REVIEW ARTICLE

The Landscape of Research on Prior Knowledge and Learning: a Bibliometric Analysis

André Bittermann[1] · Danielle McNamara[2] · Bianca A. Simonsmeier[3] · Michael Schneider[3]

Abstract

Understanding the role of prior knowledge in human learning is essential for predicting, improving, and explaining competence acquisition. However, the size and breadth of this field make it difficult for researchers to glean a comprehensive overview. Hence, we conducted a bibliometric analysis of 13,507 relevant studies published between 1980 and 2021. Abstracts, titles, and metadata were analyzed using text mining and network analysis. The studies investigated 23 topics forming five communities: Education, Learning Environments, Cognitive Processes, Nonacademic Settings, and Language. The investigated knowledge was diverse regarding its types, characteristics, and representations, covering more than 25 academic and non-academic content domains. The most frequently referenced theoretical backgrounds were the 3P Model, Cognitive Load Theory, and Conceptual Change approaches. While our results indicate that prior knowledge is a widely used cross-sectional research topic, there remains a need for more integrative theories of when and how prior knowledge causally affects learning.

Keywords Educational psychology · Learning sciences · Prestorage · Instruction · Research topics

Researchers have hypothesized that prior knowledge is among the strongest determinants of learning (e.g., Greve et al., 2019; Hambrick & Engle, 2002; Moehring et al., 2018). Prior knowledge can stimulate interest, guide attention, help interpret new information, aid memory encoding, enable logical inference, and guide problem-solving (Ormrod, 2019). It aids perception, reading comprehension, numerical

✉ André Bittermann
  abi@leibniz-psychology.org

1 ZPID – Leibniz Institute for Psychology, Universitaetsring 15, D-54296 Trier, Germany

2 Arizona State University, Tempe, AZ, USA

3 Trier University, Trier, Germany

thinking, and many other competencies that are foundations for success in school, work, everyday life, and participation in society (OECD, 2019). Accordingly, researchers have concluded that the concept of prior knowledge is central to neuro-cognitive, psychological, and educational research on learning. In line with previous studies, we define *prior knowledge* as information in a person's long-term memory at the onset of learning (e.g., Alexander et al., 1991; Greene et al., 2016; Kendeou & O'Brien, 2015; McCarthy & McNamara, 2021; Simonsmeier, et al., 2021). It can include facts, concepts, skills, beliefs, experiences, etc., and relations between these knowledge elements (e.g., diSessa et al., 2004; Medin et al., 2000). Prior knowledge can be explicit (i.e., conscious and verbalizable) or implicit (i.e., pre-conscious and nonverbalizable) (Broaders et al., 2007). In the learning sciences, the term *prior knowledge* includes knowledge that is correct as well as knowledge that is incorrect from a normative scientific view (i.e., misconceptions; e.g., Brod, 2021).

The fact that the concept of prior knowledge is relevant in many areas of neuro-cognitive, psychological, and educational research on learning has positive and negative implications. A positive implication is that the concept shows where different strands of learning research might connect and, perhaps, can be partly integrated. A negative implication is that it is almost impossible for a researcher to keep track of all new publications relevant to the study of prior knowledge. This makes it difficult to evaluate the innovativeness of research questions and projects, write comprehensive reviews of the field, summarize empirical evidence in meta-analyses, develop integrative theories of how prior knowledge affects learning, and derive general recommendations for how teachers should address prior knowledge in their lessons to aid learning.

We attempt to mitigate these problems by conducting a comprehensive bibliometric analysis of studies on prior knowledge and learning. In bibliometric analyses, the characteristics of a large number of studies are quantitatively summarized with statistical methods. Bibliometric analyses are aided by the fact that literature databases usually do not only include the name and reference of each study but also categorize each study in terms of metadata, for example, key concepts, study designs, investigated age groups, and other characteristics. Additional information can sometimes be inferred from the abstracts or article full texts, for example, which words commonly appear together in the same contexts. In the present study, we employed bibliometric analyses to summarize the landscape of research on prior knowledge. In the following sections, we briefly summarize the scope of research on prior knowledge, including recent developments and open questions, before deriving our research questions.

The Scope of Research on Prior Knowledge

In the bibliometric analyses presented in this study, we investigated the research topics and the theoretical background of the included studies as well as the content domains, types, characteristics, and representations of the investigated prior knowledge. We briefly introduce each of these aspects here.

The *research topics* in studies on prior knowledge are numerous. Example topics are the roles of prior knowledge in perception, human memory, skill acquisition, computer-assisted instruction, teacher education, medical training, and expert performance. Within each topic, several sub-topics are investigated. For example, within research on computer-assisted instruction, researchers investigate sub-topics such as the roles of prior knowledge in learning with animations, serious computerized games, online reading training programs, and collaboration in online small-group learning.

The *content domain* of the investigated knowledge is often a school subject or an academic discipline, such as language, mathematics, or medicine. In a broader sense, a content domain is any "clearly defined body of knowledge, skills, abilities, aptitudes, or tasks that may be measured with an appropriately constructed test" (Colman, 2015, p. 164). So, for example, gardening or football can also be considered content domains in a broader sense.

Knowledge can be described in terms of *knowledge types*. A common distinction is between declarative and procedural knowledge (e.g., Ullman, 2004). Declarative knowledge is verbalizable knowledge about facts and their interrelations. Procedural knowledge represents sequences of operators to reach goals, which are the cognitive foundations of skills. Sometimes declarative knowledge is further divided into semantic knowledge (also termed conceptual knowledge) about general concepts and their interrelations (Rittle-Johnson et al., 2015) and episodic knowledge about the learner's previous experiences and their situational contexts (Renoult et al., 2019). Another common distinction is between explicit knowledge, which is consciously accessible and verbalizable, and implicit knowledge, which is not consciously accessible and thus not verbalizable (Batterink et al., 2015). Many other types of knowledge have been proposed, for instance, tacit knowledge, mechanistic knowledge, and experiential knowledge (cf. de Jong & Ferguson-Hessler, 1996).

Knowledge can also be described in terms of its *characteristics* (see de Jong and Ferguson-Hessler, 1996, for an overview). Knowledge can be verbal or pictorial. It can be domain-general, for example, generic learning strategies, or domain-specific, that is, relating to central principles of a domain, such as the concept of force in physics or the concept of gene mutation in biology. Knowledge can be concrete, that is, tied to the learner's experiences, or abstract, for instance, the concept of energy conservation or the idea of justice (Wiemer-Hastings & Xu, 2005). Knowledge is fragmented when there are few interrelations between the knowledge elements in long-term memory. It is integrated when there are many connections between knowledge elements in long-term memory so that the learner understands how the elements interrelate. Procedural knowledge can be automatized to different degrees depending on how much it has been practiced. Compared to non-automated procedures, automatized procedures can be carried out faster and with fewer cognitive resources that tax working memory capacity (Tenison et al., 2016).

Studies on prior knowledge differ in their *theoretical backgrounds*. Numerous theories and models include assumptions about the effects of prior-knowledge on learning (Alexander & Dochy, 1995). Some of these theories and models focus on cognitive processes, others on school instruction or developmental patterns. For example, in the 3P (presage, process, product) Model of Learning, prior knowledge

is conceptualized as part of the presage, which determines how students engage in learning processes, for example, deep-level or surface-level learning. These processes then give rise to learning outcomes such as facts, skills, or involvement (Biggs, 1993). In the Cognitive Theory of Multimedia Learning, prior knowledge is recalled from long-term memory and helps to interpret and integrate pictorial and verbal information in working memory (Mayer, 2014). Within theories of text and discourse comprehension, such as the Construction-Integration model (Kintsch, 1988) and the Landscape Model (van den Broek et al., 1999), prior knowledge is necessary to understand the explicit words and sentences as well as to integrate ideas and generate inferences to build a coherent mental model of the deeper meaning of the discourse (McNamara & Magliano, 2009).

Examples of Recent Developments and Open Questions

Research on prior knowledge is an active field of research. For instance, research on prior knowledge has branched out from investigating knowledge in the domains of language, mathematics, and science to a broader range of domains, including health, computer science, and the social sciences (Simonsmeier, et al., 2021; Vosniadou, 2013). There is also increased interest in the relationship between prior knowledge and intelligence (e.g., Schneider & McGrew, 2012). Whereas it has long been known that intelligence comprises a fluid and a crystallized component, researchers have recently begun to use confirmatory factor analyses to investigate whether crystallized intelligence can be modeled as a second-order factor with knowledge in different domains as indicators (Watrin et al., 2022).

There is increasing awareness that not only the quantity but also the quality of prior knowledge determines learning. For instance, in research on science and mathematics learning, it has been suggested that a fragmented structure of knowledge can explain many learners' struggles in concept acquisition and conceptual change (diSessa, 2017). As a result of knowledge fragmentation, learners can simultaneously hold (incorrect) naive and (correct) scientific concepts in their long-term memory. These concepts interfere during recall, and the strength of this interference correlates with school achievement (Shtulman & Valcarcel, 2012; Stricker et al., 2021). The integration of fragmented knowledge into coherent knowledge structures has thus been described as a central goal of school instruction (Linn, 2006).

Relatedly, the Multidimensional Knowledge in Text Comprehension (MDK-C) framework describes in detail how the amount, accuracy, specificity, and coherence of prior knowledge determine learners' understanding of texts (McCarthy & McNamara, 2021). Accordingly, amount refers to how many relevant concepts a reader knows, whereas accuracy refers to the extent to which this knowledge is correct or factual according to normative views. Specificity refers to the degree to which the knowledge is related to information in a target text or discourse. Coherence refers to the interconnectedness between concepts and ideas in prior knowledge. The MDK-C framework assumes that prior content knowledge can be represented as a multidimensional mental model or semantic network (e.g., comprised of nodes and links). Based on spreading activation, information in prior knowledge that is more semantically-related and more

interconnected with the new incoming information from a text is more likely to be activated and to be activated more quickly. The underlying motivation for the MDK-C framework is that providing clearer definitions and boundaries around these four (and other) dimensions of prior knowledge will afford more nuanced predictions about how knowledge is used during text and discourse comprehension.

In sum, both conceptual change approaches and the MDK-C framework agree that knowledge has a network structure that can be coherent or integrated but also incoherent and fragmented in learners. The partly fragmented or incoherent nature of prior knowledge has inspired new methodological research. It has been shown that a low internal consistency (i.e., a low Cronbach's alpha) of knowledge tests can naturally arise from the heterogeneous nature of the construct and does not necessarily indicate high amounts of random measurement error (Stadler et al., 2021; Taber, 2017). Researchers are increasingly turning to multivariate methods to statistically model heterogeneous knowledge structures and their changes over time, for instance, latent factor models for continuous changes (Schneider et al., 2011) or latent transition models for discontinuous changes (Hickendorff et al., 2018). Another notable methodological development is the emerging use of brain imaging studies for the investigation of knowledge, for instance, fMRI (Gao et al., 2019) and fNIRS (Liu et al., 2019). Brain imaging studies have made progress in uncovering neural correlates of prior knowledge effects on learning (see Brod et al., 2013, for a review).

A recent meta-analysis of how strongly the amount of domain-specific prior knowledge affects learning (Simonsmeier, Flaig, Deiglmayr, et al., 2021) distinguished between two relevant indices that are partly independent of each other: The correlation between prior knowledge and posttest knowledge reflects the stability of individual differences in knowledge over time. This stability was high ($r_{\mathrm{P}}^{+} = .53$) averaging over 7772 effect sizes. The correlation between prior knowledge and knowledge gains from pretest to posttest reflects the effect of prior knowledge on learning something new. This effect size was virtually zero ($r_{\mathrm{NG}}^{+} = -.06$), averaging over 697 effect sizes. The effect sizes followed a normal distribution around this mean and varied strongly from $r_{\mathrm{NG}}^{+} = -.94$ to $r_{\mathrm{NG}}^{+} = .93$. This demonstrates that prior knowledge can have strong positive or strong negative effects on learning but that there are also many instances of learning where prior knowledge has only a weak or negligible effect. Notably, this result might be expected or even hoped for if an intervention were intended to overcome the impact of prior knowledge, but only a few publications included in the meta-analysis indicated that this was the aim of the study. Thus, the meta-analytic findings indicate that the commonly referred to knowledge-is-power hypothesis in its generality is incorrect and that future research needs to identify the conditions under which prior knowledge has strong positive, strong negative, or negligible effects on learning and how the impact of prior knowledge can be mitigated via instruction or training.

The Current Study

Overall, the effects of prior knowledge on learning have been investigated for decades and in a multitude of literature strands that continue to evolve. The number and breadth of studies on prior knowledge make it difficult for researchers to keep track

of new developments and evaluate the merit of ideas for new studies and projects. The continuous proliferation of studies renders it challenging to conduct comprehensive literature reviews and meta-analyses. It also hampers the derivation of general conclusions about how prior knowledge affects learning that can be communicated to practitioners and policy-makers.

To provide a more comprehensive orientation and overview of research on prior knowledge, we conducted a bibliometric analysis of the research literature on prior knowledge. In bibliometric analyses, a corpus of a large number of publications is analyzed quantitatively in an automatized way. This method allowed us to address 12 research questions (RQs).

The first two RQs are concerned with the detection and analysis of research topics to gain an interdisciplinary overview of prior knowledge research. The topic identification also served as an automated eligibility screening tool: Publications included by the search query but addressing topics unrelated to prior knowledge were removed from the corpus.

RQ 1: What are the main topics of prior knowledge research?

RQ 2: How do research topics relate to each other in terms of semantic similarity?

RQs 1 and 2 targeted publication topics in an exploratory manner. We accompanied this bottom-up approach with frequency analyses of known and theoretically relevant terms to further investigate specific contents and theoretical constructs in prior knowledge research:

RQ 3: What is the content domain of the investigated knowledge?

RQ 4: What is the frequency of the investigated types, characteristics, and representations of knowledge?

RQ 5: What theoretical backgrounds are referred to in the studies?

The next three research questions aim at mapping prior knowledge research from a methodological point of view:

RQ 6: Which age groups and educational stages are investigated?

RQ 7: Which measures are employed?

RQ 8: What research designs are used?

The following two research questions address publication outlets for studies on prior knowledge. This is important background knowledge for researchers, lecturers, and students aiming to inform themselves about where to publish and look for the latest research and discussions.

RQ 9: Which journals are the most common outlets of prior knowledge research?

RQ 10: How do journals relate to each other in terms of topical similarity?

Finally, we investigated the scientific disciplines involved in prior knowledge research and looked at psychological subfields:

RQ 11: Which scientific disciplines address human learners' prior knowledge, and how do they differ regarding publication output?

RQ 12: Which subfields of psychology address prior knowledge, and how do they differ regarding publication output?

This study was preregistered (https://doi.org/10.23668/psycharchives.4903). In a few minor cases, we deviated from the preregistration and whenever we do so we

report the reasons. In particular, the research questions are an updated version of the research questions preregistered for this study. We have reordered the preregistered research questions. RQ 1 now refers to topics instead of subtopics for a more straightforward presentation. We have rephrased RQ 5 to comprise theories as well as models. RQs 11 and 12 (formerly: RQ 13) are now stated more precisely with respect to either psychological subfields or scientific disciplines. Results on the topical similarity of disciplines and psychological subfields (formerly: RQ 12; cf. Figure S1) are provided in the Electronic Supplementary Material (ESM).

Method

This section outlines how we queried the databases and analyzed the data. The analysis code can be found in the ESM. Due to copyright reasons, we cannot publish the full corpus, but we provide basic bibliographic information and precomputed R objects such as document-feature matrix, topic model, and networks.

Data

Our research questions relate to how humans use prior knowledge in learning, not how animals use prior knowledge or how computer programs use it in artificial intelligence. Therefore, we searched databases for psychological and educational research, but no databases for biology or computer science. We gathered publication records from *PsycInfo*, *PSYNDEX*, and *ERIC*. Initially, we planned to include MEDLINE for the medical sciences, as prior knowledge is a common construct in patient education. However, after a screening of MEDLINE data, we decided not to explicitly include the medical sciences in our analyses. Patient education was also addressed in publications indexed in psychological databases and those MEDLINE publications not being duplicates did not have an actual focus on prior knowledge in learning contexts. The search string included synonyms of prior knowledge listed by Dochy and Alexander (1995, p. 227) together with terms indicating learning:

(("prior knowledge" OR prestorage OR "permanent stored knowledge" OR "prestored knowledge" OR "knowledge store" OR "implicit knowledge" OR "archival memory" OR "experiential knowledge" OR "background knowledge" OR "world knowledge" OR "pre-existing knowledge" OR "preexisting knowledge" OR "personal knowledge")

AND

(learn* OR instruct* OR educat* OR development* OR train* OR teach* OR school* OR lesson* OR achiev* OR memor* OR acquisition OR practicing OR practice* OR artificial intelligence OR cognitive model* OR categor*)). ab,ea,fd,fw,hw,id,sh,su,ot,ti.

In addition, we searched the *Web of Science* using respective discipline filters with the following query:

(TS=(("prior knowledge" OR prestorage OR "permanent stored knowledge" OR "prestored knowledge" OR "knowledge store" OR "implicit knowledge" OR "archival memory" OR "experiential knowledge" OR "background knowledge" OR

"world knowledge" OR "pre-existing knowledge" OR "preexisting knowledge" OR "personal knowledge")

AND

(learn* OR instruct* OR educat* OR development* OR train* OR teach* OR school* OR lesson* OR achiev* OR memor* OR acquisition OR practicing OR practice* OR artificial intelligence OR cognitive model* OR categor*)))

AND

(WC=(Psych*) OR WC=(Education*) OR WC=("Behavioral Sciences"))

Publications not in the English language or without English translation were excluded, as well as research published before 1980 or after 2021. We accessed all databases on March 4th, 2022. For PsycInfo, PSYNDEX, and ERIC, this search strategy yielded 14,276 results, and for Web of Science 4467. We exported all the records (including titles, abstracts, authors, keywords, and metadata such as, for instance, year of publication, age group (childhood, adolescence, adulthood), or methodology (quantitative study, literature review, etc.) for subsequent bibliometric analysis. Given that we queried four databases, a systematic check for duplicates is paramount. Here, we performed a stepwise identification of duplicates by making use of digital object identifiers, titles, abstracts, and publication year. In case of duplicates, we merged complementing information from different databases (see ESM for more details).

Analytical Procedure

Topic Identification

To identify topics of prior knowledge research (RQ 1), we first joined titles, abstracts, and keywords to a text corpus. We then preprocessed the text data by tokenizing, transforming to lower case, and removing both standard stopwords (e.g., "the", "and") and stopwords for scientific abstracts ("study", "investigated", "results")[1]. The stopword lists are provided in the ESM. Moreover, punctuation, numbers, symbols, and separators were removed, and finally, tokens were lemmatized[2].

Next, we employed structural topic modeling (STM; Roberts et al., 2014) for topic identification. In topic modeling, clusters[3] of frequently co-occurring terms are detected using an unsupervised machine learning algorithm. These clusters are referred to as *topics*. Each publication can be assigned to multiple topics with

---

[1] In contrast to our preregistration, we did not exclude terms that define the corpus (e.g., "prior" or "knowledge"), as keeping them in the topic model did not result in an omnipresence across topics as we initially expected.

[2] Lemmatization was not performed for controlled keywords (i.e., APA thesaurus, MeSH terms), as they are standardized and hence, no morphological variations existed for these terms in the corpus.

[3] To be precise: Each topic consists of the full corpus vocabulary but with different term probabilities. For topic interpretation, "clusters" defined by the *n* most probable terms are used (along with inspecting the most representative documents). Common practice is to report the top 5 to 10 terms per topic.

different probabilities because it can include several terms related to differing topics. For instance, a publication about memory processes during online learning will have high document[4]-topic probabilities for the topics about memory and computer-assisted instruction and low probabilities for the other topics in the model. The specific topic modeling variant STM considers topic correlations in the topic detection process and thus yields well-separated topics—which is of a particular advantage given the topical specificity of our corpus (i.e., prior knowledge). In addition, topic reliability (i.e., reproducibility of topics across different runs of the algorithm) is warranted, as the model initialization is deterministic in STM by using a spectral decomposition of the word co-occurrence matrix instead of random initialization.

To find the optimal number of topics, we examined several candidate models regarding statistical metrics (i.e., semantic coherence, Mimno et al., 2011; exclusivity, Bischof & Airoldi, 2012), qualitative investigation of intersubjective topic interpretability, granularity level (semantic broadness vs. specificity), clustering solutions, and topic validity (using database content classification metadata for external validation). Based on our experience with comparable corpora, we set the initial range to 5–50 topics.

Final Model and Final Sample

Based on the topic quality and our interpretation of topic terms and most representative publications per topic, we concluded that a model with 25 topics had the best overall fit on the data. From this model, we excluded two topics as they revealed artifacts of our search query. For instance, some publications did not investigate prior knowledge. Still, their abstracts contained phrases such as "This review extends prior knowledge of…" or "No background knowledge is needed for studying this book." Leveraging this detection of false positives in our corpus, we removed all documents from the corpus that predominantly addressed these artifact topics (i.e., document-topic probabilities $> .5$). By doing so, we removed $n = 160$ irrelevant publications. Totals of $N = 13,507$ publications and $k = 23$ topics remained and were analyzed further.

Topic Networks

To analyze topic similarity (RQ 2), we used pairwise topic correlations based on the document-topic-probability distributions to build a network with topics as nodes and their bivariate correlations as edges. Document-topic-probabilities are based on the share of terms assigned to the topic. For instance, if the algorithm assigns 40 of 100 terms in a document to Topic X, the document-topic-probability is 40%. To group similar topics, community detection with the multi-level modularity optimization algorithm (Blondel et al., 2008) was applied, following the suggestions by Yang et al. (2016). With this algorithm, communities are detected by dividing the

---

[4] In the following, we use "document" as in topic modeling terminology, i.e., text units included in the corpus (in our case: abstracts, titles, and keywords of the publication references).

whole network into subnetworks with many edges within these subnetworks and few between them. This "modularity" of the network is found by iteratively optimizing the results from two phases: First, local communities are found by sequentially building communities from node neighbors until there is no more gain in modularity. Second, these communities are aggregated into a network of meta-communities. These two phases are repeated until a maximum of modularity is reached. All networks in our study were visualized using the Fruchterman-Reingold layout (Fruchterman & Reingold, 1991). In this force-directed graph drawing algorithm, nodes are assumed to repel each other, while edges attract the connected nodes, similar to springs. By doing so, the network can be plotted in a two-dimensional space with similar nodes being shown near each other.

Frequency Analyses

For counting the content domains (RQ 3), types, characteristics, and representations of knowledge (RQ 4), theoretical constructs (RQ 5), and measured (RQ 7), we used precompiled lists and/or keywords-in-context analysis (KWIC) for pattern matching. For instance, we inspected the word context for the keyword "theory" to retrieve "Sweller's cognitive load theory". Age groups and educational stages (RQ 6), study research designs (RQ 8), journals (RQs 9 and 10), as well as scientific (sub-)disciplines (RQ 11 and 12) were analyzed using database fields. See more details in the ESM.

Results

The results are based on 13,507 analyzed studies. Figure 1 shows the number of studies by year and database. The number of publications per year on prior knowledge and learning increased strongly from 1980 to 2021. The increase was more pronounced in the psychological literature (i.e., in PsycInfo) than in the educational literature (i.e., ERIC). Currently, about 600 new publications on prior knowledge are published every year (upper line in Figure 1). The temporal trends follow the overall publication increase in the databases: The correlations between prior knowledge publications per year and all database records per year range between $r = .87$ (ERIC) and $r = .99$ (PsycInfo). The respective plots can be found in Figure S2.

Research Question 1: Topics in Prior Knowledge Research

The 23 topics of the final model are listed in Table 1. We labeled the most prevalent topic *Reading Instruction* because the terms characterizing this topic were *literacy, reading instruction, organizer, strategy,* and *reading strategies* and because the publications that are representative of this topic could be subsumed under this term. As indicated by our algorithm, a representative example publication for this topic had the title *Building Student Literacy Through Sustained Silent Reading*. We assigned the labels of the other topics in a similar way. The prevalence of this topic was 5.35%. The prevalence indicates the mean share of
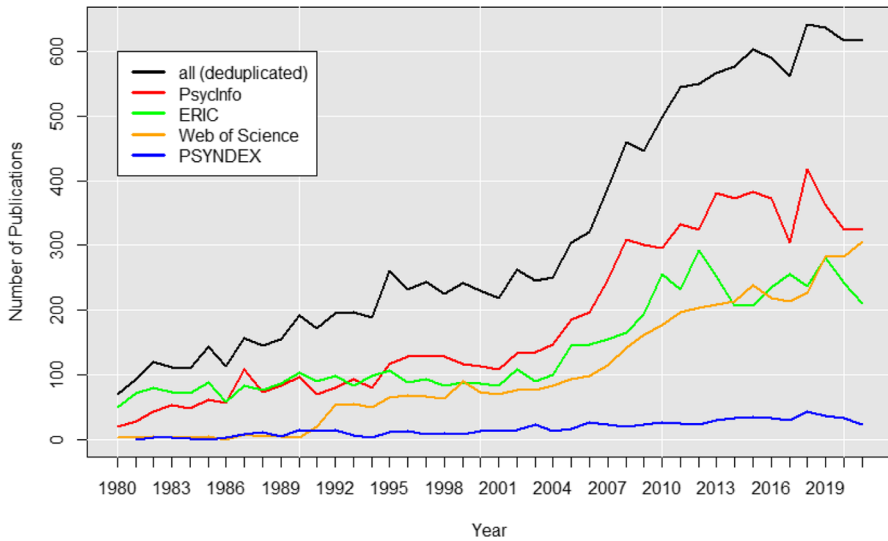
Figure 1 Number of publications on prior knowledge by year and database

terms in our corpus relating to this topic and should not be misinterpreted as the number of publications on this topic in our database.

The range of topics is very broad. The topics relate to various content domains (e.g., language learning, mathematics, science, patient education), educational levels (e.g., preschool, school, higher education), learner populations (e.g., school students, teachers, workers, patients, consumers), psychological mechanisms (e.g., perceptions, categorization, memory processes, multimedia learning, comprehension, self-regulation, social judgment, communication), and levels of analysis (e.g., brain, individual learners, groups, organizations, culture, and society).

Overall, the topic model is a relatively comprehensive representation of the lines of prior knowledge research that we know. All dominant themes in prior knowledge research that we are aware of are represented in the topic list. All topics listed in Table 1 represent well-known and well-investigated strands of the literature on prior knowledge. Lines of research investigating knowledge acquisition and learning missing from the topic list are research on the relation between intelligence and knowledge (Watrin et al., 2022), studies on the acquisition of skills and expert performance (Ullén et al., 2015), research on infants' core knowledge (Spelke & Kinzler, 2007), and studies on psychoeducation and psychotherapy (Simonsmeier, et al., 2021). Studies on these topics were in our database, but they were so rare that our algorithm did not detect them as separate topics. Major *general* lines of research in the learning sciences not on the topic list are personality, motivation, emotion, and interests. Such studies are among the publications analyzed in our study. Some of them are highly influential and have high citation rates (e.g., Pintrich et al., 1993; Tobias, 1994). However, the number of terms relating to these lines of research in our text corpus was so small that the topic model did not identify them as topics of research on prior knowledge.

Table 1 Topics in prior knowledge research

Topic community and topic	Description	Characteristic terms	Mean prevalence of characteristic terms (%)	Representative publication title
Language				
Reading Instruction	Literacy, reading instruction, reading diagnostic assessment	Literacy, reading instruction, organizer, strategy, reading strategies	5.35	"Building Student Literacy through Sustained Silent Reading"
Reading Comprehension	The role of prior knowledge in text comprehension	Text, expository, passage, comprehension, reading comprehension	5.31	"Third grade readers' use of prior knowledge in inferential comprehension"
Foreign Language	Second- and third-language learning, cross-linguistic comparisons	English, second language learning, language, EFL, English (second language)	3.19	"English Prototyped Small Clauses in the Interlanguage of Chinese/Taiwanese Adult Learners"
Learning Environments				
Academic Achievement	Measurement, predictors, and causes of academic achievement	Achievement, score, group, post-test, pretest	5.12	"Effects of Prior Knowledge of Topics and the Instructional Objectives on Students' Achievement in Literature-in-English"
Self-Regulated Learning	Self-regulated learning, learning strategies, and their roles in (computer) learning environments	Learn, feedback, tutor, scaffold, self-regulated	5.01	"Self-regulation in computer-based learning environments: Effects of learner characteristics and instructional support"
Multimedia Learning	Multimedia learning and Cognitive Load Theory	Load, multimedia, transfer, animation, condition	3.71	"Learning from instructional animations: How does prior knowledge mediate the effect of visual cues?"
Information Literacy	Search, organization, and evaluation of Information, use of databases	Internet, digital, web, online, navigation	2.78	"Combining Design Patterns and Elements of Social Computing for the Design of User Centered Online Help Systems"
Education				

Table 1 (continued)

Topic community and topic	Description	Characteristic terms	Mean prevalence of characteristic terms (%)	Representative publication title
Teacher Knowledge	Teacher knowledge, how it is affected by treacher training, and how it affects teaching practices	Preservice, teacher, classroom, teach, pedagogy	5.08	"Assessing Teachers' Beliefs to Facilitate the Transition to a New Chemistry Curriculum: What Do the Teachers Want?"
Student Knowledge in Higher Education	Students' knowledge in higher education classes	Chemistry, music, instructor, lecture, laboratory	3.90	"Students' Understanding of External Representations of the Potassium Ion Channel Protein"
Organizational Knowledge in Higher Education	Faculty development and knowledge management in higher education	Faculty, game, program, engineer, module	3.67	"Development of university-industry partnerships in railroad engineering education"
Social & Cultural Issues	Social, culture-, and race-related issues in schools	African, culturally, district, American, race	3.72	"Latina/o students' experiences in a small high school and college access through a critical race theory perspective and community cultural wealth model."
Science Education and Conceptual Change	Science education, conceptual change, and students' misconceptions	Science, scientific, analogy, science education, physic	3.42	"Diagnosing and Dealing with Student Misconceptions: Floating and Sinking"
Mathematics Learning	Mathematical learning and problem solving, understanding of numbers, mathematical expressions, and word problems	Mathematical, mathematics, algebra, problem solving, solve	2.88	"Confounding whole-number and fraction concepts when building on informal knowledge"
Nonacademic Settings				
Communication	Communication in real-life contexts, communicative actions, pragmatics of language	Action, discourse, situation, conversation, body	4.73	"Navigating difficult conversations: The role of self-monitoring and reflection-in-action"

Table 1 (continued)

Topic community and topic	Description	Characteristic terms	Mean prevalence of characteristic terms (%)	Representative publication title
Businesses and Organizations	Knowledge management in businesses and organizations	Coach, experiential, team, organizational, innovation	4.59	"Managing organisational knowledge through recruitment: Searching and selecting embodied competencies"
Patient and Medical Education	Education of patients about their health issues and training of health care professionals	Nurse, health, medical, illness, care	4.08	"Knowledge, attitudes, risk perception, and cancer screening behaviors among cancer survivors"
Preschool	Learning in preschool, children's play	Children's, child, preschool, childhood, story	2.60	"Teaching and preschoolers' ability to infer knowledge from mistakes"
Cognitive processes				
Machine Learning	Machine learning algorithms drawing on prior knowledge	Algorithm, propose, computational, machine, bayesian	4.59	"Interpretable interval type-2 fuzzy predicates for data clustering: A new automatic generation method based on self-organizing maps"
Task Performance	Factors affecting performance in, e.g., complex problem-solving tasks and writing tasks	Variable, validity, predictor, measure, intelligence	4.35	"Anxiety and cognitive functioning: Effects on performance in academic tasks"
(Neuro-)Cognition of Memory	Memory processes and their neural bases	Memory, episodic, encode, hippocampus, cortex	4.29	"Differential roles for medial prefrontal and medial temporal cortices in schema-dependent encoding: From congruent to incongruent"
Implicit and Explicit Knowledge	Acquisition and development of implicit and explicit knowledge and their roles in sequence learning, consciousness, and automatization	Implicit, explicit, sequence, acquire, knowledge level	4.18	"Evaluating the relationship between explicit and implicit knowledge in a sequential reaction time task"

Table 1 (continued)

Topic community and topic	Description	Characteristic terms	Mean prevalence of characteristic terms (%)	Representative publication title
Perception and Categorization	Perception and categorization in humans and primates	Object, categorization, perceptual, category, stimulus	4.08	"Familiar Size Effects on Reaction Time: When Congruent is Better"
Social Cognition	Social judgment and decision making, consumer choices	Causal, consumer, decision, food, epistemic	3.91	"If You Judge, Investigate! Responsibility Reduces Confirmatory Information Processing in Legal Experts"

The topic and community labels were assigned by the authors based on top terms and the most representative documents of each topic. The five most characteristic terms of each topic are lifted by both frequency and exclusivity. Topic prevalence is the mean share of terms in the corpus relating to the topic.

Most topics detected by the algorithm were homogeneous regarding their representative terms and publications. The topic that we labeled Social Cognition was heterogeneous in that it included terms relating to social judgment, decision-making, and consumer choices. The topic *Patient Education and Medical Education* was heterogeneous only in that it included studies on medical education (i.e., the training of medical professionals) and studies on patient education (i.e., the teaching of patients). These two literatures were merged by the algorithm because they use similar terms notwithstanding the fact that the training of medical professionals and the education of patients differ in their content and the required instructional techniques.

All 23 topics in Table 1 have been frequently investigated, but they differ in the mean prevalences of terms relating to this topic in our text corpus. The five most prevalent topics (i.e., *Reading Instruction, Reading Comprehension, Academic Achievement, Teacher Knowledge, Self-Regulated Learning*) relate to the role of prior knowledge in school and related learning settings. Among the relatively less prevalent topics were studies on *Mathematics Learning, Information Literacy,* and *Preschool* education. As shown in Table 1 and Figure 2, a relatively small number of studies investigated mathematical knowledge. The reason for the low prevalence of the Topic *Mathematics Learning* is, thus, not a lack of studies on this topic, but the fact that mathematical content is oftentimes used to investigate design principles for effective learning environments, cognitive memory processes, neurocognitive processing of numerical magnitudes in the brain. Thus, these studies might be represented by topics other than *Mathematics Learning* in our topic model. Studies on clinical psychology and psychotherapy research were rare (3.2 %; see also Figure S3) and thus did not form their own topic in our analyses. Most included studies from these subfields investigated psychoeducational questions and are represented by the topic *Patient and Medical Education.*

Research Question 2: Relations Between the Topics

Figure 2 visualizes how the 23 topics relate to each other in terms of similarity. The node diameter indicates the topic prevalence. The edge thickness indicates the similarity of the connected topics. A high similarity means that the two topics are similar in how frequently they refer to each of the 4,722 terms in our corpus. Each topic is somewhat similar to each other's topic, for example, because they all refer to prior knowledge. Only similarities greater than .05 are visualized by edges in Figure 2 to aid the interpretation. A high topic similarity could, for example, be due to the fact that two topics are integrated on the theoretical level, use related methodologies, or cite the same studies.

The community detection algorithm (Fruchterman & Reingold, 1991) indicated that there are five communities of topics, each one shown in a differently colored shape in Figure 2. The four (blue) rhombi in the middle form a community which we labeled Learning Environments. The topics in this community are concerned with the multimedia design of learning environments, how learning environments can measure and foster self-regulated learning strategies, how achievement can be predicted and measured in learning environments, and how learners find and
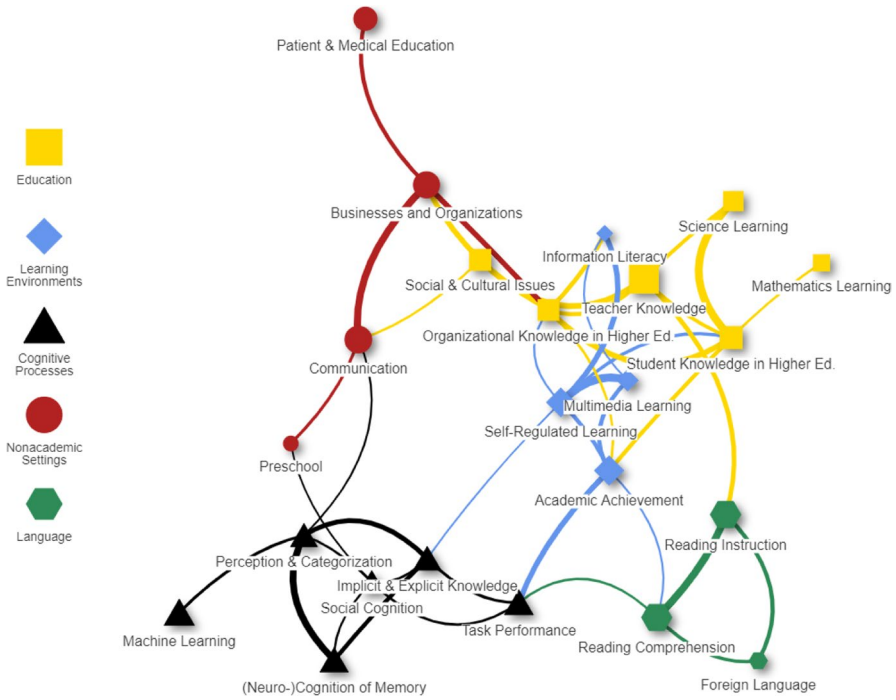
Figure 2 Topic similarity network with nodes representing the 23 topics. *Note*. The node diameter is proportional to the topic prevalence. The edge thickness indicates the topic similarity. Only edges with correlations > .05 are shown. The shapes indicate groups of similar topics as found by a community detection algorithm. For details on bivariate correlations, please find an interactive and zoomable version of this figure and a table of topic correlations in the ESM

evaluate information in learning environments. This topic community was closely related to a second community, which we labeled Education. The topics in this community were concerned with school teachers' knowledge, learning and conceptual change in mathematics and science, and social and cultural issues in school. It also included two topics on knowledge in higher education. A third community, indicated by (green) hexagons, related to the first two communities and was labeled Language. It included studies on reading comprehension, reading instruction, and foreign language learning. We labeled the fourth community Cognitive Processes. It is indicated by (black) triangles and includes research on the (neuro-)cognition of memory, explicit and implicit knowledge, prior knowledge in perception and categorization, social cognition, and related topics. The fifth community, indicated by (red) circles, thematized prior knowledge in more practical contexts out of K12-schools and higher education, for example, in preschools, the health care system, and the workplace. We labeled this community Nonacademic Settings.

The network edges between the topic communities indicate that research on education is informed by research on the design of learning environments and language which are drawing on research on the underlying cognitive processes.

Research on prior knowledge in nonacademic settings related to some aspects of education and cognitive processes, but these links were rare and mostly weak. The topic community *Nonacademic Settings* was not directly linked to the communities *Learning Environments and Language*. The community *Education* was not directly linked to the community *Cognitive Processes.*

Research Question 3: Content Domains of Knowledge

At least one content domain could be assigned to 40.83% of the publications included in our analyses. More specifically, 32.92% investigated knowledge in only one content domain, 6.82% in two domains, 0.98% in three domains, and 0.11% in even more domains. Figure 3 shows the content domains of the investigated prior knowledge. As expected, most studies investigated prior knowledge in the domains of reading, mathematics, and science. These are school subjects that provide foundations for learning in all other domains. For these reasons, large international student achievement studies also focus on these subjects, for example, the tri-annual PISA studies (OECD, 2019). Other domains on the list are the subjects of popular study programs in higher education (e.g., psychology, engineering, environmental sciences), other school subjects (e.g., computer science, art, music, physical education), relate to patient and medical education (i.e., medical, health, patient, nursing) or to social issues (i.e., multicultural, social studies, citizenship). The frequency of studies addressing STEM (science, technology, engineering, mathematics) knowledge was surprisingly low, but this resulted from the fact that most studies more specifically stated whether they investigated science knowledge, technology knowledge, and so on. If these numbers were added, STEM would be among the two most frequently investigated content domains in our database. Similarly, three domains in Figure 3 relate to forms of language learning and reading. Their combined frequency would be above 15%. Knowledge about cultural issues was among the five most frequently investigated content domains, which is interesting because this is not a common discipline of formal instruction. As shown by Figure 2, the topic *Social and Cultural Issues* relates to four topics: *Businesses and Organizations*, *Communication*, *Teacher Knowledge*, and *Organizational Knowledge in Higher Education.*

Research Question 4: Types, Characteristics, and Representations of Knowledge

Table 2 lists the relative frequencies of empirical studies investigating specific types, characteristics, and representations of knowledge. A majority of 69% of the articles in our database did not mention any of the terms listed in Table 2 in their abstracts. Among the 8434 included empirical studies, only 21% explicitly stated that they investigated a type, characteristic, or representation of knowledge. No single type, characteristic, or representation of knowledge was mentioned in more than 4% of the studies. The relatively highest frequencies were found for the terms explicit and implicit knowledge, experiential knowledge, conceptual and procedural knowledge, and mental representations, for example, scripts, schemas, propositions, and mental models.
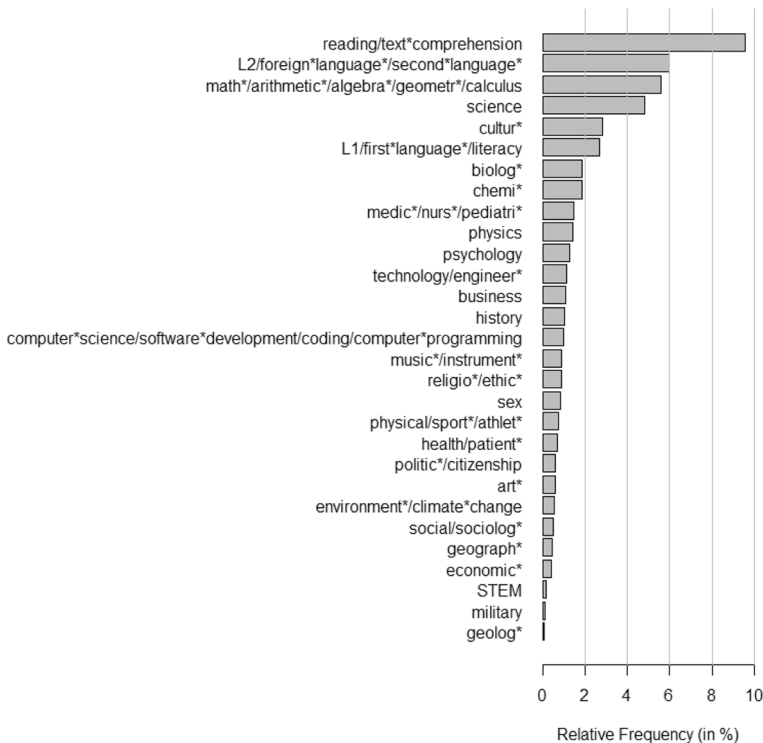
Figure 3 Relative frequency of the content domain of the investigated prior knowledge

Research Question 5: Theoretical Backgrounds of the Studies

As shown in Figure 4, the 3P (presage, process, product) model of learning (used in 2.69 % of the 13,507 publications in the corpus), Cognitive Load Theory (1.84 %) and Conceptual Change approaches (1.51 %) were the most frequently mentioned theoretical backgrounds in abstracts, titles, and metadata.

Most studies included in our analysis had a rather atheoretical approach in that they referred to neither of the models in Figure 4 in their abstracts, titles, and metadata. Even the most frequently mentioned background, the 3P model, was referred to in only less than 3% of the studies, that is, in 363 publications. This shows that there is neither one generally used theory of prior knowledge, nor a small group of competing theories that dominate the field of learning.

The numbers in Figure 4 might slightly underestimate how frequently the studies referenced a theory or model. It is possible that studies cite publications about a theory without explicitly mentioning the name of the theory. In this case, the reference to the theory would have been missed by our analysis. However, even when figuring in this potential bias, the results still indicate that many publications about prior knowledge do not reference specific theories or theoretical models.

Overall, the theories and models in Figure 4 are heterogeneous, ranging from theories making very specific statements of how prior knowledge modulates learning, for example, the Cognitive Load Theory, to very broad research areas

Table 2 Percent of publications mentioning a type, characteristic, or representation of knowledge at least once

Types		Characteristics		Representations	
Implicit	2.30	implicit	2.30	script*	3.55
Experiential	1.92	explicit	0.99	representation*	3.41
Explicit	0.99	domain-specific	0.18	schema*	1.42
Conceptual	0.53	tacit	0.17	brain*	1.06
Procedural	0.30	meta-cognitive	0.11	proposition*	0.55
Semantic	0.21	structur*	0.09	mental model*	0.52
Declarative	0.18	abstract*	0.04	imagery	0.30
Strategic	0.05	integrat*	0.04	neural network*	0.10
Category	0.03	fragment*	0.02	mental image*	0.05
Conditional	0.02	situation*	0.01	semantic network*	0.04
Episodic	0.01	verbal	0.01	production rule*	0.01
		deep	0.01		
		domain-general	0.01		

Elements with a share of 0 % are omitted. The asterisks (*) were used to take account of grammatical variations. The terms were searched in method descriptions of study abstracts. Please note that these results refer to text data (titles and abstracts) and metadata provided by the literature databases: An automated search in full texts might yield larger shares

that are primarily concerned with other issues that knowledge acquisition and only investigate prior knowledge as one of many relevant variables, for instance, *critical race theory*.

Research Question 6: Age Groups and Educational Levels

As shown in Table 3 (upper), most research was conducted with adult subjects (40.71 % of the 2604 included studies with sample age information). Most participants were young adults, primarily because a large amount of research used convenience samples of students in higher education. Children up to the age of 12 years comprised the second-largest share (36.06 %), followed by secondary school children (23.23 %).

A closer look at the participants' educational levels (Table 3, lower) revealed that postsecondary education (e.g., higher education, continuing education) was most frequently addressed (46.31 % of the 7812 studies with available information in database records), followed by secondary education (29.78 %), elementary education (16.59 %), and early childhood education (7.31 %). The hierarchy of educational levels is based on the ERIC database classification.

Research Questions 7 and 8: Measures and Research Designs

In the 416 database records with available information, the most frequently mentioned measures (Figure 5) were of a rather generic nature: fMRI (12.74 %), multiple-choice test (7.69 %), transfer test (7.45 %), and retention test (6.25 %). The
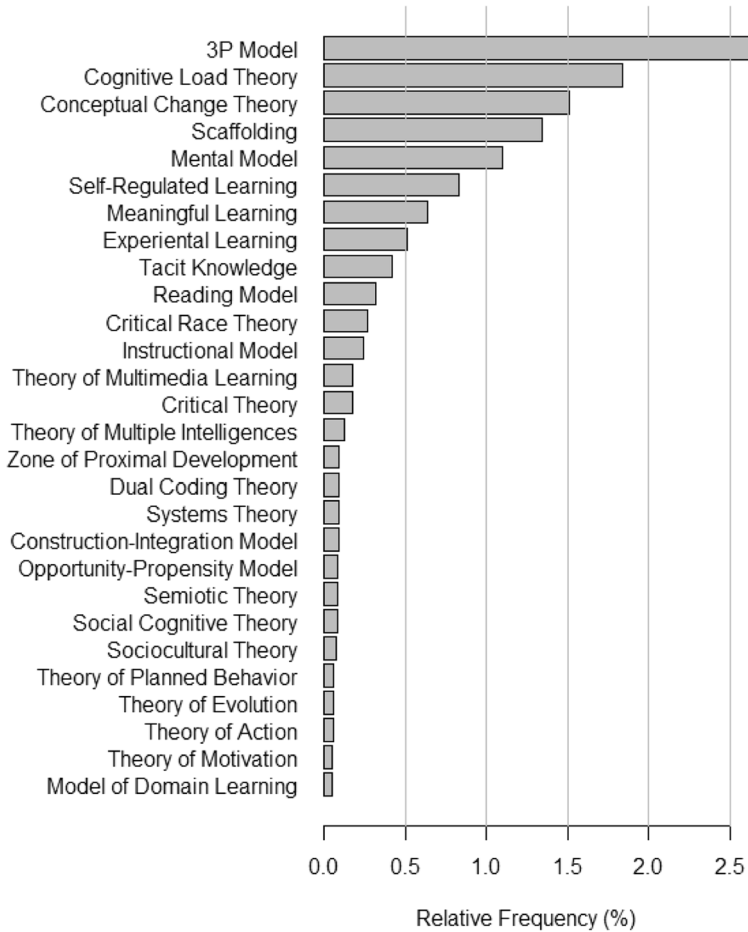
Figure 4 Theoretical backgrounds most frequently mentioned in study abstracts, titles, or keywords. *Note*. Only terms with a minimum relative frequency of 0.05 % in the corpus are shown

Table 3 Relative frequencies of investigated sample age and educational levels

Category	Relative frequency
Sample age	
Childhood (Birth – 12 years)	36.06 %
Adolescence (13–17 years)	23.23 %
Adulthood (18 years and older)	40.71 %
Educational level	
Early childhood education	7.31 %
Elementary education	16.59 %
Secondary education	29.78 %
Postsecondary education	46.31 %

Information on sample age was available in $n = 2604$ database records and on educational levels in $n = 7812$ records, respectively
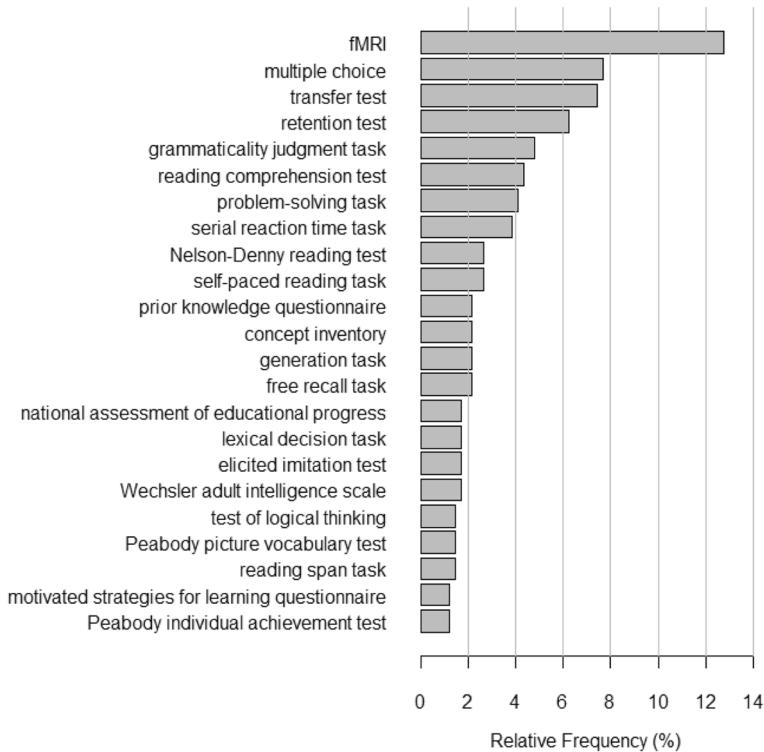
Figure 5 Relative frequencies of measures in studies with available information in the database record (*n* = 416). *Note*. Only measures with a minimum relative frequency of 1 % are shown

categories of measures in Figure 5 are not mutually exclusive. For example, a transfer test can include multiple-choice items. Not all listed assessments operationalize knowledge. Some are typical predictors or covariates of knowledge, for example, the Wechsler adult intelligence scale. It might seem surprising that fMRI is the most frequently used measure in prior knowledge research. However, its relative frequency was below 13%, and all other operationalizations listed in Figure 5 are various types of behavioral measures. Their combined frequencies show that more than 85% of the studies on prior knowledge and learning use behavioral measures.

Among the study designs (Figure 6), quantitative empirical research was the predominant category (42.50% of the 8207 studies with available information on research design). The overall share of the qualitative research category is 21.98 % (including, e.g., 12.89% interviews, 7.69% unspecified qualitative research, and 1.51% focus groups). There was only a small number of research syntheses: meta-analyses comprised 0.43%, systematic reviews 0.35%, and meta-syntheses 0.04% of the studies with available information research methodology.
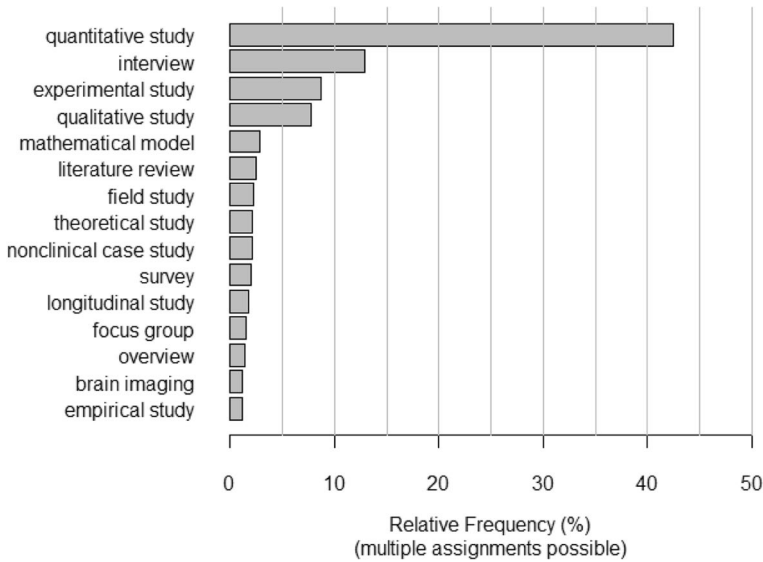
Figure 6 Relative frequencies of research designs in studies with available information in the database record (*n* = 8207). *Note*. Only designs with a minimum frequency of 1 % are shown

Research Questions 9 and 10: Journals

The most common publication outlets for prior knowledge research are shown in Figure 7. The *Journal of Educational Psychology* published 1.21 % of all 9,158 journal publications in the corpus, followed by *Learning and Instruction* (1.19 %) and *Computers in Human Behavior* (1.05 %). For about 28 % of publications in our database, there was no journal title, for example, because they were books (as indicated by metadata in the databases).

Using bivariate correlations (based on mean document-topic-probabilities by journal), we further investigated journal similarity of the most common outlets. Using a network representation of these correlations and community detection, we found the five groups of journals. Based on the included journal titles we labeled the categories *Learning and Instruction* (comprising 10.29% of all journal contributions), *Cognitive Processes* (7.17%), *STEM learning* (4.09%), *Language Learning* (4.08%), and *Other Journals* (0.90%). These percentages add up to about 27%, as we focused on the most common outlets. An interactive and zoomable network visualization of the journal communities (Figure S4) as well as Table S1 with the most frequent journal titles is provided in the ESM.

Research Questions 11 and 12: Scientific Disciplines and Subdisciplines

Table 4 shows the relative frequencies of the scientific disciplines of the 13,507 publications in our corpus. Scientific disciplines other than psychology and education
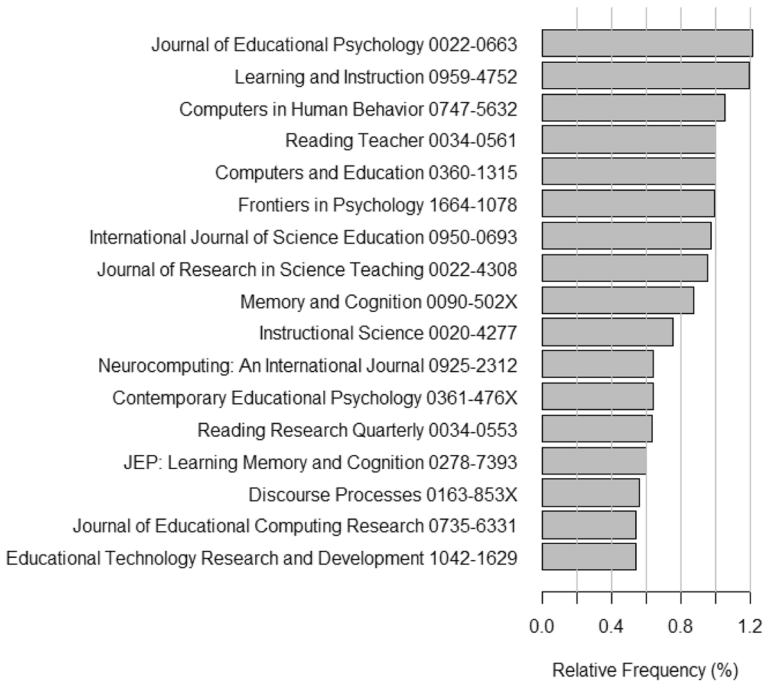
Figure 7 Title and ISSN of journals in which studies on prior knowledge are most frequently published. *Note*. Only journals with a minimum relative frequency of 0.5 % are shown

sciences are classified according to Web of Science Research Areas. Only disciplines with a minimum relative frequency of 0.2 % of all publications are shown. The results in Table 4 are shaped by the fact that we were interested only in studies on prior knowledge in human learners. Accordingly, we only searched databases for psychological and educational research. We did not search databases for research on prior knowledge in computer programs or databases, for example, in artificial intelligence research. We included studies from domains other than psychology and education (e.g., from journals on physics or computer science) in our corpus when these studies were listed in the (psychological and educational) databases we searched.

Among the included studies, psychology comprised a larger share (64.52%) than educational research (35.82%). The most frequently collaborating disciplines were communication (2.17%), engineering (1.26%), linguistics (0.77%), and neurosciences and neurology (0.70%). Engineering is linked to research on prior knowledge, because many systems are designed and built by engineers to be used by humans, for example, driver-assistance systems and power plant control stations. The users' prior knowledge needs to be taken into account in the system design to ensure that the users understand the systems, can use them effectively and safely, and are satisfied with the product (Gopher & Kimchi, 1989; Wickens et al., 2021). Closely related relevant disciplines are *Human Factors Research* and *Cognitive Ergonomics*. Presumably, physics is listed as a discipline in Table 4 because journals for physical research sometimes also publish articles about how the research findings can be taught in higher education.

Table 4 Disciplines of the included studies on prior knowledge and their relative frequency

Discipline	Relative frequency
Psychology	64.52 %
Education Sciences	35.82 %
Communication	2.17 %
Engineering	1.26 %
Linguistics	0.77 %
Neurosciences and Neurology	0.70 %
Behavioral Sciences	0.47 %
Rehabilitation	0.39 %
Computer Science	0.37 %
Health Care Sciences and Services	0.29 %
Psychiatry	0.25 %
Business and Economics	0.21 %
Physics	0.20 %

Multiple assignments are possible when different disciplines collaborate for a publication. Only disciplines with a minimum relative frequency of 0.2 % of all publications are shown

It was a deliberate decision that our literature search did not include databases for the disciplines of computer science and artificial intelligence. The question of how prior knowledge about a domain can be formally modeled and used in the design of computer software is a relatively large field of research. For example, many studies investigated algorithms to detect objects in pictures, for example, tumors in MRI scans or military airplanes in radar data. Prior knowledge, for example, about typical but irrelevant image artifacts or about the appearances of the target object under different conditions, can strongly improve these systems. Yet, these studies were excluded from our literature search and are thus also not shown in Table 4, because the topics, theories, and methods in this branch of research strongly differ from research on prior knowledge in human learners.

The information in our database (i.e., the more fine-grained classifications in PsycInfo and PSYNDEX) allowed us to investigate the subdisciplines of Psychology. The subfields were derived from the APA classification system and Web of Science Research Areas (see the ESM for details). Within psychology, prior knowledge is mainly addressed in educational psychology (27.32 % of all publications in our corpus), experimental psychology (13.44 %), followed by developmental psychology (4.92 %), and clinical psychology (3.2 %); cf. Figure S3. These results agree with the results of the topic analysis (Figure 2) and the results of the journal analysis (Table S1), which also indicate that educational research on learning and instruction as well as cognitive experimental research are central in prior knowledge research.

Discussion

Our bibliometric analysis allowed us to investigate the terms used in the titles, abstracts, and metadata of 13,507 publications on prior knowledge and human learning. Due to this unusually broad scope of our analyses, the implications of

our findings are also on a broad and abstract level. There were seven main findings. We discuss these main findings and their implications for theory, practice, or research methodology, in turn.

Prior Knowledge Is a Cross-Sectional Research Topic in Research on Learning.

The first main finding concerned the most common topics in research on prior knowledge (Research Question 1). The results indicated that prior knowledge is not one research topic among many others; rather, it is a cross-sectional topic that has been investigated in the context of 23 research topics and their many subtopics. The 23 research topics (listed in Table 1) cover many major research areas of psychology and educational science and related disciplines. Accordingly, prior knowledge research has been published in a wide variety of mostly psychological and educational journals. Within Psychology, there were topics from all major subfields, that is, educational, experimental, developmental, clinical, industrial and organizational, health, and social psychology.

These findings support the conclusion that there is a broad interest in the research on prior knowledge. This conclusion aligns with qualitative literature reviews stipulating that the influence of prior knowledge on learning is "perhaps one of the most influential ideas to emerge in cognitive psychology during the past 25 years" (Hambrick & Engle, 2002, p. 339) and that "it would be difficult to find an educational study conducted within the past two decades that does not implicitly or explicitly acknowledge the powerful impact of one's existing knowledge base on subsequent learning and development" (Dochy & Alexander, 1995, p. 225).

Prior Knowledge Research Can Help to Establish Bridges Between Neurocognitive, Cognitive, and Educational Research.

A second main finding was that, in Figure 2, the topic community *Education* was not directly linked to the topic community *Cognitive Processes*. Instead, the topic community *Education* was linked to the psychological topic community *Learning Environments*, which was linked to the neuro-cognitive topic community *Cognitive Processes*. This finding is plausible because educational studies rarely directly refer to (neuro-)cognitive studies and more frequently refer to educational-psychological research about learning environments. Educational-psychological studies frequently refer to cognitive (and sometimes neuro-cognitive) studies as theoretical backgrounds. Thus, educational-psychological research connected applied educational research and foundational (neuro-)cognitive research in our network.

This pattern mirrors a suggestion made by Bruer (1997) in the seminal article "Education and the Brain: A Bridge Too Far." Bruer argues that research on the neurocognitive foundations of learning and on classroom learning differ so strongly in their theories, methods, and levels of analysis that neurocognitive findings usually do not have *direct* implications for the design of effective school instruction (see Bowers, 2016, for a more recent literature review supporting this

argument). Bruer suggests that two bridges are needed to connect neurocognitive and educational research: Brain-imaging research on neurocognitive processes can inform behavioral psychological research on cognitive learning mechanisms and how they can be stimulated by learning environments. This psychological research can then inform studies on the design of educational interventions in applied settings, for example, school instruction.

Our findings imply that research on prior knowledge is highly useful for building such bridges from Neuroscience to Psychology and from Psychology to Education because the term prior knowledge is a cross-sectional research topic used in neurocognitive, behavioral psychological, and applied educational research, even though these three fields differ in theories and methods. For example, a review of the neurocognitive mechanisms underlying the acquisition and application of prior knowledge concluded that the medial temporal lobe and the medial prefrontal cortex play central roles in these processes (Brod et al., 2013). This finding has no direct instructional implications, but it helps to better understand the cognitive processes through which humans learn and use abstract knowledge. This understanding can then be used to derive hypotheses about how the acquisition of abstract and generalizable prior knowledge can be fostered through learning environments and, ultimately, through school instruction (e.g., Goldwater & Schalk, 2016). Since there are still large gaps between neurocognitive, psychological, and educational research (Horvath & Donoghue, 2016), future research should increase efforts to use the concept of prior knowledge to further the integration of these fields.

There Are Potentials for the Further Integration of Research on Prior Knowledge.

If researchers want to use research on prior knowledge to better integrate scientific disciplines, then it is important that research on prior knowledge is integrated into itself. Even though research on prior knowledge, in general, is represented by a closely interlinked topic network (Figure 2), our findings indicate several areas wherein prior knowledge research could profit from better integration of literature strands. Such integration would be beneficial because good scientific theories explain as many phenomena as possible with as few assumptions as possible. Integrating previously unconnected lines of research can help to increase the explanatory breadth and parsimony of theories. When two topics are unconnected in our network this does not prove that they can and should be integrated, only that it might be beneficial to investigate to what extent they might profit from a stronger integration.

One area that might profit from a greater integration of topics is research on reading literacy, mathematical literacy, and science literacy. These competencies are foundations for solving everyday life problems, work success, and participation in society. For example, they are indispensable for understanding payment plans for cell phone contracts, information on how to reduce the risk of life-threatening infections, or the outcomes of elections. For these reasons, about 80 countries triannually assess reading, mathematics, and science literacy in hundred-thousands of 15-year-old students in the PISA study (OECD, 2019). Given this foundational importance

of reading, mathematics, and science, we were surprised that these topics were relatively isolated in our topics network (Figure 2). Together, they only had six edges connecting them with the other 18 topic nodes in the network. There was not a single direct connection between research on reading, mathematics, and science and any of the topics in the community *Nonacademic Settings* (i.e., *Businesses and Organizations, Patient and Medical Education, Communication, and Preschool*). Publications connecting these topics exist and were in our database (e.g., Elster, 1995), but they numbered too few to influence the algorithm constructing the topic network. These results indicate that the field of research on prior knowledge can be moved forward by better integration of studies on reading, mathematics, and science with studies on the design of learning environments and learning in applied settings. One example, among many, is the recent demonstration that patient education and communication can be made more effective by tailoring instructors' and physicians' language to patients' prior knowledge and health literacy (Schillinger et al., 2017, 2021).

At least three areas of prior knowledge research, that is, intelligence, infants' core knowledge, and psychoeducation, are not represented in our topic network and accordingly are not well integrated with the research topics in our network. Prior knowledge is relevant in all three fields. For example, Cattel's investment theory posits that learners invest their general intelligence to acquire crystallized intelligence and domain-specific knowledge, all of which influence further learning (Watrin et al., 2022). Infants' core knowledge, for example, about objects and quantities, manifests as early as the first weeks and months of infants' lives, long before infants understand language (Spelke & Kinzler, 2007). Core knowledge is hypothesized to have evolved because it guides attention toward relevant features of situations, thus accelerating further learning (Stahl & Feigenson, 2019). Whereas many studies show that infants have prior knowledge, only a few studies directly investigate how this prior knowledge influences further learning. This might be why the term prior knowledge is seldom used in this literature and was not detected as a topic by our algorithm. Psychoeducation effectively improves patients' health and well-being and is a component of many psychotherapeutic approaches (Lukens & McFarlane, 2006; Simonsmeier, et al., 2021). Previous studies suggest that psychoeducation needs to build on patients' prior knowledge (Harvey et al., 2014), but more research on this question is needed. In sum, the concept of prior knowledge plays potentially important roles in the works of literature on intelligence, core knowledge, and psychoeducation, respectively, but these works of literature rarely explicitly refer to the prior knowledge literature synthesized in the present study. We see it as an important task for future research to better integrate these topics with prior knowledge research.

Research on skill acquisition and expert performance were not identified as topics in our analyses because the term prior knowledge is rarely used explicitly in these contexts. Notwithstanding this fact, these research areas are already well-integrated with other areas of research on prior knowledge and learning. The role of prior knowledge in the acquisition of skills and expert performance has been investigated for decades and is well-understood (Gobet, 2005; Huang et al., 2023).

There is also a need for better integration of prior knowledge research with research on non-cognitive (or socio-emotional) learning processes. Empirical findings show

that knowledge acquisition is intertwined with motivation (Song et al., 2015), emotion (Pekrun et al., 2017), and attitudes (Cacioppo et al., 1992). Studies on the relations of prior knowledge with these non-cognitive constructs have been published and are among the publications analyzed in our study. Some of these publications are highly influential and have high citation rates (e.g., Pintrich, 1993; Tobias, 1994). However, the number of terms relating to these lines of research in our text corpus was so small that the topic model did not identify them as separate topics. We see it as a task for future research to better integrate prior knowledge research with these lines of research.

Domain-Specific Prior Knowledge Is of Domain-General Importance.

The fact that prior knowledge is a cross-sectional research topic is also reflected by the content domains of the investigated knowledge. Our fourth main finding is that prior knowledge had been investigated in more than 20 content domains. Among them were mathematics, reading, and sciences, which many see as foundational for successful learning in all other academic domains (OECD, 2019). There were also many other school subjects (e.g., computer science, social studies), topics of study programs in higher education (e.g., medicine, psychology, engineering), topics of cultural and societal importance (e.g., health, the environment, and social studies), and non-academic domains (i.e., art, music, and sex). The meta-analysis by Simonsmeier, et al. (2021) found that the correlations between prior knowledge and learning outcomes (posttest knowledge or knowledge gains) did not differ or differed only minimally between content domains. Overall, we are not aware of any domain where domain-specific knowledge is irrelevant to learning in that domain.

How is it possible that domain-specific prior knowledge is relevant for so many research topics and in so many content domains, even though they differ so strongly?—A likely answer is that learning processes in all of these domains have a common denominator. This common denominator is the brain with its information-processing architecture which humans use for learning in all domains (Anderson et al., 2004; Sweller et al., 1998). For example, domain-specific prior knowledge guides attention and supports memory encoding. These mechanisms are deeply rooted in the brain and, thus, have been observed in numerous content domains (Bellana et al., 2021; Brod et al., 2013; Kim & Rehder, 2010). Essentially, it is not possible to process information with zero prior knowledge, and thus, prior knowledge drives all human processing tasks.

More general reviews of cognitively grounded domain-general principles for the facilitation of domain-specific learning processes are given, for example, by Ericsson et al. (1993; see also Ericsson & Kintsch, 1995), Anderson and Schunn (2000), and Schneider and Stern (2010a). In sum, in all domains, domain-specific knowledge affects domain-specific learning processes. In this sense, domain-specific prior knowledge is of domain-general relevance for understanding learning. This insight has theoretical and methodological implications. It explains why many studies investigating domain-specific learning processes derive their hypotheses from domain-general theories (see Figure 4). It also explains why it can be justified to average over and compare effect sizes obtained in different content domains to search for

domain-general patterns in domain-specific learning processes (e.g., as done in the recent meta-analysis on prior knowledge effects; Simonsmeier, et al., 2021).

Many Theories Refer to Prior Knowledge, but There Are No Theories of Prior Knowledge.

A fifth main finding is the high importance of prior knowledge for understanding learning on a theoretical level. This is demonstrated by the high number of theories that refer to prior knowledge effects on learning. Figure 4 lists 28 theoretical backgrounds used in studies of prior knowledge. These theoretical approaches are highly diverse. Some are theories of knowledge acquisition processes, others are more general learning theories that are not primarily concerned with knowledge acquisition, and still others are even more general fields of research not even primarily concerned with learning. This wide variety further supports our Conclusion 1, that prior knowledge is a widely used cross-sectional research topic.

However, this theoretical eclecticism also demonstrates that there is not a singular theory of prior knowledge that is generally used in prior knowledge research because it would explain the most relevant phenomena. There is also not a small number of competing theories dominating the field. The theoretical approaches listed in Figure 4 were not even formulated to explain the effects of prior knowledge on learning. They were formulated to explain more general phenomena, for instance, multimedia learning, categorization, self-regulation, racism, semiotics, or action regulation. They include prior knowledge as one of many variables to explain these more general phenomena. Consequently, theoretical insights about the role of prior knowledge in learning are fragmented. Researchers trying to predict or explain the effects of prior knowledge on learning need to draw insights from various theoretical accounts to get the full picture. Our Conclusion 3 was that there is a need to further integrate strands of the empirical research literature on prior knowledge. Conclusion 5 extends this insight by showing that there is also a need for further theoretical integration of prior knowledge research.

Types, Characteristics, and Representations of Prior Knowledge Affect Learning But Are Difficult to Operationalize.

A sixth main finding was that only a minority of studies on prior knowledge and learning referred to types, characteristics, and representations of knowledge in their titles and abstracts. This finding is surprising because there are good theoretical reasons and empirical support for the view that the quality of prior knowledge sometimes influences learning stronger than the sheer quantity of prior knowledge (Brod, 2021; diSessa et al., 2004; de Jong & Ferguson-Hessler, 1996; McCarthy & McNamara, 2021). One possible explanation for the relatively infrequent references to types, characteristics, and representations of prior knowledge is that they are difficult to measure. For example, what percentage of a learner's knowledge in a domain is explicit or implicit, conceptual or procedural, fragmented or integrated, abstract or concrete? Such a question is difficult to answer because knowledge is a theoretical

construct that cannot be directly observed. Knowledge and its properties need to be inferred from patterns of overt behavior, for example, answers to test questions. This is difficult for two reasons. First, when learners construct the answer to a test question, they can combine several pieces of knowledge through reasoning. In particular, with realistic, complex learning materials, the number of these pieces of knowledge can be high. The pieces can differ in their types, characteristics, and representations. This makes it difficult to infer the underlying knowledge from overt behavior. Second, when answering a test question, learners do not exclusively use the knowledge they already have in their long-term memory. They can derive new conclusions through reasoning during answer construction. The characteristics of this spontaneously created knowledge are not necessarily the same as the characteristics of the knowledge stored in long-term memory. For example, when learners answer a question designed to measure procedural knowledge in their long-term memory, they can derive new procedures from their conceptual knowledge (Schneider & Stern, 2010b). When learners can explain the relation between two concepts in a test, it remains unclear whether their explanation indicates the integration of these concepts in their network of knowledge in long-term memory or whether the participants spontaneously created the explanation in response to the test question (diSessa et al., 2004). Similar measurement problems have been discussed for virtually all types, characteristics, and representations of knowledge (e.g., Keren & Schul, 2009; Maie & DeKeyser, 2019; Melnikoff & Bargh, 2018).

Researchers are increasingly turning to item-response theory and latent variable modeling to infer the underlying knowledge structures from manifest behavior. However, despite some progress in this area (Ellis & Roever, 2017; Hickendorff et al., 2018; Schneider et al., 2011), there are still no psychometrically sound, generally accepted standard instruments for measuring types, characteristics, and qualities of knowledge (see Figure 5). An implication of our findings is that the construction of such instruments remains a central challenge for future research on prior knowledge (cf. McCarthy & McNamara, 2021).

There Is a Lack of Experiments Investigating the Causal Effects of Prior Knowledge on Knowledge Gains.

The high number of studies on prior knowledge and learning and the high proportion of empirical studies among them make it seem as if the relation between prior knowledge and learning was empirically well-investigated. However, this is not the case. Less than 10% of the studies surveyed here had an experimental design (Figure 6). We cannot infer from our data whether these studies were experimental investigations of prior knowledge or experimental investigations of other variables merely using pre-existing individual differences in prior knowledge as a covariate. Information on this question can be found in the meta-analysis on prior knowledge (Simonsmeier, et al., 2021), where more than 10,000 titles and abstracts and 1491 full texts were screened for relevance. Only 493 of these studies allowed the empirical investigation of prior knowledge effects on learning because they assessed knowledge before and after learning with continuous and

objective quantitative measures. Among these studies, not a single one investigated how experimentally induced changes in prior knowledge affected gains in knowledge from pretest to posttest.

Therefore, our previous conclusion that prior knowledge is a widely used cross-sectional research topic should not be misinterpreted as implying the effects of prior knowledge on learning are well investigated and well understood. This is not the case. Many of the basic cognitive processes underlying knowledge acquisition processes are well understood, for example, category learning, executive functions, attention, and the neurobiological bases of memory (e.g., Bisaz et al., 2014; Schlegelmilch et al., 2021; Shing & Brod, 2016). However, the question under which conditions and how strongly prior knowledge causally affects learning with complex, realistic materials in field settings has not been fully answered yet. The question cannot be answered by correlational studies because prior knowledge correlates with intelligence, socio-economic status, interest, and many other variables which might induce spurious correlations between prior knowledge and learning outcome measures. The question also cannot be answered by studies using posttest knowledge as dependent variables because the association of pretest knowledge with posttest knowledge merely indicates the stability of differences in prior knowledge, not the predictive power of prior knowledge for learning (Simonsmeier, et al., 2021).

Overall, these findings show a dire need for more studies investigating how experimentally induced differences in prior knowledge causally affect knowledge gains in a subsequent learning phase. Ideally, these studies would be conducted as field experiments with complex, realistic materials to preserve the ecological validity of their findings. In addition, future studies should consider the multidimensionality of prior knowledge, going beyond estimates of amounts of knowledge, toward a better understanding of how different types and characteristics of knowledge impact processing, understanding, and learning.

Limitations of our Approach

We queried the databases via OVID and Web of Science, which offer to search in titles, abstracts, and metadata such as keywords. As no full-text search was available, we could only detect research papers that included the entities of interest (such as "prior knowledge" but also specific theories and models) in these data fields. Access to full texts would clearly raise the frequencies of references to specific theories such as Kintsch's Construction-Integration model (Kintsch, 1988). This means that our frequency analyses are limited to publications reporting the entities of interest "top level" (i.e., titles, abstracts, or keywords) rather than only in the manuscript. Therefore, our results underestimate the *true* number of frequencies. Although Google Scholar does perform searches in full texts, it does not have the option to export the results and is known to be prone to false positives (e.g., Bramer et al., 2016; García-Pérez, 2010). The recently published and freely accessible "General Index" (cf. Else, 2021) might be a data corpus to overcome this limitation. However, its quality and eligibility for scientometric endeavors such as the present study are yet to be determined. Besides, our study exemplifies the limitations of literature databases for content-level frequency analyses and emphasizes the merits

of the open science movement: With freely accessible full-text or at least text fragments such as n-grams, more insights are possible. Nevertheless, our results indicate relations and rankings of various entities and constructs and pave the way for future research. For instance, the theoretical constructs detected by our strictly *bottom-up* approach could be the starting point for subsequent *top-down* reviews of theories in the context of prior knowledge.

For our explorative frequency analyses, we used binary checks of indicating terms (e.g., "secondary education" is mentioned yes/no) and a semi-automated approach of keywords-in-context (KWIC) analysis and selection of valid n-grams (e.g., three_p_model) in addition to available database fields (e.g., age groups). While this approach is scalable to corpora of any size, it must be noted that it is prone to false positives (e.g., the study has a focus on adolescents although both adolescents and adults were addressed) or missings due to spelling variants or more indirect mentions (e.g., "our sample comprised young learners as well as their parents"; "young learners" and "parents" would not have been detected with our pattern-matching-based approach). If, for instance, the age of participants in empirical studies of prior knowledge is of particular interest, manual coding of respective studies would be necessary. Therefore, the results of this exploratory bibliometric study should be regarded as a general overview of the field, not as an in-depth analysis with definite frequencies of theories, sample characteristics, and so on.

Regarding the KWIC approach, abstract passages referring to methods were determined using text windows after terms like "METHOD:," "investigated," "examined"). This of course is an approximation to extracting the concrete method sections in scientific abstracts of variable structure as observed in our corpus. While several natural language processing approaches for sentence classification have been developed (e.g., Gonçalves et al., 2019; Jiang et al., 2019), they were evaluated using PubMed records of clinical trials or computer science abstracts. Hence, a generalizable model (or at least trained with abstracts from different scientific disciplines) along with ready-to-implement software would be of particular interest to entity recognition in bibliometric analyses such as the present study.

Finally, our study is limited in that we took a bird-eye perspective on research on prior knowledge and learning. We aggregated information from 13,507 publications. This forced us to leave out many details of these studies. Thus, the aim of our study cannot be to replace the 13,507 publications. Instead, it is our aim to draw attention to these original studies and to help our readers to find the original studies most relevant for their research. Seen this way, there is no tension between broad research syntheses and precise empirical studies. Their respective strengths and weaknesses complement each other well.

Conclusion

Bibliometric analysis proved to be a valuable tool for examining studies of prior knowledge across multiple domains and disciplines. It provided a relatively unbiased view of the literature because it relies on automated methods of synthesizing the vast landscape of literature on the topic. Given the largeness of the research on

prior knowledge, it would be impossible to glean the characteristics of the research and the connections (and lack thereof) between the disciplines, as well as the gaps in the literature without bibliometric analyses. Within a single discipline (e.g., education), one might perceive prior knowledge as a relatively well-understood construct. By examining the construct from multiple lenses and across multiple pieces of literature, educators and researchers can glean a broader understanding of prior knowledge and what the construct implies. The findings indicate that prior knowledge is a cross-sectional research topic widely referred to in theories of learning and practical approaches for improving learning. Prior knowledge is investigated in many strands of the neurocognitive, psychological, and educational literature and has the potential for better integrating these fields. The findings further highlight the need for better integration of theoretical approaches of prior knowledge and learning, the need for psychometrically sound measures of types, characteristics, and representations of knowledge, and for randomized controlled field experiments testing the causal effects of prior knowledge on changes in knowledge over time.

Declarations

This study was preregistered on PsychArchives: https://doi.org/10.23668/psycharchives.4903.

Conflict of Interest The authors declare no competing interests.

References

Alexander, P. A., & Dochy, F. J. R. C. (1995). Conceptions of knowledge and beliefs: A comparison across varying cultural and educational communities. *American Educational Research Journal, 32*(2), 413–442. https://doi.org/10.3102/00028312032002413

Alexander, P. A., Schallert, D. L., & Hare, V. C. (1991). Coming to terms: How researchers in learning and literacy talk about knowledge. *Review of Educational Research, 61*(3), 315–343. https://doi.org/10.3102/00346543061003315

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review, 111*(4), 1036–1060. https://doi.org/10.1037/0033-295x.111.4.1036

Anderson, J. R., & Schunn, C. D. (2000). Implications of the ACT-R learning theory: No magic bullets. In R. Glaser (Ed.), *Advances in instructional psychology: Educational design and cognitive science* (Vol. 5, pp. 1–33). Lawrence Erlbaum Associates Publishers.

Batterink, L. J., Reber, P. J., Neville, H. J., & Paller, K. A. (2015). Implicit and explicit contributions to statistical learning. *Journal of Memory and Language, 83*, 62–78. https://doi.org/10.1016/j.jml.2015.04.004

Bellana, B., Mansour, R., Ladyka-Wojcik, N., Grady, C., & Moscovitch, M. (2021). The influence of prior knowledge on the formation of detailed and durable memories. *Journal of Memory and Language, 121*, 104264. https://doi.org/10.1016/j.jml.2021.104264

Biggs, J. B. (1993). From theory to practice: A cognitive systems approach. *Higher Education Research & Development, 12*(1), 73–85. https://doi.org/10.1080/0729436930120107

Bisaz, R., Travaglia, A., & Alberini, C. M. (2014). The neurobiological bases of memory formation: From physiological conditions to psychopathology. *Psychopathology, 47*(6), 347–356. https://doi.org/10.1159/000363702

Bischof, J. M., & Airoldi, E. M. (2012). Summarizing topical content with word frequency and exclusivity. In *Proceedings of the 29th International Conference on Machine Learning* (pp. 201–208).

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment, 2008*(10), P10008. https://doi.org/10.1088/1742-5468/2008/10/p10008

Bowers, J. S. (2016). The practical and principled problems with educational neuroscience. *Psychological Review, 123*(5), 600–612. https://doi.org/10.1037/rev0000025

Bramer, W. M., Giustini, D., & Kramer, B. M. R. (2016). Comparing the coverage, recall, and precision of searches for 120 systematic reviews in Embase, MEDLINE, and google scholar: A prospective study. *Systematic Reviews, 5*(1). https://doi.org/10.1186/s13643-016-0215-7

Broaders, S. C., Cook, S. W., Mitchell, Z., & Goldin-Meadow, S. (2007). Making children gesture brings out implicit knowledge and leads to learning. *Journal of Experimental Psychology: General, 136*(4), 539–550. https://doi.org/10.1037/0096-3445.136.4.539

Brod, G. (2021). Toward an understanding of when prior knowledge helps or hinders learning. *Npj Science of Learning, 6*, 24. https://doi.org/10.1038/s41539-021-00103-w

Brod, G., Werkle-Bergner, M., & Shing, Y. L. (2013). The influence of prior knowledge on memory: A developmental cognitive neuroscience perspective. *Frontiers in Behavioral Neuroscience, 7*. https://doi.org/10.3389/fnbeh.2013.00139

Bruer, J. T. (1997). Education and the brain: A bridge too far. *Educational Researcher, 26*(8), 4–16. https://doi.org/10.3102/0013189x026008004

Cacioppo, J. T., Marshall-Goodell, B. S., Tassinary, L. G., & Petty, R. E. (1992). Rudimentary determinants of attitudes: Classical conditioning is more effective when prior knowledge about the attitude stimulus is low than high. *Journal of Experimental Social Psychology, 28*(3), 207–233. https://doi.org/10.1016/0022-1031(92)90053-m

Colman, A. M. (2015). Content domain. In A Dictionary of Psychology (4th ed., p. 164). Oxford University

diSessa, A. A. (2017). Knowledge in pieces: An evolving framework for understanding knowing and learning. In T. G. Amin & O. Levrini (Eds.), *Converging Perspectives on Conceptual Change*. Routledge.

diSessa, A. A., Gillespie, N. M., & Esterly, J. B. (2004). Coherence versus fragmentation in the development of the concept of force. *Cognitive Science, 28*(6), 843–900. https://doi.org/10.1207/s15516709cog2806_1

Dochy, F. J. R. C., & Alexander, P. A. (1995). Mapping prior knowledge: A framework for discussion among researchers. *European Journal of Psychology of Education, 10*(3), 225–242. https://doi.org/10.1007/bf03172918

Ellis, R., & Roever, C. (2017). The measurement of implicit and explicit knowledge. *The Language Learning Journal, 49*(2). https://doi.org/10.1080/09571736.2018.1504229

Else, H. (2021). Giant, free index to world's research papers released online. *Nature*. https://doi.org/10.1038/d41586-021-02895-8

Elster, C. (1995). Importations in preschoolers' emergent readings. *Journal of Reading Behavior, 27*(1), 65–84. https://doi.org/10.1080/10862969509547869

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102*(2), 211–245. https://doi.org/10.1037/0033-295X.102.2.211

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*(3), 363–406. https://doi.org/10.1037/0033-295x.100.3.363

Fruchterman, T. M. J., & Reingold, E. M. (1991). Graph drawing by force-directed placement. *Software: Practice and Experience, 21*(11), 1129–1164. https://doi.org/10.1002/spe.4380211102

Gao, C., Baucom, L. B., Kim, J., Wang, J., Wedell, D. H., & Shinkareva, S. V. (2019). Distinguishing abstract from concrete concepts in supramodal brain regions. *Neuropsychologia, 131*, 102–110. https://doi.org/10.1016/j.neuropsychologia.2019.05.032

García-Pérez, M. A. (2010). Accuracy and completeness of publication and citation records in the web of science, PsycINFO, and google scholar: A case study for the computation of h indices in psychology. *Journal of the American Society for Information Science and Technology, 61*(10), 2070–2085. https://doi.org/10.1002/asi.21372

Gobet, F. (2005). Chunking models of expertise: Implications for education. *Applied Cognitive Psychology, 19*(2), 183–204. https://doi.org/10.1002/acp.1110

Goldwater, M. B., & Schalk, L. (2016). Relational categories as a bridge between cognitive and educational research. *Psychological Bulletin, 142*(7), 729–757. https://doi.org/10.1037/bul0000043

Gonçalves, S., Cortez, P., & Moro, S. (2019). A deep learning classifier for sentence classification in biomedical and computer science abstracts. *Neural Computing and Applications, 32*(11), 6793–6807. https://doi.org/10.1007/s00521-019-04334-2

Gopher, D., & Kimchi, R. (1989). Engineering Psychology. *Annual Review of Psychology, 40*(1), 431–455. https://doi.org/10.1146/annurev.ps.40.020189.002243

Greene, J. A., Sandoval, W. A., & Bråten, I. (2016). An introduction to epistemic cognition. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Handbook of Epistemic Cognition* (pp. 1–15). Routledge.

Greve, A., Cooper, E., Tibon, R., & Henson, R. N. (2019). Knowledge is power: Prior knowledge aids memory for both congruent and incongruent events, but in different ways. *Journal of Experimental Psychology: General, 148*(2), 325–341. https://doi.org/10.1037/xge0000498

Hambrick, D. Z., & Engle, R. W. (2002). Effects of domain knowledge, working memory capacity, and age on cognitive performance: An investigation of the Knowledge-Is-Power hypothesis. *Cognitive Psychology, 44*(4), 339–387. https://doi.org/10.1006/cogp.2001.0769

Harvey, A. G., Lee, J., Williams, J., Hollon, S. D., Walker, M. P., Thompson, M. A., & Smith, R. (2014). Improving outcome of psychosocial treatments by enhancing memory and learning. *Perspectives on Psychological Science, 9*(2), 161–179. https://doi.org/10.1177/1745691614521781

Hickendorff, M., Edelsbrunner, P. A., McMullen, J., Schneider, M., & Trezise, K. (2018). Informative tools for characterizing individual differences in learning: Latent class, latent profile, and latent transition analysis. *Learning and Individual Differences, 66*, 4–15. https://doi.org/10.1016/j.lindif.2017.11.001

Horvath, J. C., & Donoghue, G. M. (2016). A bridge too far – revisited: Reframing Bruer's neuroeducation argument for modern science of learning practitioners. *Frontiers in Psychology, 7*. https://doi.org/10.3389/fpsyg.2016.00377

Huang, Y., Brusilovsky, P., Guerra, J., Koedinger, K., & Schunn, C. D. (2023). Supporting skill integration in an intelligent tutoring system for code tracing. *Journal of Computer Assisted Learning, 39*(2), 477–500. https://doi.org/10.1111/jcal.12757

Jiang, X., Zhang, B., Ye, Y., & Liu, Z. (2019). A hierarchical model with recurrent convolutional neural networks for sequential sentence classification. *Natural Language Processing and Chinese Computing*, 78–89. https://doi.org/10.1007/978-3-030-32236-6_7

Kendeou, P., & O'Brien, E. J. (2015). Prior knowledge: Acquisition and revision. In P. Afflerbach (Ed.), *Handbook of Individual Differences in Reading: Reader, Text, and Context* (pp. 151–163). Routledge. https://doi.org/10.4324/9780203075562

Keren, G., & Schul, Y. (2009). Two is not always better than one: A critical evaluation of Two-System theories. *Perspectives on Psychological Science, 4*(6), 533–550. https://doi.org/10.1111/j.1745-6924.2009.01164.x

Kim, S., & Rehder, B. (2010). How prior knowledge affects selective attention during category learning: An eyetracking study. *Memory & Cognition, 39*(4), 649–665. https://doi.org/10.3758/s13421-010-0050-3

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review, 95*(2), 163–182. https://doi.org/10.1037/0033-295x.95.2.163

Linn, M. C. (2006). The knowledge integration perspective on learning and instruction. In R. K. Sawyer (Ed.), *The Cambridge handbook of: The learning sciences* (pp. 243–264). Cambridge University Press.

Liu, J., Zhang, R., Geng, B., Zhang, T., Yuan, D., Otani, S., & Li, X. (2019). Interplay between prior knowledge and communication mode on teaching effectiveness: Interpersonal neural synchronization as a neural marker. *NeuroImage, 193*, 93–102. https://doi.org/10.1016/j.neuroimage.2019.03.004

Lukens, E. P., & McFarlane, W. R. (2006). Psychoeducation as evidence-based practice: Considerations for practice, research, and policy. In A. R. Roberts & K. R. Yeager (Eds.), *Foundations of Evidence-based Social Work Practice* (pp. 291–313). Oxford University Press.

Maie, R., & DeKeyser, R. M. (2019). Conflicting evidence of explicit and implicit knowledge from objective and subjective measures. *Studies in Second Language Acquisition, 42*(2), 359–382. https://doi.org/10.1017/s0272263119000615

Mayer, R. E. (2014). Cognitive theory of multimedia learning. *The Cambridge Handbook of Multimedia Learning*, 43–71. https://doi.org/10.1017/cbo9781139547369.005

McCarthy, K. S., & McNamara, D. S. (2021). The multidimensional knowledge in text comprehension framework. *Educational Psychologist, 56*(3), 196–214. https://doi.org/10.1080/00461520.2021.1872379

McNamara, D. S., & Magliano, J. P. (2009). Towards a comprehensive model of comprehension. In B. Ross (Ed.), *The psychology of learning and motivation* (Vol. 51). Elsevier. https://doi.org/10.1016/S0079-7421(09)51009-2

Medin, D. L., Lynch, E. B., & Solomon, K. O. (2000). Are there kinds of concepts? *Annual Review of Psychology, 51*(1), 121–147. https://doi.org/10.1146/annurev.psych.51.1.121

Melnikoff, D. E., & Bargh, J. A. (2018). The mythical number two. *Trends in Cognitive Sciences, 22*, 280–293. https://doi.org/10.1016/j.tics.2018.02.001

Mimno, D., Wallach, H. M., Talley, E. M., Leenders, M., & McCallum, A. (2011). Optimizing semantic coherence in topic models. In R. Barzilay & M. Johnson (Eds.), *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing* (pp. 262–272). Association for Computational Linguistics.

Moehring, A., Schroeders, U., & Wilhelm, O. (2018). Knowledge is power for medical assistants: Crystallized and fluid intelligence as predictors of vocational knowledge. *Frontiers in Psychology, 9.* https://doi.org/10.3389/fpsyg.2018.00028

OECD. (2019). *PISA 2018 results (Volume I): What students know and can do.* https://doi.org/10.1787/5f07c754-en

Ormrod, J. E. (2019). *Human learning* (8th ed.). Pearson.

Pekrun, R., Lichtenfeld, S., Marsh, H. W., Murayama, K., & Goetz, T. (2017). Achievement emotions and academic performance: Longitudinal models of reciprocal effects. *Child Development, 88*(5), 1653–1670. https://doi.org/10.1111/cdev.12704

Pintrich, P. R., Marx, R. W., & Boyle, R. A. (1993). Beyond cold conceptual change: The role of motivational beliefs and classroom contextual factors in the process of conceptual change. *Review of Educational Research, 63*(2), 167–199. https://doi.org/10.3102/00346543063002167

Press de Jong, T., & Ferguson-Hessler, M. G. M. (1996). Types and qualities of knowledge. *Educational Psychologist, 31*(2), 105–113. https://doi.org/10.1207/s15326985ep3102_2

Renoult, L., Irish, M., Moscovitch, M., & Rugg, M. D. (2019). From knowing to remembering: The Semantic–Episodic distinction. *Trends in Cognitive Sciences, 23*(12), 1041–1057. https://doi.org/10.1016/j.tics.2019.09.008

Rittle-Johnson, B., Schneider, M., & Star, J. R. (2015). Not a One-Way street: Bidirectional relations between procedural and conceptual knowledge of mathematics. *Educational Psychology Review, 27*(4), 587–597. https://doi.org/10.1007/s10648-015-9302-x

Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S. K., Albertson, B., & Rand, D. G. (2014). Structural topic models for open-ended survey responses. *American Journal of Political Science, 58*(4), 1064–1082. https://doi.org/10.1111/ajps.12103

Schillinger, D., Duran, N. D., McNamara, D. S., Crossley, S. A., Balyan, R., & Karter, A. J. (2021). Precision communication: Physicians' linguistic adaptation to patients' health literacy. *Science Advances, 7*(51). https://doi.org/10.1126/sciadv.abj2836

Schillinger, D., McNamara, D., Crossley, S., Lyles, C., Moffet, H. H., Sarkar, U., Duran, N., Allen, J., Liu, J., Oryn, D., Ratanawongsa, N., & Karter, A. J. (2017). The next frontier in communication

and the ECLIPPSE study: Bridging the linguistic divide in secure messaging. *Journal of Diabetes Research, 2017*, 1–9. https://doi.org/10.1155/2017/1348242

Schlegelmilch, R., Wills, A. J., & von Helversen, B. (2021). A cognitive category-learning model of rule abstraction, attention learning, and contextual modulation. *Psychological Review*. https://doi.org/10.1037/rev0000321

Schneider, M., Rittle-Johnson, B., & Star, J. R. (2011). Relations among conceptual knowledge, procedural knowledge, and procedural flexibility in two samples differing in prior knowledge. *Developmental Psychology, 47*(6), 1525–1538. https://doi.org/10.1037/a0024997

Schneider, M., & Stern, E. (2010a). The cognitive perspective on learning: Ten cornerstone findings. In Organisation for Economic Co-Operation and Development (OECD),*The nature of learning: Using research to inspire practice* (pp. 69–90). OECD.

Schneider, M., & Stern, E. (2010b). The developmental relations between conceptual and procedural knowledge: A multimethod approach. *Developmental Psychology, 46*(1), 178–192. https://doi.org/10.1037/a0016701

Schneider, W. J., & McGrew, K. S. (2012). The Cattell-Horn-Carroll model of intelligence. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary Intellectual Assessment: Theories, Tests, and Issues* (pp. 99–144). Guilford Press.

Shing, Y. L., & Brod, G. (2016). Effects of prior knowledge on memory: Implications for education. *Mind, Brain, and Education, 10*(3), 153–161. https://doi.org/10.1111/mbe.12110

Shtulman, A., & Valcarcel, J. (2012). Scientific knowledge suppresses but does not supplant earlier intuitions. *Cognition, 124*(2), 209–215. https://doi.org/10.1016/j.cognition.2012.04.005

Simonsmeier, B. A., Flaig, M., Deiglmayr, A., Schalk, L., & Schneider, M. (2021). Domain-specific prior knowledge and learning: A meta-analysis. *Educational Psychologist, 57*(1), 31–54. https://doi.org/10.1080/00461520.2021.1939700

Simonsmeier, B. A., Flaig, M., Simacek, T., & Schneider, M. (2021). What sixty years of research says about the effectiveness of patient education on health: A second order meta-analysis. *Health Psychology Review, 16*(3), 450–474. https://doi.org/10.1080/17437199.2021.1967184

Song, H., Kalet, A., & Plass, J. (2015). Interplay of prior knowledge, self-regulation and motivation in complex multimedia learning environments. *Journal of Computer Assisted Learning, 32*(1), 31–50. https://doi.org/10.1111/jcal.12117

Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental Science, 10*(1), 89–96. https://doi.org/10.1111/j.1467-7687.2007.00569.x

Stadler, M., Sailer, M., & Fischer, F. (2021). Knowledge as a formative construct: A good alpha is not always better. *New Ideas in Psychology, 60*. https://doi.org/10.1016/j.newideapsych.2020.100832

Stahl, A. E., & Feigenson, L. (2019). Violations of core knowledge shape early learning. *Topics in Cognitive Science, 11*(1), 136–153. https://doi.org/10.1111/tops.12389

Stricker, J., Vogel, S. E., Schöneburg-Lehnert, S., Krohn, T., Dögnitz, S., Jud, N., Spirk, M., Windhaber, M. C., Schneider, M., & Grabner, R. H. (2021). Interference between naïve and scientific theories occurs in mathematics and is related to mathematical achievement. *Cognition, 214*, 104789. https://doi.org/10.1016/j.cognition.2021.104789

Sweller, J., van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*(3), 251–296. https://doi.org/10.1023/a:1022193728205

Taber, K. S. (2017). The use of Cronbach's Alpha when developing and reporting research instruments in science education. *Research in Science Education, 48*(6), 1273–1296. https://doi.org/10.1007/s11165-016-9602-2

Tenison, C., Fincham, J. M., & Anderson, J. R. (2016). Phases of learning: How skill acquisition impacts cognitive processing. *Cognitive Psychology, 87*, 1–28. https://doi.org/10.1016/j.cogpsych.2016.03.001

Tobias, S. (1994). Interest, prior knowledge, and learning. *Review of Educational Research, 64*(1), 37–54. https://doi.org/10.3102/00346543064001037

Ullén, F., Hambrick, D. Z., & Mosing, M. A. (2015). Rethinking expertise: A multifactorial gene–environment interaction model of expert performance. *Psychological Bulletin, 142*(4), 427–446. https://doi.org/10.1037/bul0000033

Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition, 92*(1–2), 231–270. https://doi.org/10.1016/j.cognition.2003.10.008

van den Broek, P., Young, M., Tzeng, Y., & Linderholm, T. (1999). The landscape model of reading: Inferences and the online construction of memory representation. In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 71–98). Lawrence Erlbaum Associates Publishers.

Vosniadou, S. (Ed.). (2013). *International handbook of research on conceptual change*. Routledge. https://doi.org/10.4324/9780203154472

Watrin, L., Schroeders, U., & Wilhelm, O. (2022). Structural invariance of declarative knowledge across the adult lifespan. *Psychology and Aging, 37*(3), 283–297. https://doi.org/10.1037/pag0000660

Wickens, C. D., Helton, W. S., Hollands, J. G., & Banbury, S. (2021). *Engineering psychology and human performance* (5th ed.). Routledge. https://doi.org/10.4324/9781003177616

Wiemer-Hastings, K., & Xu, X. (2005). Content differences for abstract and concrete concepts. *Cognitive Science, 29*(5), 719–736. https://doi.org/10.1207/s15516709cog0000_33

Yang, Z., Algesheimer, R., & Tessone, C. J. (2016). A comparative analysis of community detection algorithms on artificial networks. *Scientific Reports, 6*(1). https://doi.org/10.1038/srep30750