Medicine*, *49*(9), 1532–1544. https://doi.org/10.1017/s0033291718002106

Laferton, J. A. C., Kube, T., Salzmann, S., Auer, C. J., & Shedden-Mora, M. C. (2017). Patients' Expectations Regarding Medical Treatment: A Critical Review of Concepts and Their Assessment. *Frontiers in Psychology*, *8*, 233. https://doi.org/10.3389/fpsyg.2017.00233

Luque, D., López, F. J., Marco-Pallares, J., Càmara, E., & Rodríguez-Fornells, A. (2012). Feedback-related brain potential activity complies with basic assumptions of associative learning theory. *Journal of Cognitive Neuroscience*, *24*(4), 794–808. https://doi.org/10.1162/jocn_a_00145

McCain, K. W. (1990). Mapping authors in intellectual space: A technical overview. *Journal of the American Society for Information Science. American Society for Information Science*, *41*(6), 433–443. https://doi.org/10.1002/(sici)1097-4571(199009)41:6<433::aid-asi11>3.0.co;2-q

McKim, T. H., Bauer, D. J., & Boettiger, C. A. (2016). Addiction History Associates with the Propensity to Form Habits. *Journal of Cognitive Neuroscience*, *28*(7), 1024–1038. https://doi.org/10.1162/jocn_a_00953

Miltner, W. H. R., Braun, C. H., & Coles, M. G. H. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task: evidence for a "generic" neural system for error detection. *Journal of Cognitive Neuroscience*, *9*(6), 788–798. https://doi.org/10.1162/jocn.1997.9.6.788

Montague, P. R., Dayan, P., & Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *The Journal of Neuroscience*, *16*(5), 1936–1947. https://doi.org/10.1523/jneurosci.16-05-01936.1996

Moustafa, A. A., Gilbertson, M. W., Orr, S. P., Herzallah, M. M., Servatius, R. J., & Myers, C. E. (2013). A model of amygdala–hippocampal–prefrontal interaction in fear conditioning and extinction in animals. *Brain and Cognition*, *81*(1), 29–43. https://doi.org/10.1016/j.bandc.2012.10.005

Mueller, E. M., Burgdorf, C., Chavanon, M.-L., Schweiger, D., Wacker, J., & Stemmler, G. (2014). Dopamine modulates frontomedial failure processing of agentic introverts versus extraverts in incentive contexts. *Cognitive, Affective, & Behavioral Neuroscience*, *14*(2), 756–768. https://doi.org/10.3758/s13415-013-0228-9

Mullen, P. D., & Ramírez, G. (2006). The promise and pitfalls of systematic reviews. *Annual Review of Public Health*, *27*(1), 81–102. https://doi.org/10.1146/annurev.publhealth.27.021405.102239

Nader, K., Schafe, G. E., & Le Doux, J. E. (2000). Fear memories require protein synthesis in the amygdala for reconsolidation after retrieval. *Nature*, *406*(6797), 722–726. https://doi.org/10.1038/35021052

Newman, M. E. J., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, *69*(2 Pt 2), 026113. https://doi.org/10.1103/physreve.69.026113

O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H., & Dolan, R. J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron*, *38*(2), 329–337. https://doi.org/10.1016/s0896-6273(03)00169-7

Panitz, C., Endres, D., Buchholz, M., Khosrowtaj, Z., Sperl, M. F. J., Mueller, E. M., Schubö, A., Schütz, A. C., Teige-Mocigemba, S., & Pinquart, M. (2021). A revised framework for the investigation of expectation update versus maintenance in the context of expectation violations: The ViolEx 2.0 model. *Frontiers in Psychology*, *12*, 726432. https://doi.org/10.3389/fpsyg.2021.726432

Panitz, C., Sperl, M. F. J., Hennig, J., Klucken, T., Hermann, C., & Mueller, E. M. (2018). Fearfulness, neuroticism/anxiety, and COMT Val158Met in long-term fear conditioning and extinction. *Neurobiology of Learning and Memory*, *155*, 7–20. https://doi.org/10.1016/j.nlm.2018.06.001

Pasadeos, Y., Phelps, J., & Kim, B.-H. (1998). Disciplinary Impact of Advertising Scholars: Temporal Comparisons of Influential Authors, Works and Research Networks. *Journal of Advertising*, *27*(4), 53–70. https://doi.org/10.1080/00913367.1998.10673569

Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, *87*(6), 532–552. https://doi.org/10.1037/0033-295x.87.6.532

Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*, *442*(7106), 1042–1045. https://doi.org/10.1038/nature05051

Pfabigan, D. M., Alexopoulos, J., Bauer, H., & Sailer, U. (2011). Manipulation of feedback expectancy and valence induces negative and positive reward prediction error signals manifest in event-related brain potentials. *Psychophysiology*, *48*(5), 656–664. https://doi.org/10.1111/j.1469-8986.2010.01136.x

Pinquart, M., & Block, H. (2020). Coping with broken achievement-related expectations in students from elementary school: An experimental study. *International Journal of Developmental Science*, *14*(1–2), 9–17. https://doi.org/10.3233/dev-200001

Pinquart, M., Endres, D., Teige-Mocigemba, S., Panitz, C., & Schütz, A. C. (2021). Why expectations do or do not change after expectation violation: A comparison of seven models. *Consciousness and Cognition*, *89*, 103086. https://doi.org/10.1016/j.concog.2021.103086

Pinquart, M., Rothers, A., Gollwitzer, M., Khosrowtaj, Z., Pietzsch, M., & Panitz, C. (2021). Predictors of Coping with Expectation Violation: An Integrative Review. *Review of General Psychology: Journal of Division 1, of the American Psychological Association*, *25*(3), 321–333. https://doi.org/10.1177/10892680211024123

Pittig, A., & van den Berg, L. (2016). The key role of extinction learning in anxiety disorders: behavioral strategies to enhance exposure-based treatments. *Current Opinion in Psychiatry*, *29*(1), 39–47. https://doi.org/10.1097/yco.0000000000000220

Proulx, T., & Inzlicht, M. (2012). The five "A"s of meaning maintenance: Finding meaning in the theories of sense-making. *Psychological Inquiry*, *23*(4), 317–335. https://doi.org/10.1080/1047840x.2012.702372

Proulx, T., Inzlicht, M., & Harmon-Jones, E. (2012). Understanding all inconsistency compensation as a palliative response to violated expectations. *Trends in Cognitive Sciences*, *16*(5), 285–291. https://doi.org/10.1016/j.tics.2012.04.002

Proulx, T., Sleegers, W., & Tritt, S. M. (2017). The expectancy bias: Expectancy-violating faces evoke earlier pupillary dilation than neutral or negative faces. *Journal of Experimental Social Psychology*, *70*, 69–79. https://doi.org/10.1016/j.jesp.2016.12.003

R Core Team. (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Randles, D., Inzlicht, M., Proulx, T., Tullett, A. M., & Heine, S. J. (2015). Is dissonance reduction a special case of fluid compensation? Evidence that dissonant cognitions cause compensatory affirmation and abstraction. *Journal of Personality and Social Psychology*, *108*(5), 697–710. https://doi.org/10.1037/a0038933

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79–87. https://doi.org/10.1038/4580

Rapp, A. M., Grammer, J. K., Tan, P. Z., Gehring, W. J., Chavira, D. A., & Miller, G. A. (2021). Collectivism is associated with enhanced neural response to socially salient errors among adolescents. *Social Cognitive and Affective Neuroscience*, *16*(11), 1150–1159. https://doi.org/10.1093/scan/nsab065

Rescorla, R. A., & Wagner, A. R. (1972). A Theory of Pavlovian Conditioning: Variations in the Effectiveness of Reinforcement and Nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). Appleton-Century-Crofts.

Rief, W., & Glombiewski, J. A. (2016). Expectation-focused psychological interventions (EFPI). *Verhaltenstherapie*, *26*(1), 47–54. https://doi.org/10.1159/000442374

Rief, W., Glombiewski, J. A., Gollwitzer, M., Schubö, A., Schwarting, R., & Thorwart, A. (2015). Expectancies as core features of mental disorders. *Current Opinion in Psychiatry*, *28*(5), 378–385. https://doi.org/10.1097/yco.0000000000000184

Riels, K., Ramos Campagnoli, R., Thigpen, N., & Keil, A. (2022). Oscillatory brain activity links experience to expectancy during associative learning. *Psychophysiology*, *59*(5), e13946. https://doi.org/10.1111/psyp.13946

Roesch, M. R., Esber, G. R., Li, J., Daw, N. D., & Schoenbaum, G. (2012). Surprise! Neural correlates of Pearce-Hall and Rescorla-Wagner coexist within the brain. *European Journal of Neuroscience*, *35*(7), 1190–1200. https://doi.org/10.1111/j.1460-9568.2011.07986.x

Roese, N. J., & Sherman, J. W. (2007). Expectancy. In A. W. Kruglanski & E. T. Higgins (Eds.), *Social Psychology: Handbook of Basic Principles* (2nd ed., pp. 91–115). Guilford Press.

Sambrook, T. D., & Goslin, J. (2015). A neural reward prediction error revealed by a meta-analysis of ERPs using great grand averages. *Psychological Bulletin*, *141*(1), 213–235. https://doi.org/10.1037/bul0000006

San Martín, R. (2012). Event-related potential studies of outcome processing and feedback-guided learning. *Frontiers in Human Neuroscience*, *6*, 304. https://doi.org/10.3389/fnhum.2012.00304

Santesso, D. L., Evins, A. E., Frank, M. J., Schetter, E. C., Bogdan, R., & Pizzagalli, D. A. (2009). Single dose of a dopamine agonist impairs reinforcement learning in humans: evidence from event-related potentials and computational modeling of striatal-cortical function. *Human Brain Mapping*, *30*(7), 1963–1976. https://doi.org/10.1002/hbm.20642

Schiffer, A.-M., & Schubotz, R. I. (2011). Caudate nucleus signals for breaches of expectation in a movement observation paradigm. *Frontiers in Human Neuroscience*, *5*, 38. https://doi.org/10.3389/fnhum.2011.00038

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology*, *80*(1), 1–27. https://doi.org/10.1152/jn.1998.80.1.1

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*(5306), 1593–1599. https://doi.org/10.1126/science.275.5306.1593

Schultz, W., & Dickinson, A. (2000). Neuronal coding of prediction errors. *Annual Review of Neuroscience*, *23*(1), 473–500. https://doi.org/10.1146/annurev.neuro.23.1.473

Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2013). The expected value of control: an integrative theory of anterior cingulate cortex function. *Neuron*, *79*(2), 217–240. https://doi.org/10.1016/j.neuron.2013.07.007

Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for Information Science*, *24*(4), 265–269. https://doi.org/10.1002/asi.4630240406

Sommerville, J. A., Schmidt, M. F. H., Yun, J.-E., & Burns, M. (2013). The development of fairness expectations and prosocial behavior in the second year of life. *Infancy*, *18*(1), 40–66. https://doi.org/10.1111/j.1532-7078.2012.00129.x

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement Learning*: *An Introduction* (1st ed.). The MIT Press.

Träuble, B., Marinović, V., & Pauen, S. (2010). Early Theory of Mind Competencies: Do Infants Understand Others' Beliefs? *Infancy*, *15*(4), 434–444. https://doi.org/10.1111/j.1532-7078.2009.00025.x

Ullsperger, M., Danielmeier, C., & Jocham, G. (2014). Neurophysiology of performance monitoring and adaptive behavior. *Physiological Reviews*, *94*(1), 35–79. https://doi.org/10.1152/physrev.00041.2012

Ullsperger, M., Fischer, A. G., Nigbur, R., & Endrass, T. (2014). Neural mechanisms and temporal dynamics of performance monitoring. *Trends in Cognitive Sciences*, *18*(5), 259–267. https://doi.org/10.1016/j.tics.2014.02.009

Van de Cruys, S., Evers, K., Van der Hallen, R., Van Eylen, L., Boets, B., de-Wit, L., & Wagemans, J. (2014). Precise minds in uncertain worlds: Predictive coding in autism. *Psychological Review*, *121*(4), 649–675. https://doi.org/10.1037/a0037665

van Eck, N. J., & Waltman, L. (2009). How to normalize cooccurrence data? An analysis of some well-known similarity measures. *Journal of the American Society for Information Science and Technology*, *60*(8), 1635–1651. https://doi.org/10.1002/asi.21075

van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, *84*(2), 523–538. https://doi.org/10.1007/s11192-009-0146-3

van Eck, N. J., & Waltman, L. (2014). Visualizing Bibliometric Networks. In Y. Ding, R. Rousseau, & D. Wolfram (Eds.), *Measuring Scholarly Impact: Methods and Practice* (pp. 285–320). Springer International Publishing. https://doi.org/10.1007/978-3-319-10377-8_13

Walsh, M. M., & Anderson, J. R. (2014). Navigating complex decision spaces: Problems and paradigms in sequential choice. *Psychological Bulletin*, *140*(2), 466–486. https://doi.org/10.1037/a0033455

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070. https://doi.org/10.1037/0022-3514.54.6.1063

Weinberg, A., Dieterich, R., & Riesel, A. (2015). Error-related brain activity in the age of RDoC: A review of the literature. *International Journal of Psychophysiology*, *98*(2 Pt 2), 276–299. https://doi.org/10.1016/j.ijpsycho.2015.02.029

Weinberg, A., Meyer, A., Hale-Rude, E., Perlman, G., Kotov, R., Klein, D. N., & Hajcak, G. (2016). Error-related negativity (ERN) and sustained threat: Conceptual framework and empirical evaluation in an adolescent sample. *Psychophysiology*, *53*(3), 372–385. https://doi.org/10.1111/psyp.12538

Yon, D., de Lange, F. P., & Press, C. (2019). The Predictive Brain as a Stubborn Scientist. *Trends in Cognitive Sciences*, *23*(1), 6–8. https://doi.org/10.1016/j.tics.2018.10.003

Yu, A. J., & Dayan, P. (2005). Uncertainty, neuromodulation, and attention. *Neuron*, *46*(4), 681–692. https://doi.org/10.1016/j.neuron.2005.04.026

Zingora, T., Vezzali, L., & Graf, S. (2020). Stereotypes in the face of reality: Intergroup contact inconsistent with group stereotypes changes attitudes more than stereotype-consistent contact. *Group Processes & Intergroup Relations*, *24*(8), 1284–1305. https://doi.org/10.1177/1368430220946816

Zupic, I., & Čater, T. (2015). Bibliometric Methods in Management and Organization. *Organizational Research Methods*, *18*(3), 429–472. https://doi.org/10.1177/1094428114562629

Supplementary Materials

Peer Review History

Download: https://collabra.scholasticahq.com/article/73830-expectation-violations-expectation-change-and-expectation-persistence-the-scientific-landscape-as-revealed-by-bibliometric-network-analyses/attachment/154399.docx?auth_token=mZUKlQ7B7ASCPpSkELx2

Supplemental Materials

Download: https://collabra.scholasticahq.com/article/73830-expectation-violations-expectation-change-and-expectation-persistence-the-scientific-landscape-as-revealed-by-bibliometric-network-analyses/attachment/154400.pdf?auth_token=mZUKlQ7B7ASCPpSkELx2