Mapping the landscape of behavioral reinforcement learning research

Anna I. Thoma[1†], Florian Bolenz[1,2†], Kevin Tiede[1,3,4], Yujia Yang[1,5], Stefano Palminteri[6,7], Ralph Hertwig[1], Dirk U. Wulff[1,8]

[1]Max Planck Institute for Human Development, Berlin, Germany.
[2]Science of Intelligence, Berlin, Germany.
[3]University of Erfurt, Erfurt, Germany.
[4]Bernhard Nocht Institute for Tropical Medicine, Hamburg, Germany.
[5*]Humboldt Universität zu Berlin, Berlin, Germany.
[6] École Normale Supérieure, Paris, France.
[7]Institut National de la Santé et de la Recherche Médicale, Paris, France.
[8]Universität Basel, Basel, Switzerland.

†These authors contributed equally to this work. Correspondence should be addressed to Anna I. Thoma thoma@mpib-berlin.mpg.de

Abstract

As global research output increases, maintaining a comprehensive overview becomes challenging, particularly in high-volume research fields spanning multiple disciplines and research traditions. Behavioral reinforcement learning—an influential approach to understanding how people learn from interactions with the environment—is one such field. What characterizes this research landscape, and how is it interconnected? Here, we introduce a novel bibliometric approach—combining computational semantics, large language models, and clustering methods—to explore article clusters within behavioral reinforcement learning. Our analysis provides a comprehensive map of the field, highlighting broad lines of behavioral and neuroscientific research and documenting a wide array of interdisciplinary topics. We characterize research clusters by frequent research topics and methods, key journals, and publication timeline. Moreover, we examine the relationships between clusters and visualize the distribution of research topics and methods across the landscape. Finally, we discuss implications for facilitating scientific exchange and present an online tool for exploring the landscape.

With ever more academic articles being published each year [1–3], maintaining a comprehensive overview of a research field is becoming increasingly challenging. Overlooking relevant new work may result in duplicate efforts and inefficient resource allocation, while neglecting previous work may lead to an increasingly fragmented and siloed research landscape with diminished innovative potential [4]. Research at the intersection of multiple disciplines with a highly dynamic evolution of topics may be particularly challenging to navigate. A prime example of a research field at risk of fragmentation is behavioral reinforcement learning, which formalizes learning and decision-making processes in humans and animals as they learn about optimal behavior from feedback. Reinforcement learning is frequently used as an umbrella term for learning algorithms, paradigms, and for the diverse research field as a whole [5]. The field spans multiple disciplines, including psychology, neuroscience, economics, and psychiatry, and contributes to important applications in areas such as psychiatric treatments [6–8], public policy [9–11], and education [12, 13]. Research on behavioral reinforcement learning dates back to work on conditioning in the early 20th century [14, 15]; however, efforts have intensified considerably over recent decades (see Figure S1 in Supplementary Material B). Numerous research streams have emerged, investigating, for instance, the cognitive and neural bases [16–23], the development trajectories of reinforcement learning [24–27], or, very recently, the emerging reinforcement learning capabilities of large language models (LLMs) [28, 29]. Consequently, the research landscape of behavioral reinforcement learning has become complex and challenging to navigate. In this article, we introduce a novel, LLM-powered bibliometric approach and an interactive online tool to map this research landscape, characterize its composition, and explore its connections and thematic similarities.

A fragmented research landscape threatens cumulative scientific progress [30–32]. Although research on behavioral reinforcement learning has a strong history of advancing new academic and practical applications, there is concern that highly specific paradigms or computational approaches may limit the generalizability of findings across populations and contexts [33–35]. The long research history and interdisciplinarity of the field pose an additional challenge: Important terminology was only introduced in certain periods (e.g., 'computational psychiatry' in the 2010s; see [8, 36]), making it difficult to trace emerging trends back to their scientific roots before a certain terminology was popularized. Given that academic incentive structures often favor the proliferation of terminology and measures [31, 37], it is crucial to develop tools that facilitate research integration. Facilitating cumulative scientific progress through bibliometric analysis can lead to more efficient resource allocation, avoid redundancy, and help to develop overarching theoretical and methodological frameworks, ultimately encouraging replication across varied contexts.

Bibliometric analyses offer a powerful approach to evaluate and visualize hidden relationships in bibliographic data (for an overview of bibliometric analyses, see [38]) and can inform new hypotheses in fragmented fields [39]. Bibliometric tools such as the VOSviewer software [40] and bibliometrix R package [41] draw on different types of data, such as bibliographic coupling (i.e., whether two articles cite the same source) or co-occurrence metrics (e.g., of authors, citations, or words). Although these approaches have helped to map out thousands of research landscapes, integrating recent methodological advances can address some of their important limitations.

One limitation of past bibliometric approaches is that they typically rely on a single dimension to construct a research landscape, which may introduce biases. Consider a research landscape based solely on the co-occurrence of authors across articles: lines of research investigating similar topics may appear separated due to the differing timelines of researchers' careers. Similarly, citation-based research landscapes may miss emerging topics if an article has not yet been widely cited. Integrating multiple information dimensions into a single research landscape may generate more robust insights. A second limitation is that traditional bibliometric analyses often use simple co-occurrences of words to examine the distribution of research topics. While traditional natural language processing techniques detect basic semantic similarities, leveraging state-of-the-art LLMs offers more nuanced insights [42, 43]. LLMs can disambiguate the meaning of words based on context and uncover the semantic similarity among emerging topics even before their keywords frequently co-occur. In particular, in research environments in which terminologies are proliferating, merely analyzing word co-occurrences may not be enough to detect important similarities [37]. For instance, in reinforcement learning, researchers from different disciplines may habitually refer to structurally equivalent tasks by different names (e.g., multi-armed bandit, repeated choice task, partial-feedback paradigm). Using an LLM-based approach to evaluate the semantic content of articles can provide a deeper understanding of the research topics and methods in a research field.

We developed a novel bibliometric approach that combines LLMs, computational semantics, and clustering methods. Our approach integrates three information sources to compute the similarity between articles (see Figure 1): author details, cited references, and the semantic content extracted from article titles and abstracts. Applied to behavioral reinforcement learning, our analysis visualizes clusters of similar articles as research 'continents' and 'countries', highlighting important topics, methods, leading journals, and publication timelines. To explore opportunities for scientific exchange, we analyze the similarity profiles between article clusters and examine the distribution of research topics and methods across the landscape. Additionally, we complement our analyses with an interactive online tool (see online material at https://mpib.berlin/vFVqU) that allows researchers to explore

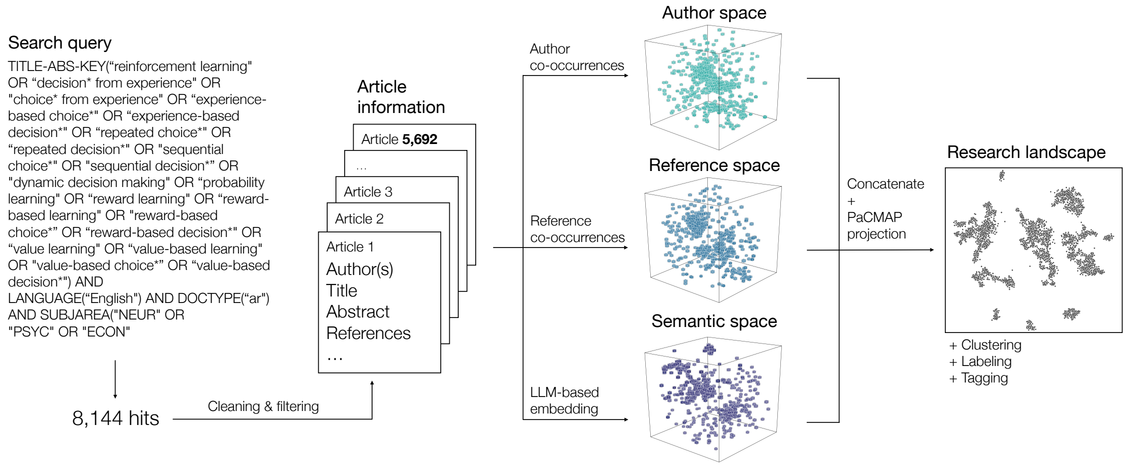

Figure 1 Overview of the bibliometric procedure used to map the landscape of behavioral reinforcement learning research. LLM, large language model; PaCMAP, Pairwise Controlled Manifold Approximation Projection.

the landscape in detail. Our analysis seeks to foster cross-disciplinary dialogue and integration within research on behavioral reinforcement learning and serves as a blueprint for similar efforts in other research fields.

1 Results

1.1 Mapping the research landscape

Our analysis approach to map the research landscape of behavioral reinforcement learning consisted of three stages (see Figure 1). First, we conducted a systematic literature search, employing a broad array of search terms to identify relevant work beyond exact matches of 'reinforcement learning' (see Methods). To focus on human learning and decision making and relevant animal models, we limited our search to the disciplines psychology, neuroscience, and economics, and applied a classifier to identify articles that were out of scope. Our final article database consisted of 5,692 articles published between 1970 and 2025. Second, we evaluated the similarity of each article pair in the database based on three information sources, resulting in three high-dimensional similarity spaces: authors, references, and semantic content. Specifically, we evaluated the similarity of authors and references based on co-occurrence frequencies and used LLMs to analyze the semantic similarity of articles' titles and abstracts. Third, we combined the three similarity spaces and projected them onto two dimensions using a non-linear dimension reduction algorithm to produce a map of the research landscape. Figure 2A presents the research landscape, where each point represents one article, and the distance between articles reflects their similarity. Articles closer together are generally more similar, while those further apart are less similar. However, due to the information loss and potential distortions in non-linear dimensionality reduction techniques, exact distances between articles should be interpreted with caution.

Finally, we performed hierarchical clustering to identify 30 research countries, grouped into ten research continents. We chose the number of research countries (i.e., the number of article clusters) to strike a balance between obtaining detailed insights and ensuring sufficient integration of different lines of research. Research countries define the diversity of the landscape on a more granular level, whereas continents represent broader lines of research. We employed LLMs to extract and label the main topics and methods as reported in the articles' titles and abstracts, here referred to as research 'tags'.

1.2 Overview of research continents

Like geographical continents, research continents differ in size and the number of countries (i.e., smaller, meaningful article clusters) they encompass. Five of the 10 research continents comprise multiple research countries, each separated by a white line (see Figure 2A and online materials for an

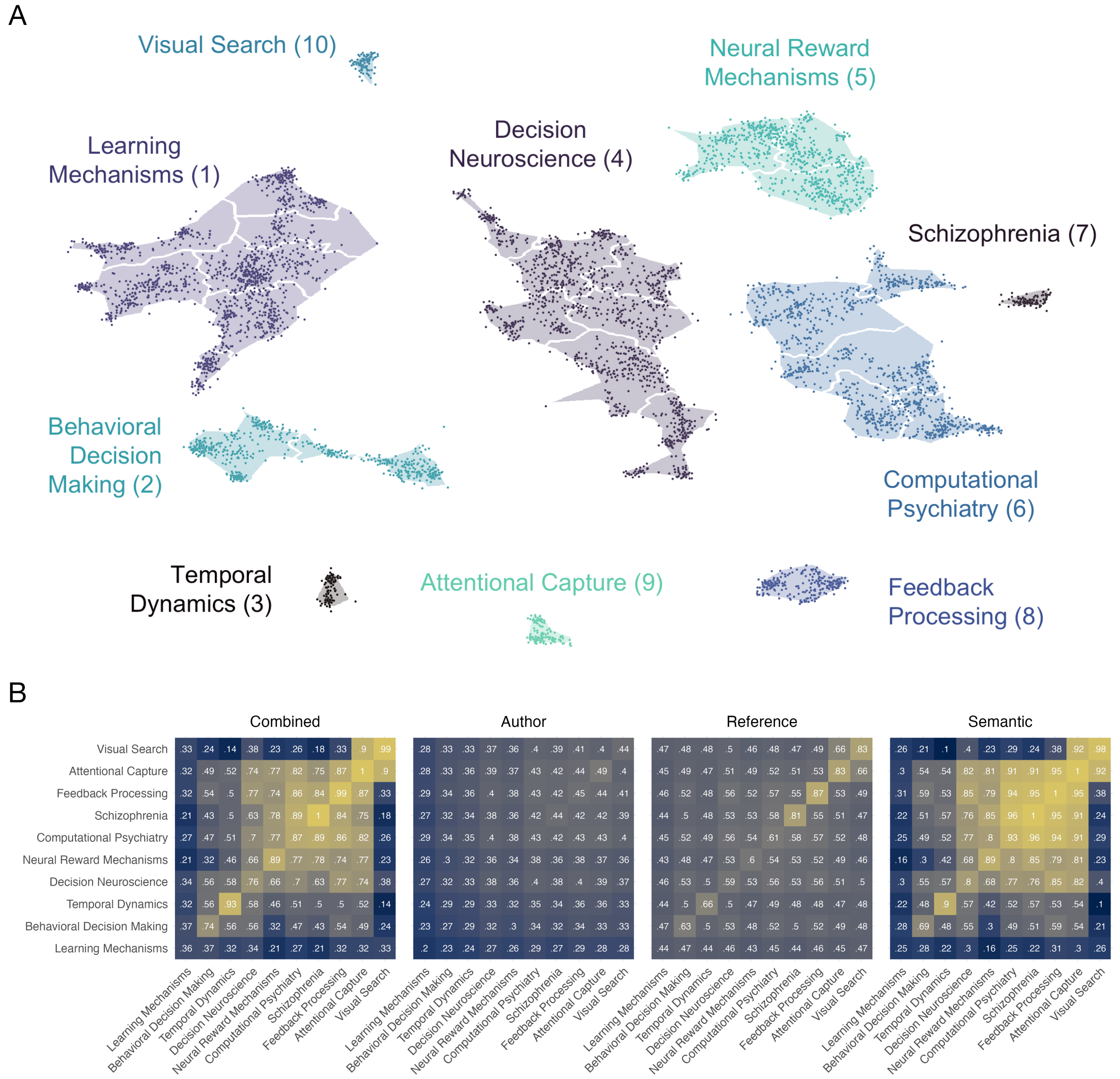

Figure 2 Research landscape and continent relations. Panel A displays the research landscape of behavioral reinforcement learning. Each point represents one article. Point placement reflects the aggregate similarity of articles with respect to their authors, cited references, and semantic content. Using hierarchical clustering methods, we grouped similar research articles into ten research continents which, in turn, break down into 30 research countries (separated by white lines). The distance between continents is shaped by their internal cohesion. Continent labels were assigned based on a manual inspection of their most informative research tags; the number in parentheses indicates a continent's numeric label. Continent colors do not reflect similarity. See online materials for an interactive version of the map. Panel B displays the proportion of high-similarity article pairs (i.e., pairs exceeding the respective median similarity) by similarity space. Values on the diagonal reflect the similarity of article pairs within continents (i.e., their internal cohesion); values off the diagonal reflect the similarity between continents. A value of 1 indicates that all article pairs are above the median similarity; a value of 0 indicates that no article pair is above the median similarity.

interactive version). The remaining five continents comprise a single country and appear as islands. Figure 2B displays the internal cohesion of research continents on the diagonals as the proportion of article pairs with strong similarities (i.e., exceeding the median similarity; see Methods for details). Across the aggregate similarity space and the author, reference, and semantic similarity spaces, single-country continents, on average, reveal a higher proportion of article pairs with strong similarity than continents with multiple countries. Yet, it is important to note that a high internal cohesion still allows for single-country continents to be interconnected with other continents. In the following, we provide an overview of the ten research continents, reporting key metrics (e.g., size and timeline) and

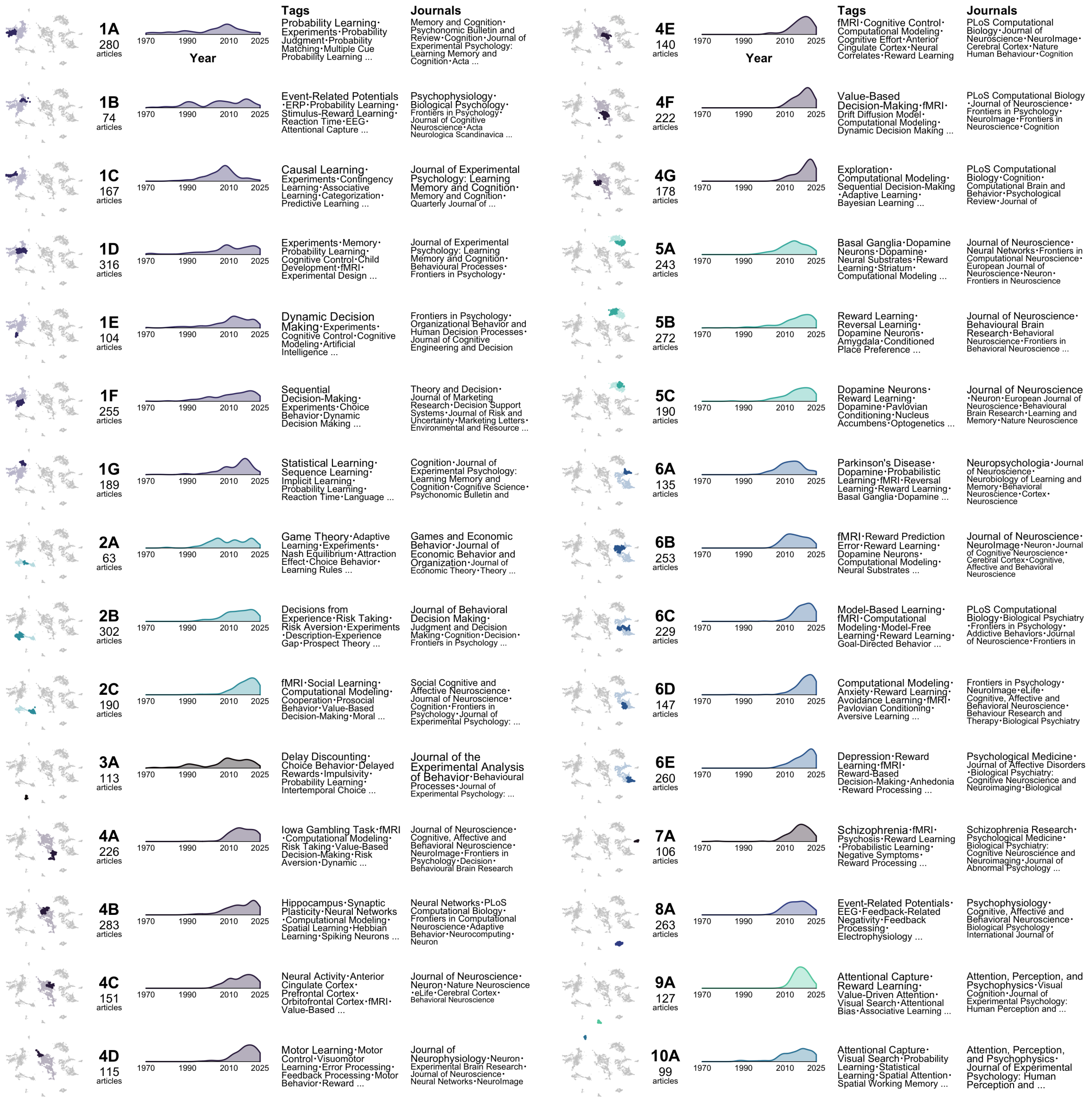

Figure 3 Overview of the 30 research countries. For each country, we present its location on the map, number of articles, publication timeline, key research tags, and journals. A country's alphanumeric label is printed in bold. Research tags are LLM-extracted topics and methods from articles' titles and abstracts. The font size of tags and journals reflects their relative frequency. Research countries grouped into the same continent appear in the same color.

important research tags (i.e., LLM-extracted topics and methods). Analogous to the description of research continents, Figure 3 displays detailed information on the 30 research countries.

The map reveals the importance of both behavioral and neuroscientific research traditions for behavioral reinforcement learning. Several continents on the left side of the map originate from historically behavioral lines of research, including Learning Mechanisms (1), Behavioral Decision Making (2), and Temporal Dynamics (3). The continent Learning Mechanisms (1) is one of the largest clusters. It hosts 1,385 articles with a median publication year of 2010; its earliest publications date back to 1970, making it one of the continents with the longest research history. It is rooted in classic experimental psychology, as documented in the frequency of tags referring to 'experiments' and 'experimental design' and in its coverage of different modes of learning (e.g., 'probability learning', 'causal learning', 'statistical learning', 'predictive learning', 'sequence learning', and 'sequential learning'). Furthermore, the tags 'dynamic decision making', 'sequential decision making', and 'choice behavior' reflect the continent's contributions to research on judgment and decision making.

A group of 555 articles forms the continent Behavioral Decision Making (2). Its earliest articles were published in 1980, with a median publication year of 2017. This continent hosts research from cognitive psychology and behavioral economics investigating 'decisions from experience'. Frequent tags include 'risk taking' and 'risk aversion', 'social learning', 'description–experience gap', 'game theory' and 'prospect theory', and commonly refer to methods such as 'experiments', 'fMRI', and 'computational modeling'.

The single-country continent Temporal Dynamics (3) comprises 113 articles dating back to 1973, with a median publication year of 2010. Research on this continent investigates temporal aspects of reinforcement learning, including 'delay discounting', 'delayed rewards', 'impulsivity', 'intertemporal choice', and varying reward schedules such as 'concurrent schedules' and 'variable-interval schedules'. Connecting to fundamental work on conditioning, research on this continent may utilize animal models, such as 'pigeons'.

In summary, research continents originating from behavioral research traditions are characterized by a focus on experimental design and different modes of learning, and represent the earliest lines of research in our article database.

Most continents at the center and on the right of the map originate from neuroscientific research traditions. Figure 2B suggests a high semantic similarity and moderate similarity of author and reference spaces between these continents. In other words, research from neuroscientific traditions seems to be closely interconnected. At the center, 1,315 articles create the continent Decision Neuroscience (4). Its median publication year is 2016, and its earliest articles date back to 1989. Articles located on this continent share methodological interests, for instance, using 'fMRI', and 'computational modeling', and investigate 'value-based decisions', 'reward learning' and 'cognitive control' in the context of 'neural activity' and 'neural correlates' and different brain areas (e.g., 'anterior cingulate cortex', 'prefrontal cortex').

In the top right corner, 705 articles establish the continent Neural Reward Mechanisms (5). The earliest publications date back to 1973, with a median publication year of 2016. The most commonly documented research tags on this continent combine neuroscientific topics—represented by a focus on 'dopamine' including 'dopamine neurons', 'dopamine signaling', 'basal ganglia', and the 'nucleus accumbens'—with behavioral and cognitive topics like 'reward learning', 'Pavlovian conditioning' and methods like 'computational modeling'.

In the center-right, 1,024 articles form the continent Computational Psychiatry (6), with a median publication year of 2018 and articles dating back to 1985—before the term 'computational psychiatry' was introduced. Common research tags cover clinical and neuroscientific aspects, such as 'depression', 'Parkinson's disease', and 'dopamine', methods like 'fMRI' and 'computational modeling', and behavioral and cognitive topics focusing on learning—for instance, 'reward learning', 'reward prediction error', 'probabilistic learning', 'model-based learning', and 'reversal learning'.

Among psychiatric disorders, one in particular seems to play a distinct role in reinforcement learning research: The single-country continent Schizophrenia (7) is located on the right side of the map. The earliest of its 106 articles dates back to 1977; its median publication year is 2018. Research articles clustering on this continent reveal a strong specialization in 'schizophrenia' (e.g., 'psychosis', 'negative symptoms', 'aberrant salience'), with aspects of learning including 'reward learning', 'probabilistic learning', and 'reward processing', and methods such as 'fMRI' and 'computational modeling'. Central scientific outlets include *Schizophrenia Research*, *Psychological Medicine*, and *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*. While this research may intuitively be viewed as part of research on Computational Psychiatry (6), the high internal cohesion between articles establishes a unique continent (see Figure 2B).

The bottom-right corner of the landscape hosts the single-country continent Feedback Processing (8). This cluster of 263 articles appears to have emerged relatively recently, with publications dating back to 2002 and a median publication year of 2016. Research on this continent combines neurophysiological and behavioral tags, such as 'event-related potentials', 'EEG', 'electrophysiology', 'feedback-related negativity', as well as 'reward prediction error' and 'reward positivity'.

Taken together, our analysis thus documents that neuroscientific and clinical research on reinforcement learning are closely intertwined. Articles highlight the role of dopamine and dopamine-related brain structures for reinforcement learning and emphasize the value of this research for understanding and treating psychiatric disorders.

The remaining two single-country continents share aspects from behavioral and neuroscientific research traditions, with a strong focus on psychophysiology and the visual system. The youngest single-country continent, Attentional Capture (9), is located in the center-bottom of the map. The 127 articles on this continent cluster closely together and were published from 2009 onward, with a median publication year of 2018. Research tags focus on 'value-driven attention', 'visual search', and 'attentional bias' in relation to 'reward learning', 'associative learning', and 'stimulus-reward learning' using 'fMRI' and 'experiments'.

At the top left of the map, 99 articles comprise the single-country continent Visual Search (10). Their earliest publication year is 1988, and their median publication year is 2017. Frequently reported research tags refer to 'spatial attention', 'spatial working memory', 'eye movements', and different modes of learning (e.g., 'probability learning', 'statistical learning', 'implicit learning', and 'spatial learning'). While there seems to be considerable overlap in the semantic and reference similarity of the last two continents, their high internal cohesion (i.e., high within-continent similarity) establishes two unique clusters of articles (see Figure 2B). Attentional Capture (9) shows a higher semantic similarity to neuroscientific continents than does Visual Search (10), suggesting a stronger overlap in research topics or methods.

Taken together, the landscape of behavioral reinforcement learning reflects the field's growing interdisciplinarity. We document influential—behavioral and neuroscientific—research traditions, trace the roots of important research areas, and identify a broad array of research topics and methods. The landscape reveals both thematic overlap between research countries and specialization in specific areas. To further investigate the composition of the landscape, we next analyze the connections between research countries.

1.3 Connections between research countries

Examining the connections between research countries across the landscape serves two goals. On the one hand, it documents the status quo of scientific exchange. On the other hand, it can help to identify promising areas for further advancing scientific exchange (e.g., expanding collaboration networks or integrating knowledge bases). To this end, we examined similarity profiles between research countries; that is, how pairs of countries compare on each of the three similarity spaces (i.e., authors, references, and semantic content).

For each similarity space, we evaluated the strength of connections between all research countries; Figure 4A displays the strongest 50% of all connections. Across the landscape, the community appears to actively engage in scientific exchange across authors, references, and semantic content, as indicated by a positive correlation between similarity spaces: Research countries with similar authors tend to cite similar references and to have similar semantic content (see also Figure 4B). However, the strength of connections between research countries varies by continent. For instance, whereas research countries on the Learning Mechanisms (1) continent display relatively few connections exceeding the 50% percentile, countries in Computational Psychiatry (6) are strongly interconnected across all three similarity spaces. An important factor limiting connectivity is the publication timeline. Recall that articles on Learning Mechanisms tend to have been published from earlier onward. In such cases, authors are less likely to have collaborated across extended periods (e.g., limited by retirement). Likewise, articles can only cite references available at the time of their publication, and changes in terminology, writing style, or methods used will inevitably lead to weaker connections in evolving research fields spanning longer time periods.

We propose two pathways by which our analysis may help advance scientific exchange between research countries. First, consider the scenario where multiple groups of researchers investigate similar topics and refer to similar references: In this case, there may be a shared research interest, and fostering new collaborations beyond existing author networks may lead to more efficient resource allocation and avoid redundant efforts. For instance, countries 2A and 2B on the continent Behavioral Decision Making (2) show considerable overlap in semantic content and cited references (see Figure 4B)—both countries seem to study 'decisions from experience'. However, examination of their key research tags and journals suggests that they originate from different disciplines: Whereas country 2A seems to investigate experience-based decision making from an economic perspective, country 2B seems to take a psychological perspective. In this case, entering new collaborations may help—for instance, to resolve methodological differences born in different disciplines, to elaborate on viewpoints, and to combine

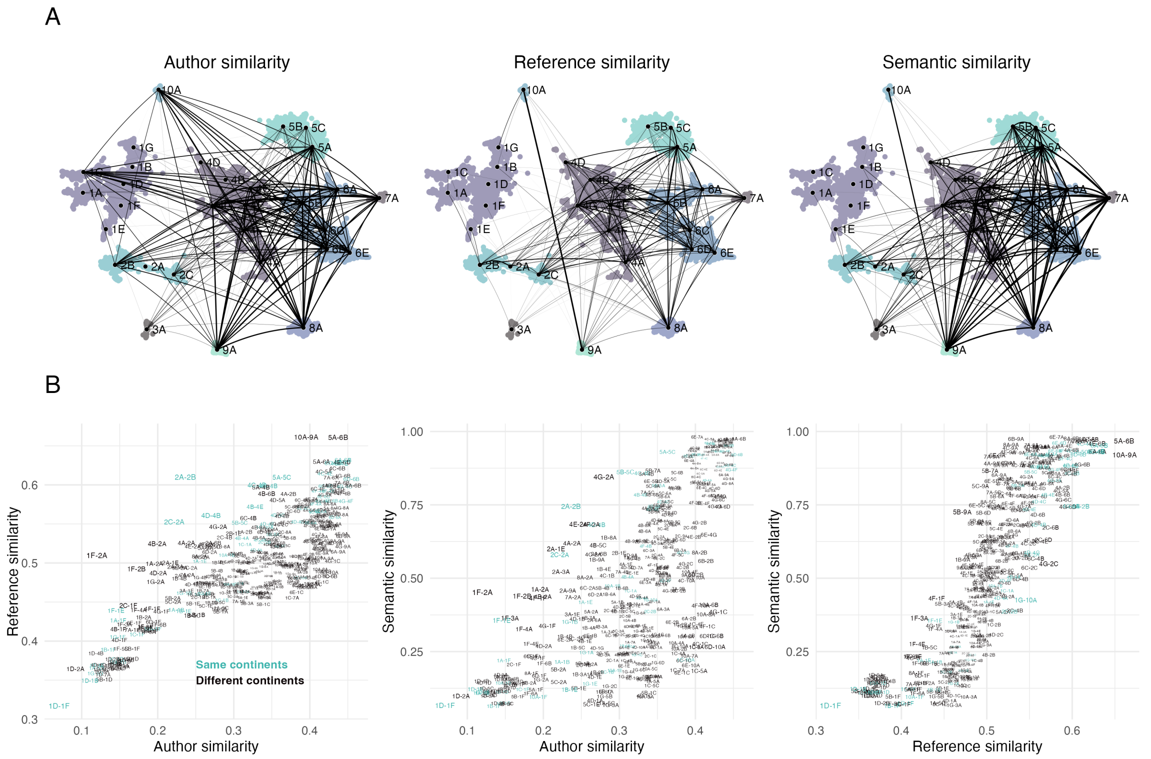

Figure 4 Connections between research countries. Panel A shows the proportion of strong connections between article pairs (i.e., article pairs exceeding the median similarity) in the author, reference, and semantic spaces between two research countries. Only the strongest 50% of connections are displayed. The stronger the line between two countries, the higher the proportion of strong connections. Panel B shows the relationships between the three similarity spaces. Each point represents the correlation of the cosine similarities between two research countries. Pairs printed in black are comparisons across continents; pairs printed in blue are from the same continent.

parallel efforts (see [44], including the peer commentary involving researchers from psychology and economics). One line of research that elaborates on methodological differences between disciplines is the description–experience gap, which describes how people's choice behavior differs depending on whether they learn about probabilistic information from written descriptions or from hands-on experience [45]. Research on the description–experience gap has uncovered systematic methodological differences between experimental psychology and behavioral economics that have hindered cumulative scientific progress since the early 1970s [46]. Increased collaboration between researchers from different backgrounds interested in similar topics may, therefore, prevent neighboring disciplines from making only incremental advances.

A second pathway to identifying promising areas for scientific exchange addresses the risk of disconnected lines of literature. When two research countries exhibit high semantic similarity but low reference similarity, the scientific community may benefit from integrating the literature to establish a more comprehensive knowledge basis. For instance, research country 5B on the Neural Reward Mechanisms continent reveals a high semantic similarity to country 7A on the Schizophrenia continent, but reference similarity scores are low. Examining their research tags suggests that the two countries may find common ground in dopamine-related structures and processes. However, whereas 5B seems to focus on basic processes and psychopharmacological interventions, 7A seems to specialize in research applied to psychiatric conditions, in particular schizophrenia and psychosis. In this case, it may be worth synthesizing the referenced literature to explore, for instance, new applications of basic research on the role of dopamine and psychopharmacological interventions to the treatment of severe psychiatric disorders.

In summary, our analysis documents considerable scientific exchange across the field of behavioral reinforcement learning. However, it also shows that the strengths of connections between research countries vary by continent, with time potentially being a driving factor. We propose that evaluating the relationship between similarity spaces can help to discover potential for collaboration between author networks and for the integration of knowledge bases.

1.4 Distribution of research tags

Another approach to fostering scientific exchange is by identifying major research topics and methods, and their occurrence across the landscape. To examine the thematic overlap in the field, we employed an LLM to extract the main research topics and methods from articles' titles and abstracts (see Methods for details). In contrast to simple word co-occurrences or evaluation of author keywords, this approach can help to identify interdisciplinary links when a common language has yet to be established within a research field or when terminology evolves over time.

Figure 5A displays the distribution of research tags assigned at least 20 times across articles as a function of their continent entropy and frequency of occurrence. Continent entropy reflects how equally or unequally a tag is distributed across the ten continents; a higher (lower) entropy indicates a more equal (unequal) spread across continents. This measure thus indicates whether a research tag is reported across the entire field or appears only in specialized parts of the landscape. Figure 5B displays a visual distribution of the 60 most frequent tags across the landscape.

The entropy and frequency of occurrence of a research tag represent its role and importance in the field of behavioral reinforcement learning; research tags that are relevant for all research continents have a high entropy and frequency. The tag 'reinforcement learning' occurs frequently and is distributed evenly across the landscape—validating the scope of our analysis. Similarly, 'computational modeling' represents a core methodology, and the inherently probabilistic nature of reinforcement learning is reflected in the tags 'feedback processing' and 'probabilistic learning'. Generally speaking, many high-frequency, high-entropy tags refer to cognitive or behavioral aspects of reinforcement learning. These tags may represent the common ground among research disciplines and serve as building blocks for unifying frameworks.

Research tags with medium or low entropy represent specialized topics or methods that may unlock new innovations when applied to different contexts. For instance, 'fMRI' appears to be an important methodological tool, but only in specific areas of the landscape; other examples are 'dopamine neurons', and larger brain areas such as the 'prefrontal cortex', 'anterior cingulate cortex', and 'ventral striatum'. In general, many neuroscientific tags have medium entropy, reflecting the specialization of the respective continents. Similarly, the historically grown emphasis on experimental methods is reflected in the stronger localization of the research tags 'experiments' and 'experimental design' on behavioral continents. Research tags of medium or lower entropy have the potential to identify shared interests across different areas, to suggest new links across disciplines, and uncover new applications.

Several research tags reveal a high semantic similarity, for instance, 'reward-based decision making' versus 'value-based decision making' or 'probabilistic learning' versus 'probability learning'. Whereas similar research tags may convey nuanced differences between distinct concepts, adopting these terms without critical reflection might further contribute to the fragmentation of the field. Our analysis suggests that different research traditions may historically favor certain terminologies, although they may be conceptually related. For instance, 'probability learning' and 'probabilistic learning' both originate from classic experimental psychology [47], but seem to be more strongly localized on behavioral and neuroscientific continents, respectively. While 'probability learning' is typically used to describe behavior in a specific paradigm, 'probabilistic learning' is more often used in the context of models or algorithms. Given the low connectedness between the research countries focusing on either terminology, but high conceptual relatedness between the research tags, this implies a risk that relevant research may be overlooked. As a compass that helps researchers navigate to and explore the distribution of related tags, the online tool allows the display of related research tags based on the semantic similarity of their embeddings (see online material at https://mpib.berlin/vFVqU).

Taken together, our LLM-based research tag analysis reveals considerable thematic overlap in the field. Assessing the entropy and distribution of research tags across the landscape can identify topics or methods that may help build overarching frameworks, establish new connections, and develop new applications. However, research tags may include semantic variations of similar concepts, which could contribute to the fragmentation of the field when used uncritically. Our online tool allows researchers to search for all 2,510 research tags identified as well as for other words using regular expressions.

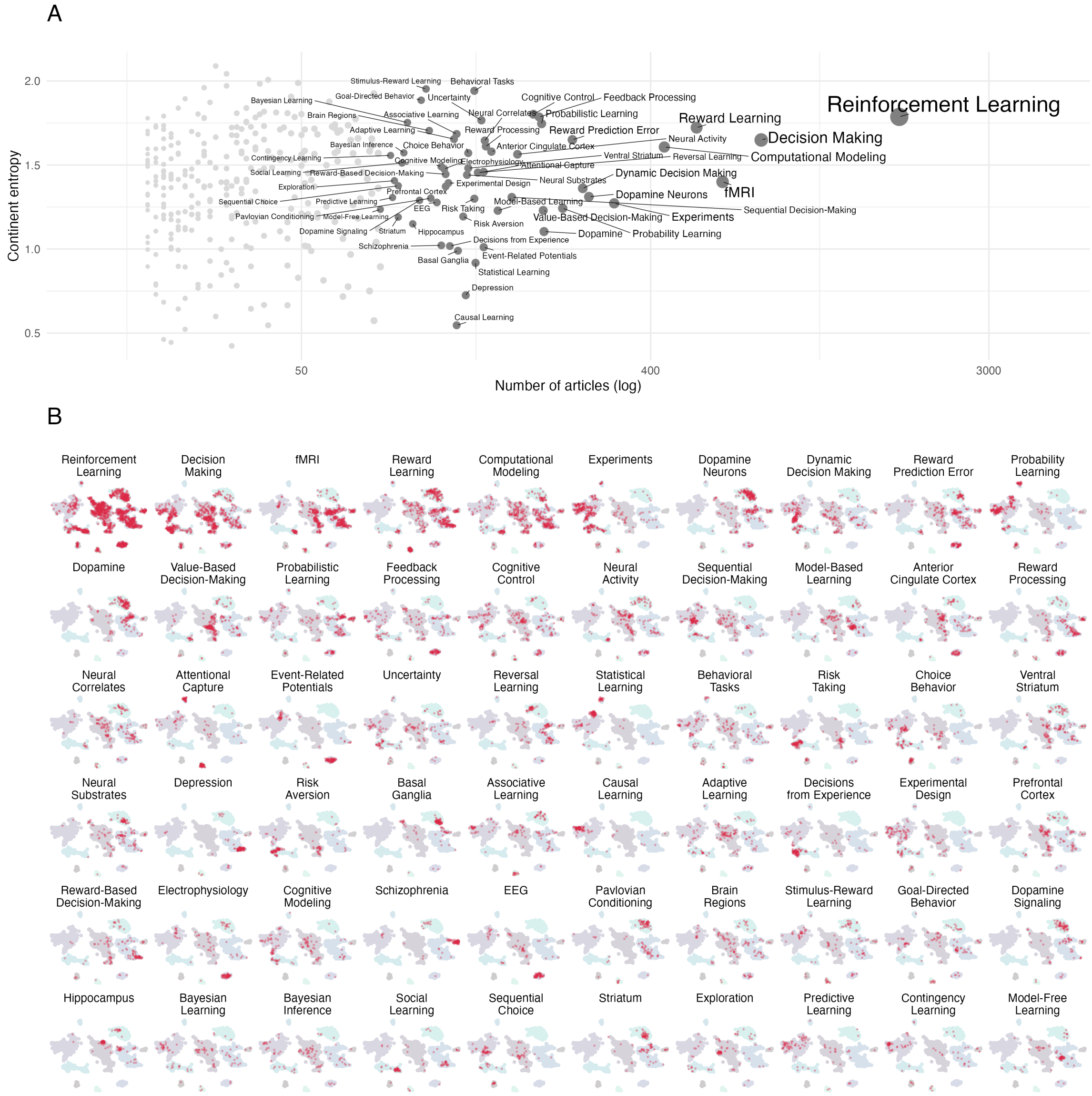

Figure 5 Distribution of research tags. Panel A displays the distribution of research tags assigned to at least 20 articles as a function of their continent entropy and number of occurrences; the 60 most frequent tags are labeled. Continent entropy indicates how evenly tags are spread across continents; higher (lower) values indicate a more (less) even distribution. Note that the frequency of occurrences on the x-axis is log-scaled. Panel B illustrates the distribution of the 60 most frequent tags across the research landscape; red points indicate articles including a tag in their title or abstract.

2 Discussion

We have presented a novel bibliometric approach to characterize and visualize research articles, applied to the field of behavioral reinforcement learning. Our analysis combines LLMs, computational semantics, and clustering methods to enable a multi-faceted characterization of research continents and countries (i.e., clusters of articles). The resulting landscape highlights the field's interdisciplinarity, documenting larger behavioral (e.g., Learning Mechanisms, Behavioral Decision Making) and neuroscientific research continents (e.g., Decision Neuroscience, Neural Reward Mechanisms, Computational Psychiatry). The neuroscientific continents, in particular, showed considerable thematic and collaborative overlap. Moreover, we explored how similarity profiles between research countries can uncover promising areas for scientific exchange. Examining the distribution of research tags (i.e., LLM-extracted topics and methods) uncovered thematic overlap—potentially serving as building blocks for unifying frameworks—and specialized topics—potentially inspiring new connections and applications. Our analyses are complemented by an interactive online tool that enables researchers to explore the landscape and to search for specific topics and methods.

Leveraging LLMs together with network and clustering techniques presents a significant opportunity to advance meta-scientific analyses. Traditional bibliometric tools often rely on co-occurrence frequencies of individual bibliometric dimensions, which can obscure conceptual overlap across article clusters. Yet, combining author, reference, and semantic content to construct a research landscape allows detailed insights into the strengths and shortcomings of scientific exchange in a research field. Our analysis serves as a testament to the ongoing exchange in behavioral reinforcement learning. However, we showed that comparing similarity profiles across multiple information dimensions can systematically assess the potential benefit of increased collaboration and integration of the literature in specific areas of the landscape. For instance, our analysis revealed article clusters that share a common research interest but have as yet forgone the opportunity to collaborate (e.g., cognitive psychology and behavioral economics) or build comprehensive knowledge bases (e.g., when applying basic research to psychiatric treatment). Similarly, our analysis reveals article clusters that appear more isolated from others due to their high internal cohesion, such as two small research continents focusing on visual attention. When article clusters are highly cohesive, in particular across multiple information dimensions, there may be a risk of research silos emerging when efforts for collaboration and theory integration wane.

Such risks may be amplified by proliferating research terminology, which is a growing concern across the cognitive and behavioral sciences [30, 32, 48]. Traditional bibliometric investigations that rely on the co-occurrence of words to evaluate semantic similarity may fail to capture thematic overlap, in particular, as research terminology proliferates. In contrast, we highlight the advantage of LLM-powered analyses to examine key research topics and methods. We identified a wide array of behavioral, cognitive, and neuroscientific research tags—with topics ranging from modes of learning, over cognitive processes, to neural correlates and psychopathology, and methodologies from behavioral experiments to computational modeling and neuroimaging. However, our results also document semantic variations of similar research topics (e.g., ‘probability learning’ and ‘probabilistic learning’) with distinct distributions on the landscape. It has been suggested that incentivization structures in academia favor the introduction of novel terminology for existing concepts [see 31], presenting a serious problem for cumulative scientific progress. Our analyses and online tool allow researchers to uncover semantic variations in terminology and to trace them to specific research continents or countries, ultimately facilitating a shared language.

Several limitations of our analyses warrant discussion. First, we acknowledge that the landscape presented is one map among many possible alternatives, each with a different organization of articles. Among other factors, the composition of the abstract database, the selection of search keywords, and the computational choices underlying the similarity spaces, including the dimensionality reduction algorithm, affect the resulting research landscape. Similar problems are common to bottom-up analyses involving clustering methods and apply to bibliometric analysis in general. Although differences between possible research landscapes will, for the most part, be superficial, the research landscape presented here should not be viewed as a canonical representation but as a heuristic tool. Similarly, we would like to emphasize the importance of carefully selecting and evaluating suitable prompts when instructing an LLM to generate research tags. LLM-based extractions typically require a trade-off between capitalizing on the LLM’s ability to synthesize topics and maintaining critical differences in terminology—this is particularly important in research fields that borrow terminology from different disciplines. Here, we addressed this trade-off by instructing the LLM to return tags in verbatim and synthesized tags based on a semantic-similarity threshold.

Second, we offer a static representation of the research landscape that does not explicitly account for temporal developments. Such analyses come with significant challenges, as removing articles from the database (e.g., filtering for specific publication years) will alter the cluster structure in ways that make comparisons across time difficult [49]. Another challenge arises from changes in language and terminology over time. For instance, tags such as ‘reinforcement learning’ and ‘computational modeling’ were not widely used in early research on behavioral reinforcement learning, which limits the possibilities for thematic overlap over time. Nevertheless, we want to emphasize that the risk of similar research clusters appearing disjointed due to their temporal dynamics is mitigated by using

LLM-based topic extraction and integrating multiple information sources into their characterization. Third, although our approach integrates multiple bibliometric dimensions, several possibilities for further extensions remain; for instance, entering co-citations of articles, the academic or geographic location of authors, or the semantic content of full texts. In particular, including full article texts may facilitate more detailed insights into methodological differences, as task design or computational models are usually presented in methods sections rather than in titles or abstracts. More broadly, our approach may serve as a blueprint that can be adapted to characterize other research fields or to explore, for instance, the diversity of topics within research institutions (i.e., based on institution-wide research output), journals, or conferences. LLM-based semantic analysis can also be tailored to specific goals—for instance, to prioritize the detection of semantic variations or harmonize research terminology.

In summary, we have introduced a novel LLM-powered bibliometric approach to visualize clusters of research articles and examine research landscapes. Applying this approach to the field of behavioral reinforcement learning, we highlighted its interdisciplinary nature, uncovering larger lines of research and thematic overlap, and proposing pathways to identify promising areas for enhanced scientific exchange. In particular, our analysis emphasizes that the field is larger than its parts explicitly focusing on reinforcement learning algorithms and traces its extensive research traditions. With our bibliometric approach, we aim to pave the way for future efforts to overcome fragmentation, develop overarching scientific agendas, and venture into novel domains.

3 Materials and Methods

The data and code necessary to reproduce the analyses are available online at https://osf.io/kxz9s/.

3.1 Systematic literature search

We conducted a systematic literature search to curate an article database on behavioral reinforcement learning. First, five authors discussed and defined a list of keywords from different lines of research related to behavioral reinforcement learning (e.g., 'reinforcement learning', 'decision*s from experience', 'reward learning'). Next, we used *Scopus* to retrieve articles written in English that contain one of these search terms in their titles, abstracts, or author keywords. We restricted our search to the disciplines psychology, neuroscience, and economics to maximize hits for empirical articles with human participants and relevant animal models, rather than research focusing on artificial agents. The search query is reported in Supplementary Material A. In May 2025, our systematic literature search resulted in 8,144 entries; for each entry, we extracted the title, abstract, author keywords, author names and author identifiers, publication year, referenced articles, and journal.

3.2 Database cleaning and filtering

We cleaned and filtered database entries before analysis. Editorial or publishing practices such as watermarks or text organizers can artificially inflate the semantic similarity of abstracts. To minimize this effect, we used regular expressions to remove watermarks (e.g., copyright information) and text organizers (e.g., 'Method', 'Background'). Additionally, we removed articles that were duplicates (12 articles), or that did not include an abstract (127 articles), references (379 articles), or author information (5 articles).

Manual inspection of the database revealed a large proportion of articles whose main research theme was unrelated to the scope of the current investigation (e.g., machine learning research with artificial agents). To ensure that human behavior was the main focus of articles in the database, we trained a classifier based on the thematic fit of a subset of manually labeled articles. To this end, we labeled 1001 articles as within or outside the scope of our current investigation. We then extracted the semantic embedding of each article's title and abstract using a sentence transformer model (all-MiniLM-L6-v2) [50] and submitted the embeddings as a predictor in a logistic ridge regression to evaluate the thematic fit of articles. Bootstrap cross-validation indicated an average out-of-sample accuracy of 96.0%. Applying the classifier to the full database resulted in 2,019 articles being classified as out of scope; these were removed from further analysis.

Our final database comprised 5,692 unique articles published from 1970 to 2025. Publication dates were significantly skewed, emphasizing the growing research interest in the field (see Figure S1 in

Supplementary Material B): The first quartile of articles was published from 1970 to 2010, the second from 2010 to 2016, the third from 2016 to 2021, and the fourth from 2021 to 2025. Relative to the total research output in the respective disciplines—inferred from removing the subject keywords from our Scopus search—research on behavioral reinforcement learning is increasing in volume: Articles in our database made up only 0.08% of the general research output in economics, neuroscience, and psychology in the decade from 1995 to 2004, increasing to 0.22% from 2005 to 2014 and to 0.27% from 2015 to 2024.

3.3 Mapping the landscape

A key challenge in visualizing and characterizing a research landscape is to identify meaningful similarities between articles. To offer a multi-faceted characterization of the research landscape, we evaluated article similarity across three bibliometric dimensions: authors, cited references, and the semantic content of titles and abstracts. For each dimension, we constructed a vector space encoding the similarity between articles, concatenated the spaces, and projected the aggregate similarity space onto a two-dimensional map.

To construct vector spaces encoding the similarity of authors and references, we employed an analysis pipeline inspired by latent semantic analysis [51]. First, we created an article–author and an article–reference frequency matrix, tabulating the frequencies of articles co-occurring with authors and references, respectively. Each row in these matrices represents one article, and each column one author or reference. The marginal distributions of authors and references are highly skewed, with many authors or references having only a few occurrences. To avoid the long tails of these distributions receiving undue importance, we removed authors with only one occurrence and references with less than five occurrences [52]. Next, we normalized frequencies using positive point-wise mutual information to control for possible effects of the length of author and reference lists, and for highly frequent authors or references [52]. To match the dimensionality of the semantic vector space—which is determined by the architecture of an LLM—we used singular decomposition to extract 384-dimensional vector spaces, encoding author and reference similarity, respectively.

To determine the semantic vector space, we capitalized on the nuanced understanding of a state-of-the-art open LLM—Llama-3.3-70B [53]—to fine-tune a small off-the-shelf embedding model (all-MiniLM-L6-v2). Specifically, we prompted Llama-3.3-70B to assess the similarity between 50,000 semi-randomly selected article pairs (see Supplementary Material C for the LLM prompt). Our sampling strategy to select the 50,000 pairs prioritized high-similarity pairs, which were initially determined by all-MiniLM-L6-v2. This sampling strategy provides more opportunities for the state-of-the-art LLM to learn the characteristics of high-similarity article pairs, in contrast to a fully random selection, which may comprise a high proportion of low-similarity pairs. To select the pairs based on the initial embeddings, we clustered the cosine similarity matrix of all articles into 50 clusters and sampled 42% from within and 58% from between clusters. Finally, we used the similarity assessment from the state-of-the-art LLM to fine-tune the small embedding model. We used the fine-tuned embedding model to extract semantic embeddings from articles' titles and abstracts, yielding a 384-dimensional semantic vector space that encodes the semantic similarity between articles. Compared to exclusively relying on similarity ratings by a state-of-the-art LLM for all article pairs, our analysis approach capitalizes on the LLM's refined understanding of behavioral reinforcement learning research while maintaining the computational feasibility of using an embedding model.

Next, we combined the three separate vector spaces—representing similarity between authors, references, and semantic content—into one space, encoding their combined similarity. Specifically, we concatenated the three vector spaces, yielding 1,152 dimensions, and projected this space into two dimensions using PaCMAP [54]. PaCMAP is a non-linear dimensionality reduction algorithm that preserves the global similarity structure better than other algorithms; we prioritized community structure and selected parameters accordingly (i.e., number of neighbors = 50, mid-near-pair ratio = 3, far-pair ratio = 10.0). In the final two-dimensional space, we applied hierarchical clustering using the complete linkage criterion to identify 30 research countries. These countries consist of subsets of articles that form cohesive groups based on their combined similarity across three information sources: authors, references, and semantic content of titles and abstracts. The number of research countries was selected by balancing the need for sufficient granularity to detect differences between research countries with keeping the overall complexity of the research landscape manageable. Finally, we combined clusters within the same spatial shape into ten research continents.

3.4 Quantifying connections between continents and countries

To quantify the author, reference, and semantic connections between continents and research countries, we computed cosine similarities between articles for each vector space as presented in Figure 4. Additionally, we calculated the connection between research countries as the proportion of articles above the median similarity of the respective dimension as presented in Figure 2.

3.5 Identifying research tags

To generate research tags from article titles and abstracts, we combined LLM-based topic extraction with clustering methods. First, we instructed an LLM (Llama-3.3-70b-Instruct) to generate three to seven taxonomic tags based on an article's title and abstract (see Supplementary Material D for the LLM prompt). To yield high-quality tags, we prompted the LLM to balance articles' topics and methods, to use the article's terminology, to avoid redundancy, to use a maximum of three words per tag, and to follow chain-of-thought reasoning. The LLM returned a total of 13,440 tags, many of which contained semantic variations of the same word (e.g., 'fMRI' and 'functional MRI'). These semantic variations make it difficult to differentiate between relevant differences and irrelevant semantic variations. Therefore, we combined semantically equivalent taxonomic tags by extracting their semantic embeddings using SFR-Embedding-Mistral [55] and joining tags when their cosine similarity was larger than the 95th similarity quantile. For each group of joined tags, we selected the most frequently occurring taxonomic tag as its label. After clustering, 2,510 unique tags remained. The 60 most frequent tags are displayed in Figure 5; the others can be explored in the online tool (see online material at https://mpib.berlin/vFVqU).

4 Acknowledgments

We thank Susannah Goss for editorial assistance and Maik Messerschmidt for programming support. F.B. was funded by the German Research Foundation (DFG) under Germany's Excellence Strategy – EXC 2002/1 "Science of Intelligence" – project number 390523135. S.P. acknowledges funding by the Alexander von Humboldt Foundation (Bessel Prize) and the European Research Council (Consolidator Grant: 101043804). D.W. acknowledges funding from the German Research Foundation (project number 546419617).

5 Author contributions

A.I.T. and F.B. contributed equally to this work. All authors conceptualized the study. A.I.T., F.B., and D.U.W. designed the methodology, prepared the database, and wrote the original draft. D.U.W. performed the analyses. All authors revised the manuscript.

References

1. Larsen, P. & Von Ins, M. The rate of growth in scientific publication and the decline in coverage provided by Science Citation Index. *Scientometrics* **84,** 575–603 (2010).
2. Bornmann, L., Haunschild, R. & Mutz, R. Growth rates of modern science: a latent piecewise growth curve approach to model publication numbers from established and new literature databases. *Humanities and Social Sciences Communications* **8,** 1–15 (2021).
3. Bornmann, L. & Mutz, R. Growth rates of modern science: A bibliometric analysis based on the number of publications and cited references. *Journal of the association for information science and technology* **66,** 2215–2222 (2015).
4. Park, M., Leahey, E. & Funk, R. J. Papers and patents are becoming less disruptive over time. *Nature* **613,** 138–144 (2023).
5. Sutton, R. S. & Barto, A. G. *Reinforcement learning: An introduction* 2nd ed. (ed Bach, F.) (MIT Press, Cambridge, MA, USA, 2018).

6. Maia, T. V. & Frank, M. J. From reinforcement learning models to psychiatric and neurological disorders. *Nature Neuroscience* **14,** 154–162 (2011).
7. Huys, Q. J., Maia, T. V. & Frank, M. J. Computational psychiatry as a bridge from neuroscience to clinical applications. *Nature neuroscience* **19,** 404–413 (2016).
8. Montague, P. R., Dolan, R. J., Friston, K. J. & Dayan, P. Computational psychiatry. *Trends in Cognitive Sciences* **16,** 72–80 (2012).
9. Hertwig, R. & Wulff, D. U. A description–experience framework of the psychology of risk. *Perspectives on Psychological Science* **17,** 631–651 (2022).
10. Palminteri, S. Human reinforcement learning processes and biases: computational characterization and possible applications to behavioral public policy. *Mind & Society* (2025).
11. Schulz, L. & Bhui, R. Political reinforcement learners. *Trends in Cognitive Sciences* **28,** 210–222 (2024).
12. Memarian, B. & Doleck, T. A scoping review of reinforcement learning in education. *Computers and Education Open* **6,** 100175 (2024).
13. Riedmann, A., Schaper, P. & Lugrin, B. Reinforcement Learning in Education: A Systematic Literature Review. *International Journal of Artificial Intelligence in Education,* 1–55 (2025).
14. Skinner, B. F. *The behavior of organisms: an experimental analysis* (Appleton-Century, Oxford, England, 1938).
15. Pavlov (1927), P. I. Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex. *Annals of Neurosciences* **17,** 136–141 (2010).
16. Fontanesi, L., Gluth, S., Spektor, M. S. & Rieskamp, J. A reinforcement learning diffusion decision model for value-based decisions. *Psychonomic Bulletin & Review* **26,** 1099–1121 (2019).
17. Palminteri, S. & Lebreton, M. The computational roots of positivity and confirmation biases in reinforcement learning. *Trends in cognitive sciences* **26,** 607–621 (2022).
18. Schultz, W. & Dickinson, A. Neuronal coding of prediction errors. *Annual review of neuroscience* **23,** 473–500 (2000).
19. Radulescu, A., Niv, Y. & Ballard, I. Holistic reinforcement learning: the role of structure and attention. *Trends in cognitive sciences* **23,** 278–292 (2019).
20. Bein, O. & Niv, Y. Schemas, reinforcement learning and the medial prefrontal cortex. *Nature Reviews Neuroscience* **26,** 141–157 (2025).
21. Niv, Y. Reinforcement learning in the brain. *Journal of Mathematical Psychology* **53,** 139–154 (2009).
22. Dayan, P. & Daw, N. D. Decision theory, reinforcement learning, and the brain. *Cognitive, Affective, & Behavioral Neuroscience* **8,** 429–453 (2008).
23. Lee, D., Seo, H. & Jung, M. W. Neural basis of reinforcement learning and decision making. *Annual review of neuroscience* **35,** 287–308 (2012).
24. Nussenbaum, K. & Hartley, C. A. Reinforcement learning across development: What insights can we draw from a decade of research? *Developmental Cognitive Neuroscience* **40,** 100733 (2019).
25. Thoma, A. I., Newell, B. R. & Schulze, C. Emerging adaptivity in probability learning: How young minds and the environment interact. *Journal of Experimental Psychology. General* **154,** 1523–1544 (2025).
26. Bolenz, F., Reiter, A. M. F. & Eppinger, B. Developmental changes in learning: computational mechanisms and social influences. *Frontiers in Psychology* **8,** 2048 (2017).
27. Decker, J. H., Otto, A. R., Daw, N. D. & Hartley, C. A. From creatures of habit to goal-directed learners: Tracking the developmental emergence of model-based reinforcement learning. *Psychological science* **27,** 848–858 (2016).
28. Hayes, W. M., Yax, N. & Palminteri, S. Relative value encoding in large language models: A multi-task, multi-model investigation. *Open Mind* **9,** 709–725 (2025).
29. Schubert, J. A., Jagadish, A. K., Binz, M. & Schulz, E. *In-context learning agents are asymmetric belief updaters* 2024.
30. Anvari, F. *et al.* Defragmenting psychology. *Nature Human Behaviour* **9,** 836–839 (2025).
31. Mischel, W. The toothbrush problem. *APS Observer* **21.** https://www.psychologicalscience.org/observer/the-toothbrush-problem (2008).
32. Wulff, D. U. & Mata, R. *Escaping the jingle–jangle jungle: Increasing conceptual clarity in psychology using large language models* 2025.

33. Eckstein, M. K. *et al.* The interpretation of computational model parameters depends on the context. *eLife* **11,** e75474 (2022).
34. Witte, K., Thalmann, M. & Schulz, E. Model-based exploration is measurable across tasks but not linked to personality and psychiatric assessments. *Scientific Reports* **15,** 27479 (2025).
35. Eckstein, M. K., Wilbrecht, L. & Collins, A. G. What do reinforcement learning models measure? Interpreting model parameters in cognition and neuroscience. *Current opinion in behavioral sciences* **41,** 128–137 (2021).
36. Rangel, A., Camerer, C. & Montague, P. R. A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience* **9,** 545–556 (2008).
37. Wulff, D. U. & Mata, R. Semantic embeddings reveal and address taxonomic incommensurability in psychological measurement. *Nature Human Behaviour* **9,** 944–954 (2025).
38. Donthu, N., Kumar, S., Mukherjee, D., Pandey, N. & Lim, W. M. How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research* **133,** 285–296 (2021).
39. Sabat, M., de Dampierre, C. & Tallon-Baudry, C. Evidence for domain-general arousal from semantic and neuroimaging meta-analyses reconciles opposing views on arousal. *Proceedings of the National Academy of Sciences* **122,** e2413808122 (2025).
40. Van Eck, N. J. & Waltman, L. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics* **84,** 523–538 (2010).
41. Aria, M. & Cuccurullo, C. bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics* **11,** 959–975 (2017).
42. Debelak, R., Koch, T. K., Aßenmacher, M. & Stachl, C. From embeddings to explainability: A tutorial on large-language-model-based text analysis for behavioral scientists. *Advances in Methods and Practices in Psychological Science* **8,** 25152459251351285 (2025).
43. Hussain, Z., Binz, M., Mata, R. & Wulff, D. U. A tutorial on open-source large language models for behavioral science. *Behavior Research Methods* **56,** 8214–8237 (2024).
44. Hertwig, R. & Ortmann, A. Experimental practices in economics: A methodological challenge for psychologists? *Behavioral and brain sciences* **24,** 383–403 (2001).
45. Wulff, D. U., Mergenthaler-Canseco, M. & Hertwig, R. A meta-analytic review of two modes of learning and the description-experience gap. *Psychological Bulletin* **144,** 140–176 (2018).
46. Lejarraga, T. & Hertwig, R. How experimental methods shaped views on human competence and rationality. *Psychological Bulletin* **147,** 535–564 (2021).
47. Estes, W. K. Theory of learning with constant, variable, or contingent probabilities of reinforcement. *Psychometrika* **22,** 113–132 (1957).
48. Núñez, R. *et al.* What happened to cognitive science? *Nature Human Behaviour* **3,** 782–791 (2019).
49. Wulff, D. U., Kieslich, P., Henninger, F., Haslbeck, J. & Schulte-Mecklenbeck, M. *Movement tracking of psychological processes: A tutorial using mousetrap* 2023.
50. Reimers, N. & Gurevych, I. *Sentence-BERT: Sentence embeddings using Siamese BERT-networks* in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)* (Association for Computational Linguistics, Hong Kong, China, 2019), 3980–3990.
51. Landauer, T. K., Foltz, P. W. & Laham, D. An introduction to latent semantic analysis. *Discourse Processes* **25,** 259–284 (1998).
52. Bullinaria, J. A. & Levy, J. P. Extracting semantic representations from word co-occurrence statistics: A computational study. *Behavior Research Methods* **39,** 510–526 (2007).
53. Grattafiori, A. *et al.* *The Llama 3 herd of models* Nov. 2024. http://arxiv.org/abs/2407.21783.
54. Wang, Y., Huang, H., Rudin, C. & Shaposhnik, Y. Understanding how dimension reduction tools work: An empirical approach to deciphering t-SNE, UMAP, TriMAP, and PaCMAP for data visualization. *Journal of Machine Learning Research* **22,** 1–73 (2021).
55. Meng, R. *et al.* *SFR-Embedding-Mistral: Enhance text retrieval with transfer learning* Published: Salesforce AI Research Blog. 2024. https://www.salesforce.com/blog/sfr-embedding/.

Mapping the landscape of behavioral reinforcement learning research

Anna I. Thoma[1†], Florian Bolenz[1,2†], Kevin Tiede[1,3,4], Yujia Yang[1,5], Stefano Palminteri[6,7], Ralph Hertwig[1], Dirk U. Wulff[1,8]

[1]Max Planck Institute for Human Development, Berlin, Germany.
[2]Science of Intelligence, Berlin, Germany.
[3]University of Erfurt, Erfurt, Germany.
[4]Bernhard Nocht Institute for Tropical Medicine, Hamburg, Germany.
[5*]Humboldt Universität zu Berlin, Berlin, Germany.
[6] École Normale Supérieure, Paris, France.
[7]Institut National de la Santé et de la Recherche Médicale, Paris, France.
[8]Universität Basel, Basel, Switzerland.

[†]These authors contributed equally to this work.

1 A - Search query

For the systematic literature search on Scopus, we used the following search query: 'TITLE-ABS-KEY("reinforcement learning" OR "decision* from experience" OR "choice* from experience" OR "experience-based choice*" OR "experience-based decision*" OR "repeated choice*" OR "repeated decision*" OR "sequential choice*" OR "sequential decision*" OR "dynamic decision making" OR "probability learning" OR "reward learning" OR "reward-based learning" OR "reward-based choice*" OR "reward-based decision*" OR "value learning" OR "value-based learning" OR "value-based choice*" OR "value-based decision*") AND LANGUAGE("English") AND DOCTYPE("ar") AND SUBJAREA("NEUR" OR "PSYC" OR "ECON")'

2 B - Histogram of articles across time

Figure 1 presents the distribution of articles in our database across publication years from 1970 until May 2025. To account for articles that will be published later in 2025, we extrapolated their expected number.

3 C - Semantic ratings

We used the following prompts to generate semantic relatedness ratings from Llama-3.3-70B-Instruct [1]. In place of $\{pair\}$, we pasted two articles with their titles and abstracts.

3.1 System prompt

You are an expert in the academic literature on behavioral reinforcement learning, who accurately discerns differences in specific research topics.

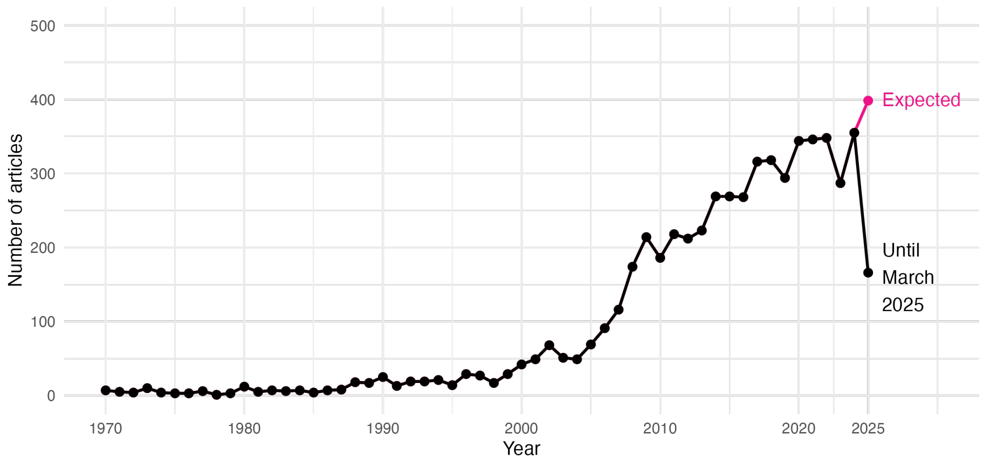

Figure 1 Distribution of research articles across publication years. The black line shows actual counts. The pink shows the expected number of articles extrapolating from May 2025 until the end of the year.

3.2 User prompt

Your primary task is to compare the following two articles (Article 1 and Article 2) based *only* on their provided titles and abstracts. Both articles operate within the general field of behavioral reinforcement learning (BRL).goal is to determine how similar their *specific research topics* are within the BRL context. Do they investigate the same sub-problem, mechanism, or research question?

Here are the articles to evaluate:pair.

First, provide your reasoning: **Reasoning:** [Provide a brief explanation here comparing the specific research topics, methodologies, or questions apparent in the titles/abstracts. Highlight similarities and differences relevant to the BRL field.]

Second, immediately following your reasoning, provide the numerical rating on a new line using the specified format. Rate the similarity of the specific research topics on a scale from 0 to 100:

- **0:** Completely different specific research topics within BRL.
- **50:** The articles share significant common ground but ultimately address distinct specific research topics within BRL.
- **100:** The articles address the same specific research topic within BRL.

Strictly format the rating line *exactly* like this, with no extra text before or after: Answer=[rating]

4 D - Taxonomic tags

We used the following prompts to generate taxonomic tags from Llama-3.3-70B-Instruct [1]. In place of $\{article\}$, we pasted the article title and abstract.

4.1 System prompt

You are a highly specialized academic research assistant with expertise in behavioral reinforcement learning. Your primary function is to meticulously analyze academic articles and generate accurate, concise, and contextually relevant taxonomic tags that capture their core subject matter and methodological approaches.

4.2 User prompt

Your objective is to analyze the provided academic article (Title and Abstract) and characterize it in terms of its core **topic** and **methodology**. Based on this analysis, you must generate between 3 and 7 **distinct and non-redundant** taxonomic tags.

Tagging Guidelines:

1. **Relevance:** Tags must accurately reflect the core topics and methods discussed.
2. **Conciseness:** Each tag should be a maximum of three words long.

3. **Source:** Tags should, as much as possible, use or be directly derived from the terminology present in the article's Title and Abstract.
4. **Specificity:** Prefer specific terms over overly broad ones, where appropriate, to best differentiate the article.
5. **Non-Redundancy:** Avoid tags that convey the same or very similar meaning. If multiple related terms exist, choose the most representative or synthesize them if possible.
6. **Coverage:** Aim for a balanced representation of both subject matter and methodology.

Article to evaluate: {*article*}
Output Structure:

1. **Reasoning (Brief):** First, provide a concise explanation of your tagging choices. Justify how your selected tags cover the article's subject and methodology, and briefly mention how you ensured non-redundancy and adherence to the article's language.
2. **Taxonomic Tags:** Immediately following your reasoning, provide the taxonomic tags. They must be separated by a semicolon (;) and enclosed within 'Answer=[...]'.

Strict Formatting Example for Tags: Answer=[tags]

References

1. Grattafiori, A. *et al. The Llama 3 herd of models* Nov. 2024. http://arxiv.org/abs/2407.21783.