Review

Empirical Findings on Learning Success and Competence Development at Learning Factories: A Scoping Review

Nine Reining * and Simone Kauffeld

Department of Industrial, Organizational and Social Psychology, Institute of Psychology, Technische Universität Braunschweig, 38106 Braunschweig, Germany
* Correspondence: n.reining@tu-braunschweig.de

Abstract: To meet the changing competence requirements for employees in engineering professions, education and training need to adapt accordingly. Learning factories offer various possibilities to design or integrate practice-oriented learning into training measures. Whether this approach in fact facilitates learning and competence development is rarely investigated. For this reason, the objective of this scoping review is to analyze and summarize the existing empirical findings on learning success and competence development in learning factories regarding their evaluation methods and results. Following standardized guidelines (PRISMA, JBI) for scoping reviews, 12 databases were researched. The literature screening led to the identification of 24 publications included in the final analysis. The results indicate that a variety of evaluation methods are used to assess learning and competences at learning factories and that criteria of all four competence facets (professional, methodological, social, and self-competence) can be enhanced at learning factories in general. As many of the identified studies show potential for improvement regarding the quality of the used methods and analysis of results, further studies on these topics are needed. Evaluations should be integrated into all training measures at learning factories to ensure learning success and competence development and to be able to readjust design, structure, and didactics where necessary.

Keywords: learning factory; learning; competence; competence development; engineering education; scoping review

1. Introduction

Increasing digitization and technologization in the professional context are leading to a change in the competence requirements for employees, with both existing areas of competence changing and new areas of competence being added (cf. [1,2]). As recent studies and literature reviews show, this explicitly applies to engineering professions (cf. [3–5]). In order to prepare future employees for this altered work environment, the required competences need to be developed during training.

Learning factories provide an innovative and practice-oriented format in which learning success and competence development are to be facilitated in educational measures such as teaching and training. Initially, learning factories were designed for the engineering sector. Today, learning factories exist around the world on a variety of topics, including lean, agile, industry 4.0, building design, product, and software development [6]. Accordingly, they can be designed very diversely. Abele [6] provides a comprehensive definition of learning factories that takes this aspect into account: A learning factory is the changeable physical and/or virtual environment of a real value chain with several technical and organizational processes, in which a product and/or service is manufactured in several steps (stations). Users can carry out their own actions at the learning factory to initiate processes and intervene in the factory's operating modes. This enables a comprehensive experience of cause and effect in a practice-oriented environment.

In many cases, however, the fundamental goal of knowledge transfer and competence development is not sufficiently considered in the design and development of learning

factories [7]. Therefore, it cannot be assumed without further consideration that learning factories in their individual designs are per se suitable to enhance learning success and competence development of the users. Against this background, it is of great importance to carry out comprehensive evaluations in the context of university teaching, education, and training at learning factories in order to assess their benefits in this respect and, if necessary, to readjust their design, structure, and didactics.

The possibilities for evaluation in the context of learning factories are very diverse. However, publications often only report on short feedbacks from participants that have been obtained after the teaching and training courses. This feedback often says little to nothing about a possible learning success and/or competence development of the participants, but mostly maps the satisfaction of the participants with the teaching or training. In contrast, solid empirical studies on the extent to which participants (sustainably) build knowledge and competences through learning factory-related educational measures seem to be rather rare. This raises the question of the extent to which learning factories are suitable in practice for promoting learning success and competence development among participants and thus supporting their preparation for the professional world. At the same time, an overview of evaluation methods used in this context can create an incentive for learning factory managers to also investigate the learning success and competence development of participants at their learning factory.

Therefore, this review aims to identify studies in which learning success and competence development in learning factories have been investigated and presented in a comprehensible way based on data and results. The focus of the review is to analyze and summarize the existing empirical findings on learning success and competence development at learning factories regarding their methods and their results.

Research Questions

1. How have learning success and competence development been empirically assessed to date in the context of learning factories?
2. Which empirical evidence already exists on learning success and competence development at learning factories?

Due to the limited amount of literature available on the topic, the quality of studies and findings is considered only secondary and is not a screening criterion. For this reason, this paper was designed in the form of a scoping review [8]. In this way, the current state of research on learning success and competence development at learning factories can be comprehensively mapped and existing research gaps can be identified. A search for reviews related to learning factories in the Scopus database and the Bielefeld Academic Search Engine (BASE) (cf. [9]) at the beginning of the subsequent review process revealed that no other review was available that addresses empirical research on learning success and/or competence development at learning factories. The literature search of the present review, which is described in detail below, also revealed no existing review on the topic.

2. Materials and Methods

Systematic reviews address specific, well-defined research questions and therefore need strict inclusion criteria and a high quality of included studies while scoping reviews have a broader scope and therefore may have more extensive inclusion criteria and focus less on the quality of included studies [8,10]. Hence, this review was designed in the form of a scoping review to address the broadly-defined research questions. To maintain scientific standards, the review process followed the guidelines on scoping reviews of the JBI Manual for Evidence Synthesis [11], which is based on the methodological framework of Arksey and O'Malley [10]. The JBI guideline is consistent with the PRISMA Extension for Scoping Reviews (PRISMA-ScR), which was first published in 2018 as an extension to the well-known PRISMA guideline for systematic reviews to further improve both methodological and reporting quality of scoping reviews [12,13]. The detailed review process is shown in Figure 1.
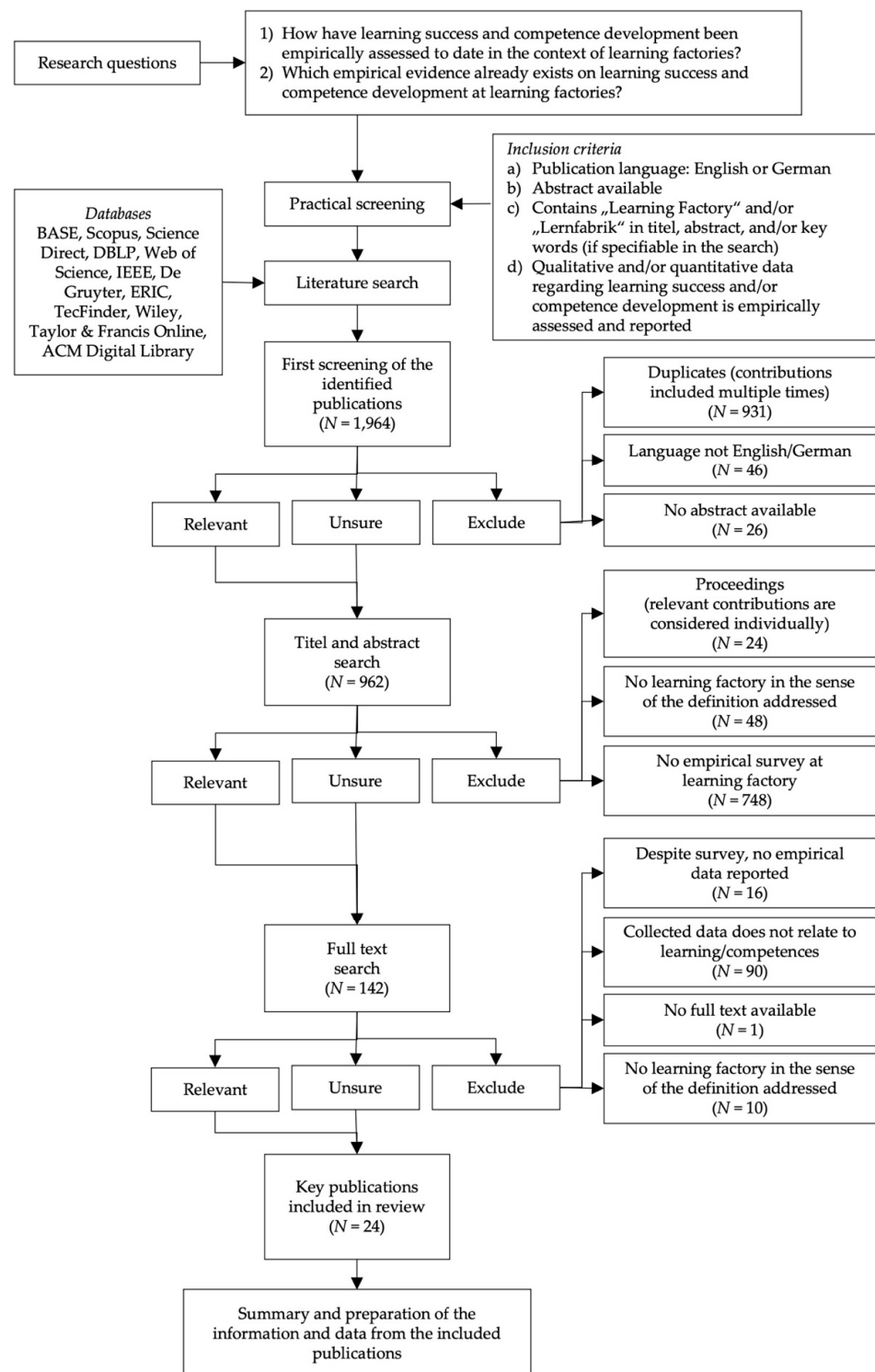
Figure 1. Review process, flowchart based on Page et al. [12].

The literature search followed three inclusion criteria: (1) The publication is written in English or German. Publications in German were included because a majority of learning factories is operated in German-speaking countries. (2) An abstract of the publication is available. (3) The publication contains empirical qualitative and/or quantitative data on learning success and/or competence development at learning factories, including results.

In [12], due to the limited amount of available literature on the topic as well as the broad, multidisciplinary research field, no explicit exclusion criteria (e.g., regard-

ing participants) were defined. The search included all types of publications, including grey literature.

For the literature search, the following databases were identified as relevant based on Gusenbauer and Haddaway [9]: ACM Digital Library, Bielefeld Academic Search Engine (BASE), Digital Bibliography and Library Project (DBLP), Education Resources Information Center (ERIC), IEEE Xplore Digital Library, Scopus, Science Direct, Web of Science, and Wiley Online Library. In addition, the databases of De Gruyter and Taylor and Francis Online were consulted for literature search, as the proceedings of relevant conferences (e.g., Conference on Learning Factories) are published here. TecFinder was further used as a relevant database for the inclusion of publications on science and technology in German. The initial search included all search entries listed in the databases which were initially retrieved and then filtered for duplicates and relevance to the research questions.

Searches were conducted from 25 May to 9 June 2022 using the search terms "Lernfabrik" and "learning factory" in titles, keywords, and abstracts, or in the entire document if the database did not allow limitations. No further limitation of the search was made, since "Lernfabrik" in German and "learning factory" in English are established terms. Moreover, the terms already contain the words "lern" (German for "learning") and "learning", respectively. Testing with different search strings on one of the databases showed that no further restrictions were possible regarding the topic area of learning success at learning factories without excluding possibly relevant publications, especially since evaluations were not the main topic of some of the possibly relevant publications. Further, no restrictions regarding the time period were made, as the aim of the scoping review is to give an overview of all relevant publications. In total, 1964 publications were identified. The review of the identified publications was performed by Reviewer 1 following Waffenschmidt et al. [14].

In a first screening step, all duplicates were identified ($N = 931$) and excluded from the further review process. In addition, publications were excluded for which no abstract was available ($N = 26$) and which were not published in German or English based on the review of title and abstract ($N = 46$).

Subsequently, the abstracts of the remaining 962 publications were reviewed. Publications that do not concern learning factories in the sense of the definition were eliminated ($N = 48$), as well as proceedings that contain a large number of publications as a collection ($N = 24$), since the publications contained therein-if relevant-have been identified individually in the search. In addition, all publications that, according to the abstract, clearly do not address an empirical assessment in the context of learning or competences were excluded ($N = 748$).

A full text analysis of the remaining 142 publications focused on identifying those publications that address empirical data on learning and/or competences at learning factories. For this purpose, full text copies of the publications were researched and requested from authors and organized in a literature database (Mendeley). For one publication, no full text could be obtained, so it could not be analyzed further and had to be excluded. In the full text analysis of the remaining publications, those were excluded in which no data collection on learning or competences in the learning factory context are reported ($N = 90$) and no learning factory in the sense of the definition is used during the collection of the data ($N = 10$). In addition, publications were excluded according to which there were data collections on the topics, but these were not reported further and/or the results were only summarized briefly (usually a maximum of two to three sentences) without data ($N = 16$). At the end of the process, 24 publications remained, which were included in the review.

From the publications included in the review, the information relevant to answering the research questions was extracted and organized in Table 1. In the following, the results of the review are systematically presented and assessed in terms of their relevance and limitations.

3. Results

This section presents the characteristics of the 24 publications that were included in the review as they empirically address learning success and/or competence development at learning factories. Following the structure of the previously mentioned research questions that guided this review, the used methods and findings of the included studies are further summarized. Table 1 displays an overview of the studies included in the review.

3.1. Characteristics of Publications and Studies

Twenty-four empirical publications on the topics of learning success and competence development at learning factories were identified. Of these publications, one was published in 1998, four were published in the five years between 2010 and 2014, ten between 2015 and 2019, and nine since 2020. The publications mostly include studies of learning factories in Europe ($n = 18$), followed by Africa ($n = 3$) and North America ($n = 2$). One study interviewed participants at learning factories in both Europe and Asia. In twenty-two cases the publication language is English and in two cases it is German.

Three publications are Master's theses and one publication is a doctoral dissertation. Of the identified publications, two publications each refer to the same underlying studies: Ogorodnyk [15] and Granheim [16], as well as Makumbe [17] and Makumbe et al. [18]. The review thus includes 24 publications with 22 studies (see Table 2).

Table 1. Overview of studies included in the review (sorted by year of publication).

Authors, Year (Source of Evidence)	Country (Type of Learning Factory)	Intervention (Duration)	Number and Origin of Participants	Type of Measurement (Measurement Time Points)	How Learning/ Competences Assessed?	Findings
Morell de Ramirez, Velez-Arocho, Zayas-Castro, and Torres, 1998 [19] (Conference paper)	Puerto Rico (physical)	Seminar (one semester)	$N = 181$; students (multi-disciplinary): $n = 122$, faculty: $n = 14$, industry: $n = 42$, other: $n = 3$	Questionnaire for (self-) evaluation (post)	Questionnaire: 5 items on learning/competences (in the questionnaire for students), 5-point scale from "strongly agree" to "strongly disagree"	Responses that "strongly agree" and "agree": • Communication skills emphasized: 89% of industry, 71% of faculty, and 80% of students • Teamwork skills emphasized: 93% of industry, 93% of faculty, and 97% of students • Better understanding of engineering: 78% of students • More confident in solving real life problems: 78% of students • More confident in their ability to teach themselves: 80% of students
Cachay, Wennemer, Abele, and Tenberg, 2012 [20] (Conference paper)	Germany (physical)	Trainings (4.5 h) Experimental group: 60 min. own experience at the learning factory Control group: 60 min. lesson with practical examples shown by instructor	$N = 25$ (students in engineering); experimental group: $n = 16$, control group: $n = 9$	Knowledge test questionnaire (pre, post); practical application (post)	Knowledge test: Eight open questions regarding action-independent knowledge and 5 open questions regarding action-substantiating knowledge, answers for analysis evaluated on 5-point scale from 1 "no answer/not correct" to 5 "correct" Practical application: Time measurement, support evaluation	Knowledge test comparison post vs. pre: • Action-independent knowledge: experimental group improved in absolute difference by 27.5% more than control group • Action-substantiating knowledge: experimental group improved in absolute difference by 47.5% more than control group Practical application (working task at the learning factory): • Experimental group: divided into three groups, each took 10 to 30 min. without support • Control group: divided into two groups, each took approximately 60 min. with much support
Kesavadas, 2013 [21] (Conference paper)	USA (virtual)	Project work (14 weeks)	$N = 38$ (Bachelor's and Master's students in engineering)	Questionnaire for self-evaluation (post)	Questionnaire: Two items on learning/competences, 3-point scale from "fully agree" to "do not agree"	Responses that "fully agree" and "partially agree": • Virtual Learning Factory helped to better understand the concepts taught in class: 42% fully agree, 47% partially agree • Virtual Factory project helped experience the progress better than the more traditional individual group/class project formats: 47% fully agree, 37% partially agree

Table 1. *Cont.*

Authors, Year (Source of Evidence)	Country (Type of Learning Factory)	Intervention (Duration)	Number and Origin of Participants	Type of Measurement (Measurement Time Points)	How Learning/ Competences Assessed?	Findings
Riffelmacher, 2013 [22] (Doctoral thesis)	Germany (physical, digital, and virtual)	Qualification (results based on several trainings; duration not specified)	$N =$ approximately 13 (professionals from industries of variant-rich series production and with background in industrial engineering)	Questionnaire for self-evaluation (post, ex post after 6 month)	Questionnaire: Post 7 to 14 questions for each learning module, ex post 6 questions in total, 6-point scale from 1 "fulfilled" to 6 "not fulfilled"	Post questionnaire: • As far as reported, the mean values of the individual items ranged between $M = 1.1$ and $M = 2.3$. Mean values of relevant items from the ex post questionnaire: • "Have you personally used methods and tools to manage turbulence?"— $M = 1.2$ • "In order to manage turbulences, have you carried out the planning steps foreseen in the qualification concept that fall within your area of responsibility?"—$M = 1.4$ • "Were you able to transfer the content and approach to turbulence management into your daily work routine and thus apply it?"—$M = 1.4$
Ovais, Liukkunen, and Markkula, 2014 [23] (Conference paper)	Finland (physical Software Factory)	Project course (7 projects at the software factory; duration not specified)	$N = 19$ (Master's students in engineering)	Questionnaire for self-evaluation (post)	Questionnaire: 8 items for improvement in eight competence areas, 5-point scale from 1 "strongly disagree" to 5 "strongly agree"	Reported mean values of the competences: • Effective task management— $M = 3.68$ • Solving complex problems—$M = 3.58$ • Sharing responsibilities—$M = 3.74$ • Developing a shared vision— $M = 3.68$ • Building a positive relationship— $M = 3.84$ • Negotiating with other groups— $M = 3.89$ • Use of rational argument to persuade others—$M = 3.32$ • Resolving conflict—$M = 3.21$
Plorin, Jentsch, Hopf, and Müller, 2015 [24] (Conference paper)	Germany (physical and digital)	Trainings (different group compositions and contents from a module pool; duration not specified)	$N = 31$ (employees from various backgrounds)	Questionnaire for self-evaluation (pre, ex post 2 month after training)	Questionnaire: 12 knowledge and competence areas queried, 6-point scale from 1 "very good" to 6 "very bad"	Improvement of the mean values of the effectiveness of knowledge transfer in all queried areas: Environmental influences, energy policy, energy needs assessment of buildings, forms of energy in general, energy balance, energy generation, energy conversion, energy distribution, energy recovery, compressed air leaks, approaches to increase energy efficiency, energy management according to DIN ISO 50001
Zinn, Güzel, Walker, Nickolaus, Sari, and Hedrich, 2015 [25] (Journal article)	Germany (physical and simulated)	Training (3 days)	$N = 35$ (trainees, dual students, inexperienced employees from mechatronics and service technology); Group 1: $n = 7$, Group 2: $n = 13$, Group 3: $n = 5$	Questionnaire for self-evaluation (pre, post); knowledge test questionnaire (pre, post); practical application (post)	Questionnaire for self-evaluation: 37 items on how confident participants feel with regard to certain areas, 5-point scale from 1 "not confident at all" to 5 "very confident" Knowledge test: 37 open questions (partly adapted based on Zinke et al. [26] and Abele et al. [27]), answers for analysis evaluated Computer simulation: Processing 4 error cases, processing documented, + answer 5 competence items	Self-attributed professional knowledge: • Participants rate themselves significantly better after the training than before ($p < 0.001$, Cohen's $d = 0.73$) Increase professional knowledge from pre to post (t-test): • Overall group: pre 48%, post 60% ($p < 0.001$, $d = 0.71$) • Group 1: pre 39%, post 52% ($p < 0.001$, $d = 0.80$) • Group 2: pre 54%, post 67% ($p < 0.001$, $d = 0.92$) • Group 3: pre 50%, post 60% ($p < 0.01$, $d = 0.55$) Increase in fault diagnosis competence from pre to post (t-test): • Overall group: pre 31%, post 46% ($p < 0.001$, $d = 0.76$) • Group 1: pre 29%, post 49% ($p < 0.01$, $d = 1.1$) • Group 2: pre 30%, post 35% ($p = 0.305$, $d = 0.30$) • Group 3: pre 35%, post 55% ($p < 0.001$, $d = 1.0$)

Table 1. *Cont.*

Authors, Year (Source of Evidence)	Country (Type of Learning Factory)	Intervention (Duration)	Number and Origin of Participants	Type of Measurement (Measurement Time Points)	How Learning/ Competences Assessed?	Findings
Granheim, 2016; Ogorodnyk, 2016 [15,16] (Master's theses)	Norway (physical)	Case study (duration not specified)	$N = 11$ (Master's students in engineering); Group 1: $n = 6$, Group 2: $n = 5$	Practical application (continuously); group interview (post, ex post after 1 week)	Time measurement: How long did assembly 1 vs. assembly 2 take (between assembly cycles, participants could adjust assembly stations according to their own preferences) Group interviews: Open questions on knowledge and application, content evaluation of the answers	Time measurement: • Group 1 (one pair of roller skis): assembly cycle 1 (first roller ski) = 25 min., assembly cycle 2 (second roller ski) = 13 min. • Group 2 (two pairs of roller skis): assembly cycle 1 = 35 min. (first roller ski) assembly cycle 2 = 16 min. (three roller ski with approximately 5 min. each) Interviews post: Participants of both groups were able to answer all questions (e.g., changes/improvements related to the theory; What does kaizen mean to you now? Is it better to use pull or push?) Interviews ex post: Participants of both groups could answer all questions (What theoretical aspects do you remember?; What types of waste do you remember?; How did you apply this part of the theory in the activity?)
Henning, Hagedorn-Hansen, and von Leipzig, 2017 [28] (Journal article)	South Africa (physical)	Games (duration not specified)	$N = 368$ (students); Beer Game: $n = 195$, Lego Car Game: $n = 78$, Train Game: $n = 12$, Off-Roader LEGO Car Game: $n = 11$, The Fresh Connection: $n = 72$	Questionnaire for self-evaluation (post)	Questionnaire: Knowledge estimation using knowledge dimension levels based on the revised Bloom's taxonomy [29] with 4 levels: 1 "factual", 2 "conceptual", 3 "procedural", 4 "etacognitive" (no information given on scales or items)	70% of participants of Train Game chose levels 3 and 4 and over 90% of participants of all other games chose levels 3 and 4, indicating that the students have learned certain terms or theories through the experience
Makumbe, 2017; Makumbe, Hattingh, Plint, and Esterhuizen, 2018 [17,18] (Master's thesis, Conference paper)	South Africa (physical)	Training (2 days plus coaching and implementation in the workplace)	$N = 26$ (mining employees); Group 1: $n = 12$ (engineers, production crew, business improvement specialists), Group 2: $n = 14$ (operations support services)	Knowledge test questionnaire (pre, post); observation (continuously); Interviews (ex post); practical application (ex post with two times of measurement)	Knowledge test: Pre 5 and post 11 open questions regarding the understanding of the 5 lean principles, answers for evaluation evaluated on 4-point scale from 1 "not understood" to 4 "well understood" Observation: By researchers using structured observation tables Interview: Unstructured open questions regarding knowledge and application Practical application: Implementation on the job, analysis of statistical process control charts	Knowledge test: • Group 1: significant improvement of 3 out of 5 lean principles, but improvement overall not significant (t-test) • Group 2: significant improvement of 4 out of 5 lean principles, improvement overall significant (t-test) Observations: • Support the results of the knowledge tests for both groups Interview ex post (both groups combined): • Participants still remembered the theoretical concepts they learned during the activity, and were able to define all of them and give examples of how they were applied during the activity • Participants indicated they could apply the knowledge they had acquired when needed as they knew how and where it could be used Process data on implementation in the company: • Group 1: Improvement resulted in reduction in variability, change in production is significant (t-test) • Group 2: No data on implementation

Table 1. *Cont.*

Authors, Year (Source of Evidence)	Country (Type of Learning Factory)	Intervention (Duration)	Number and Origin of Participants	Type of Measurement (Measurement Time Points)	How Learning/ Competences Assessed?	Findings
Glass, Miersch, and Metternich, 2018 [30] (Conference paper)	Germany (physical)	Practical exercises as part of a lecture (duration not specified)	N = 30 to 45 (Master's students in engineering); no direct control group, but grade comparison with students who did not attend the intervention	Observation (continuously); questionnaire for self-evaluation (retrospective pre, post); exam grades (post)	Questionnaire: Self-assessment of the level of knowledge, items developed based on Erler et al. [31], 6-point scale from 1 "very good" to 6 "deficient" Observation: Depending on exercise, 25–40% of the participants, work samples during the exercises evaluated using the method of Schaper [32]; Exam grades	Questionnaire: • Yamazumi: Before the exercise the grades vary, and the ratings are spread. After the exercise, the grade two was selected more than 50% of the time. Grade one, two, and three together make up around 95% of all given grades. • Value-Stream-Mapping: improvement through the exercise is rated with over one whole grade point in the practical measures from pre $M = 3.25$ to post $M = 2.10$, in "evaluation and interpretation" with a half and in the social competences with a quarter of a grade point; t-test: overall improvement from pre to post is significant Observation: • Correlation with questionnaire data is 50% Exam grades: • 90% of the exam questions addressed in the practical exercises • t-tests: students, who attended at least 75% of the exercises, achieve more points than students, who did not visit the exercises; students who attended more than 75% of all exercises did not achieve a higher score on a task with no correlating exercise
Balve and Ebert, 2019 [33] (Conference paper)	Germany (physical)	Project work (15 weeks)	N = 45 (former Bachelor's students in engineering)	Questionnaire for self-evaluation (ex post)	Questionnaire: Out of 42 competences participants asked to choose a maximum of 6 competences, that they believe were specifically strengthened using the learning factory	Overall ranking of competences of at least 30%: Organizational skills (47%), time management (42%), interdisciplinary thinking (36%), recognizing interrelations (31%), problem solving ability (31%); social competences were hardly selected
Reining, Kauffeld, and Herrmann, 2019 [34] (Conference paper)	Germany (physical)	Seminar with practical, research-based group work at the learning factory (one semester)	N = 8 (Master's students in engineering); Group 1: n = 4, Group 2: n = 4	Video data (continuously); questionnaire for self-evaluation (pre, post)	Video: Group work recorded, conversations divided into sense units and analyzed Questionnaire: Scale on affinity for technology (items negatively worded) adapted from Richter et al. [35], 5-point scale from 1 "not true at all" to 5 "very true"	Video: • Both groups: Interactions mostly address professional competences ($M = 50\%$), followed by social competences ($M = 32.2\%$), methodological competences ($M = 6.9\%$), and self-competences ($M = 3.8\%$) • Most sense units allocated to the criterion "technical knowledge, knowledge of science and mechanics" (30.7%), "analytical thinking" (10.1%), and "communication skills" (31.3%) • Differences in the distribution of addressed competences: a) between groups, b) depending on whether groups work theoretically in the seminar room, on the computer or practically at the learning factory Questionnaire: • Participants rated their affinity for technology post ($M = 1.83$, $SD = 0.60$) somewhat more positively than pre ($M = 2.04$, $SD = 0.77$) Overall: • Comparison with existing competence model for intervention: all competences of the model addressed in videos or positively assessed in questionnaire

Table 1. *Cont.*

Authors, Year (Source of Evidence)	Country (Type of Learning Factory)	Intervention (Duration)	Number and Origin of Participants	Type of Measurement (Measurement Time Points)	How Learning/ Competences Assessed?	Findings
Devika, Raj, Venugopal, Thiede, Herrmann, and Sangwan, 2020 [36] (Conference paper)	India/Germany (physical)	No specific intervention (duration not specified)	$N = 14$ (students in engineering and instructors who have practical experience at a learning factory)	Interview (ex post)	Interview: Semi-structured to identify transversal competences that can be strengthened at the learning factory, recorded, transcribed, and coded with MAXQDA	Interview: • Learning factories provide an environment to develop four transversal competences: (1) teamwork (interaction, problem-solving, leadership), (2) communication (oral presentation, foster communication and interaction, cross-cultural communication), (3) creativity and innovation (idea generation, product generation, self-exposure), and (4) lifelong learning (reflection, acquiring and learning, initiating) • Learners were able to use transversal competences in all three phases (planning, execution, reflection) at the learning factory, although to varying degrees
Juraschek, Büth, Martin, Pulst, Thiede, and Herrmann, 2020 [37] (Conference paper)	Germany (physical)	GameJam (3 days)	$N = 18$ (no further information given)	Questionnaire for self-evaluation (pre, post)	Questionnaire: No information given on scales or items, 5-point scale from 1 "low consent" to 5 "high consent"	Cumulative self-assessment of relevant competences: Value increased from pre $M = 3.0$ to post $M = 3.5$
Omidvarkarjan, Conrad, Herbst, Klahn, and Meboldt, 2020 [38] (Conference paper)	Switzerland (physical)	Training (2 days)	$N = 7$ (Bachelor's and Master's students in engineering)	Written feedback (post)	Written feedback: Participants put in writing their thoughts on their perceived learnings with regard to the agile principles, results were analyzed qualitatively	Number of mentions of learned agile principles, that are consistent with the principles taught during training: Frequent interactions = 3, test-driven development = 3, self-organizing teams = 4, iterative progression = 7, continuous improvement = 1, customer involvement = 2, accommodating change = 1, simplicity = 1
Sieckmann, Petrusch, and Kohl, 2020 [39] (Conference paper)	Germany (physical)	Training (duration not specified)	$N = 66$ (Master's students in engineering); 4 experimental groups, which differed in terms of problem (case study, own problem) and social form of interaction (small group, plenary): Group 1: $n = 18$, Group 2: $n = 15$, Group 3: $n = 16$, Group 4: $n = 17$	Questionnaire for self-evaluation (pre, post) practical application (post)	Questionnaire: 1 item each for 7 lean methods, 5-point scale from 1 "unknown" to 5 "can moderate" Practical application: Application of A3 method in individual work, participants fill out a template for processing, which is evaluated	Questionnaire (evaluation of all participants): • Significant positive change in understanding of all elements of the learning unit from pre to post • A3 method: increase in reported ability to apply the method from pre 5.4% to post 66.2%, with an additional 21.6% confident in moderating the method • Ishikawa diagram: ability to apply and moderate increased from pre 67.5% to post over 90% • Must-criteria analysis: pre unknown to 87.7%, post 60% at least could apply the method Practical application: Mean values for practical problem solving depending on the previous experimental group Group 1: small group, own problem— $M = 90.3\%$ Group 2: small group, case study— $M = 73.6\%$ Group 3: plenum, own problem— $M = 77.0\%$ Group 4: plenum, case study—$M = 68.0\%$

Table 1. *Cont.*

Authors, Year (Source of Evidence)	Country (Type of Learning Factory)	Intervention (Duration)	Number and Origin of Participants	Type of Measurement (Measurement Time Points)	How Learning/ Competences Assessed?	Findings
Adam, Hofbauer, and Stehling, 2021 [40] (Journal article)	Austria (physical)	Training (1.5 days)	$N = 234$ (employees of various backgrounds)	Questionnaire for self-evaluation (post)	Questionnaire: Three items on understanding of lean tools (scale not given), four items on challenges (4-point scale from "very simple" to "difficult"), 4 items on types of help (4-point scale from "very helpful" to "less helpful")	• Post: lean tools understood by 47% to 74% of participants • Significant correlation between understanding and a) an alternation between theory and practice and b) a do-it-yourself approach (Chi^2, both $p = 0.00$) • Over 50% of participants refuse to implement even the simplest lean tool; shop floor members have more doubts than management and office staff (Chi^2, $p = 0.00$) • Significant correlation between the intention to transfer and: (a) understanding of content (Chi^2, $p = 0.00$), (b) how easy lean tools were to apply in training (Chi^2, $p = 0.00$), (c) participants' higher prior lean experience (Chi^2, $p = 0.00$)
Mahmood, Otto, Kuts, Terkaj, Modoni, Urgo, Colombo, Haidegger, Kovacs, and Stahre, 2021 [41] (Conference paper)	not specified (virtual)	Workshop (duration not specified)	$N = 15$ (students)	Questionnaire for self-evaluation (post)	Questionnaire: Evaluation of the achieved learning level and self-assessment, 5-point scale from 1 "very dissatisfied" to 5 "very satisfied"	Questionnaire: Evaluation of the reached learning level of methods and digital tools with "very satisfied" and "satisfied": • Workflow definition: 80% • Performance evaluation: 70% • Virtual modelling: 60% • Self-evaluation: acquired competences: 80%
Roll and Ifenthaler, 2021 [42] (Journal article)	Germany (physical)	Vocational training (8 × 45 min)	$N = 71$ (trainees of electronic professions); experimental group 1 (much time at learning factory): $n = 18$, experimental group 2 (medium amount of time at learning factory): $n = 24$, control group (no time at learning factory): $n = 21$	Knowledge test questionnaire (pre, post, ex post after 4 weeks)	Knowledge test: • (a) Multidisciplinary digital competences: pre 13, post 12, and ex post 12 open questions, responses analyzed for evaluation with the help of content analyses; (b) Subject-related competences: pre 7, post 6 and ex post 10 open questions, responses rated for evaluation on 5-point scale from 0 "answer blank" to 4 "complete and perfect answer"	Multidisciplinary digital competences: • No significant interaction effect of interaction level and time of survey on multidisciplinary digital competences ($df = 4$, $SS = 10{,}091$, $H = 3.37$, $p = 0.50$, $\eta^2 = 0.02$) • Significant results of a small interaction effect of interaction level and time of survey on the competence dimension problem solving ($df = 4$, $SS = 28{,}812$, $H = 9.66$, $p = 0.05$, $\eta^2 = 0.05$), control group scored lower than both experimental groups Subject-related competences: • Interaction effect of interaction level and time of survey on technical competences was significant and had a medium strength ($df = 4$, $SS = 4436$, $H = 14.88$, $p = 0.01$, $\eta^2 = 0.08$), with control group scores mostly lower than both experimental groups' scores • Experimental group 1 and experimental group 2 each had performance peaks at post measurement and improved significantly here compared to the pre survey (experimental group 1: $Diff = -0.70$, $p = 0.00$, $r = 0.33$; experimental group 2: $Diff = -2.42$, $p < 0.00$, $r = 0.77$). • Experimental group 2 improved significantly in the long term, with a large effect at ex post survey compared to pre survey ($Diff = 1.10$, $p = 0.00$, $r = 0.60$) • Experimental group 1 improved significantly in the ex post survey compared to the pre-survey, but with negligible effect size ($Diff = -0.13$, $p = 0.00$, $r = 0.03$)

Table 1. *Cont.*

Authors, Year (Source of Evidence)	Country (Type of Learning Factory)	Intervention (Duration)	Number and Origin of Participants	Type of Measurement (Measurement Time Points)	How Learning/ Competences Assessed?	Findings
Kleppe, Bjelland, Hansen, and Mork, 2022 [43] (Conference paper)	Norway (physical)	Project work (duration not specified)	*N* = 13 (Bachelor's and Master's students in engineering)	Questionnaire for self-evaluation (post)	Questionnaire: Three items on learning/competences, 5-point scale from 1 "strongly disagree" to 5 "strongly agree"	Number of participants (out of 13) who "strongly agree" and "agree" with statements: "Project work in the IdeaLab increased my understanding for connecting design and manufacturing": *n* = 12 "Use of machines and equipment at the IdeaLab have taught me skills I would not have learned in traditional classroom settings": *n* = 13 "Assess to the IdeaLab have made it easier for me to solve cases and assignments given in the course": *n* = 13
Urgo, Terkaj, Mondellini, and Colombo, 2022 [44] (Journal article)	Italy (virtual)	Serious Game (duration not specified)	*N* = 60 (Bachelor's students) [Number of responses per level: level 1: *n* = 60, level 2: *n* = 59, level 3: *n* = 21, as not all participants completed the game]	Questionnaire for self-evaluation (post); grades (post)	Questionnaire: • 1 item on perceived knowledge, 5-point scale from 1 "strongly disagree" to 5 "strongly agree" Grade: From A to F: calculated on the basis of answers to each level as well as an assessment of the students' knowledge and skills by Moodle	Questionnaire: "After taking part to this experience, my knowledge on the topic is better than before.": participants that answered with "agree" or "strongly agree": Level 1: 72%, Level 2: 61%, Level 3: 47% Grades: Level 1: A + B = 63%, C = 20%, D + F = 17% Level 2: A + B = 53%, C = 10%, D + F = 37% Level 3: A + B = 43%, C = 33%, D + F = 24%

Table 2. Characteristics of the 22 studies addressed in the 24 included publications.

Characteristics	Number of Studies
Type of learning factory	
Physical	16
Simulated/virtual/digital	3
Physical and simulated/virtual/digital	3
Type of intervention	
Seminars, project work, lectures	7
Training, workshop, qualification, education	9
Vocational training	1
Game intervention	3
Case study	1
Not specified	1
Type of assessment on learning and competence development	
Self-assessment questionnaire	16
Practical application	5
Knowledge test	4
Interview	3
Observation	2
Grades	2
Videos	1
Written feedback	1
Experimental and control groups	
Experimental group(s) only and no control group	20
Both experimental and control groups	2 (+1)
Number of participants	
Up to 15	7
16 to 30	4
31 to 50	5
51 to 100	3
Over 100	3
Background of participants	
Students	16
Trainees	2
Professionals	6
Stakeholders	1
Not specified	1

Of the 22 studies considered in the review, 16 took place at physical learning factories and three in simulated, virtual, and/or digital learning factories. In three other studies, the learning factories used were both physical and simulated/virtual/digital. The assessed

interventions were measures that fell into the area of advanced training (training, workshop, qualification, education; *n* = 9) and measures that were specifically integrated into higher education (seminars, project work, lectures; *n* = 7) or vocational training (*n* = 1). In three studies, game interventions were explicitly investigated, one study was designed as a case study, and in one study the intervention was not specific. Few studies describe the assessed interventions in detail. It is often not stated how much time of the intervention the participants actively spent at learning factories. In terms of content, the learning factories and interventions considered in the studies address, among other things, the topics of lean, agile, industry 4.0, and product as well as software development.

In terms of the assessment of learning success and competence development (other study topics that may have been explored are not addressed in this review), most studies used self-assessment questionnaires for participants (*n* = 16). In addition, practical applications (*n* = 5), knowledge tests (*n* = 4), interviews (*n* = 3), observations (*n* = 2), grades (*n* = 2), videos (*n* = 1), and written feedback (*n* = 1) were also used for data collection. In 14 studies, only one single data collection instrument was used, which was mostly a self-assessment questionnaire (*n* = 11). Six studies, one study, and one study used two, three and four data collection instruments, respectively.

Of the 22 studies, 20 were designed with experimental group(s) only and no control group, although one study had a group by which the results could be controlled as a result of the data collection procedure. Two studies were designed with both experimental and control groups.

Of the 22 studies, seven, four, five, three, and three studies were conducted with up to 15, 16 to 30, 31 to 50, 51 to 100, and over 100 participants, respectively. Participants were mostly students (*N* = 16). In two studies and six studies, trainees and professionals were participants, respectively. One study also surveyed stakeholders who had not been working at the learning factory, and in one study there was no information on participants' background.

Regarding the statistical analysis of data and the reporting of results, the studies show major qualitative differences. In five studies, quantitative data were analyzed more comprehensively, and significances were reported in terms of differences or correlations. Of these studies, two also report effect sizes on the significances. Beyond that, reporting of quantitative data was purely based on mean and frequency data and percentage changes. In some cases, these data were visualized using graphs without exact numbers stated to them or named in the text.

Only two studies address the extent to which learning success or competence development at a learning factory had an impact on participants' working practices in their professional lives.

3.2. Assessment of Learning Success and Competence Development at Learning Factories

Of the 22 studies identified in the review process, learning success and competence development were assessed in 16 studies on the basis of participants' self-assessments using questionnaires. In nine cases, the questionnaire was answered only after the intervention (post), although in one case a retrospective assessment of the learning/competence status before the intervention was also collected. In one case, the questionnaire was answered only at a time interval after the intervention (ex post). In four studies, the self-assessment was conducted through questionnaires both before (pre) and after (post) the intervention. In one study the survey was conducted both before (pre) and at a time interval after the intervention (ex post), and in another study both after (post) and at a time interval after the intervention (ex post).

In terms of content, participants were asked, among other things, to describe their level of knowledge in general terms (e.g., "After taking part to this experience, my knowledge on the topic is better than before." [44]) or very specifically (e.g., "How confident do you consider yourself in creating pneumatic plans?" [25]), assess their competence development (e.g., "As a result of this course, I am more confident in my ability to solve real-life

problems." [19]), or even assess their transfer performance (e.g., "Were you able to transfer the content and approach to turbulence management into your daily work routine and thus apply it?" [22]). Participants answered the questions on three- to six-point scales. The scope of the used self-assessments questionnaires on learning success and/or competence development varies widely. Six studies report one to five items on this topic [19,21,40,41,43,44]. Five other studies report the use of up to 15 items [20,22–24,39]. One study used 37 items to survey learning success or competence development [25], while two studies do not report on the items or scales used [28,37]. For one study, an existing scale was adapted [34], and in another study, the authors used an existing concept of competence measurement as a guide for item design [30]. In some cases, more items on learning/competences were used than were finally reported in the studies (i.e., [43]). As the aforementioned examples of items already reveal, they are mostly very close in content to the topics covered at the learning factories. Psychometric testing of the items or scales for self-assessment of learning success or competence development was not reported in any of the studies.

Practical applications of the learnings were used in five studies to assess learning and competences. In three of these studies, participants completed practical tasks immediately after the intervention, which were assessed using time measurement and required support [20] or evaluations of solution logs, the latter either individually [39] or in combination with additional competence items [25]. In one of the five studies, the intervention consisted of two hands-on work cycles at the learning factory, using time measurements and quantitative product output to analyze participants' developmental progress between the two cycles [15,16]. In another study, a group of participants applied the intervention content practically at their workplace as a follow-up. Statistical process control charts were used to track a resulting change in production [17].

In three studies, knowledge tests were collected before (pre) and after (post) completion of the intervention. In another study, the knowledge survey was additionally conducted some time after the completion of the intervention (ex post). In all four studies, open-ended questions were used to assess knowledge, which were subsequently evaluated either using scales (e.g., from "not understood" to "well understood" [17,18]) or using scores [25]. Questions included subject-specific exam questions such as "What is the design difference between a retro-reflective sensor and a through-beam sensor?" [42] or application-related competence questions such as "When commissioning a system, you notice that a distance measurement is not working properly. Describe your approach to solving the problem in bullet points" [42].

Interviews were used in three studies, two of which were conducted some time after the intervention (ex post). In the third study, interviews were conducted both following the intervention (post) and some time after the intervention (ex post). In one study, participants were asked to report the extent to which four transversal competencies were formed through their work at the learning factory; the reporting was guided by semi-structured questions [36]. Accordingly, the interviews were geared toward qualitative inquiry. In the other two studies, the interviews were designed as group interviews and included open-ended questions about knowledge and application [15–18]. Example questions are: "Which production system is better in this case, push or pull?" [15,16] or "How did you apply this part of the theory in the activity?" [15,16]. Because at least some of the answers were evaluated for accuracy in terms of content, and as one survey also compared short-term and long-term knowledge gains, both qualitative and quantitative results could be obtained through these interviews.

In two studies, observations of the participants took place during the intervention. In one study, structured observation charts were used for this purpose [17,18]. In a second study, participants' work samples were assessed during the intervention using the observation method of Schaper [30,32].

Two studies used grades to assess learning/competence development: One study with students used the regular exam grade at the end of the semester, as approximately 90% of the exam's exercises had been addressed in the practical training at the learning

factory; thus, there was a very high correspondence in content between the exam and the intervention [30]. In the second study, grades were determined based on answers to knowledge questions as well as an automated calculation of knowledge and skill in the serious game by the Moddle platform used [44].

In one study, group work in the intervention was continuously videotaped, the conversations were then divided into sense units, and these were categorized based on the competences addressed in them. The results were aligned with an existing competence model for the intervention based on the four facets: technical, methodological, personal, and social competence [34].

In one study, participants provided their thoughts in writing after the intervention on their perceived learning success regarding agile principles [38]. The results were qualitatively analyzed, and the reported learning successes were aligned with the intended learning objectives of the training.

3.3. Empirical Findings on Learning Success and Competence Development at Learning Factories

All 22 studies included in the review report an increase in knowledge or competences among the participants as a result of the respective intervention at a learning factory. In all 16 studies, in which self-assessment questionnaires were used, the participants reported increases in knowledge/competences, or assessed these as more positive on average or with high reported frequencies after the intervention. In pure self-assessments of knowledge gains, between 47% and 100% of participants (strongly) agreed that their knowledge or understanding increased in general or on specific topics (e.g., lean tools) [19,21,30,40,41,43,44]. This is also confirmed by studies in which the self-assessment of knowledge and understanding was surveyed before and after the intervention. Here, improved mean values are evident [24,39], and in some cases significant positive changes were also confirmed using *t*-tests [25,30]. In one of the studies, a high effect was determined for this change [25]. In another study, continuous observation during the intervention was used in addition to the self-assessment questionnaire, with the observation data showing a correlation of only 50% with the questionnaire data [30]. In two studies, the assessment of grades occurred in addition to the self-assessment questionnaire: Urgo et al. [44] had grades calculated in their serious game, with results reflecting participants' self-assessment tendencies across the three levels of the game. Glass et al. [30], on the other hand, used exam grades and were able to show through *t*-tests that students who attended at least 75% of the intervention scored higher than students who did not attend the intervention. Furthermore, students who attended more than 75% of the intervention did not score higher on an exercise for which there was no corresponding task in the intervention; participants' reported improvement of their knowledge levels as a result of the intervention is supported by their exam grades [30].

Regarding the knowledge tests which were collected in four studies before and after the interventions, increases in knowledge and/or competences were reported in all cases. In one study, the reported increase in knowledge in absolute difference is between 27.5% and 47.5% [20]. In another study, it is significant for one group but not significant for a second group according to *t*-tests, with these results being validated through observations [17,18]. In a third study, the increase in professional knowledge is significant in all surveyed groups with high effects, with the knowledge items used being controlled for floor and ceiling effects and Cronbach's alpha being reported for the scale [25]. Roll and Ifenthaler [42], on the other hand, report that the interaction of time spent at the learning factory (none, medium, high) and time of survey (pre, post, ex post) is not significant with respect to multidisciplinary digital competences, but significant with respect to subject-specific technical competences, with the values of the control group being mostly lower than those of the two experimental groups. It is interesting to note that both experimental groups had their performance peak directly after the intervention and that the experimental group with medium amount of time at the learning factory improved in the long term (ex post versus pre) with a large effect, while the long-term improvement of the experimental group with most time at the learning factory was also significant but showed a negligible effect size [42].

The study also reported Cronbach's alpha and interrater reliabilities for the knowledge items used.

In terms of practical applications, improvements were found both during the intervention [15,16] and following the intervention. Sieckmann et al. [39] reported that participants scored between 68.0% and 90.3% in practical problem solving; Cachay et al. [20] were able to show that the experimental group needed half the time and less support to solve a practical work task compared to the control group, with the latter undergoing a very similar intervention without a learning factory. In another study, all groups showed significant improvement in practical application according to the *t*-test, although this had a small effect in one group and a large effect in two other groups [25]. In a fourth study, practical application followed the intervention as implementation of the learnings at the workplace. Here, process parameters were used to measure improvements (reduction of fluctuations) due to the implementation, which was significant according to the *t*-test [17,18].

Apart from the study from the publications of Makumbe [17] and Makumbe et al. [18], an implementation of the learnings only took place in the study of Riffelmacher [22]: Here, in an ex-post survey using self-assessment questionnaires, participants indicated on average that they had used learned tools and methods in their professional context, had carried out planning steps, and had transferred and thus applied content and procedures in their everyday professional life.

Group interviews show that participants recalled knowledge content after the intervention, and in one study they stated they were also able to apply it theoretically in the long term [15–18].

In the 22 studies included in the review, a variety of competences were directly addressed in questionnaires, tests, interviews, practical applications, and video studies. The results indicate that participants' competences are increased after the interventions [19,23,28,34,37,40]. With regard to the four competence facets (professional, methodological, social, and self-competence [45–47]), medium to high levels of proficiency or improvement or particular strengthening were reported for the following criteria as a result of working at a learning factory:

- Professional competences: Solving (complex, real-world) problems [19,23,33,39,42,43]; information processing, recognizing interconnections, and analytical thinking [19,33,34]
- Methodological competences: Organizational skills, task management, and time management [23,30]; lean methods [30,38,39]; creativity and innovation [36]
- Social competences: Communication skills [19,34,36]; persuasion with rational arguments [23]; teamwork [19,23,36]; negotiation skills and conflict management [23]; interdisciplinarity [33]
- Self-competences: Lifelong and/or independent learning [19,36]; openness to (new) technologies [34]

Adam et al. [40] also identified a significant relationship between understanding the intervention content and (a) alternating between theory and practice and (b) a do-it-yourself approach. In addition, the results of their study showed a significant relationship between participants' intention to transfer the learnings and (a) understanding the training content, (b) the ease with which the content could be applied in the training, and (c) a higher former experience with the training content [40].

3.4. Summary of the Results

A large number of different data collection instruments are already used in the empirical assessment of learning success and competence development at learning factories. The available studies show a very broad spectrum regarding their evaluation depth and range. While some studies only consider the topics using single items as well as mean values and frequencies (i.e., [37,38,43]), other studies use comprehensive data collection instruments and statistical analysis options (i.e., [17,18,42]).

All 22 empirical studies of learning success and competence development at learning factories identified in the review process produced positive results, suggesting that learning

factories in general are a functional teaching–learning environment. This was to be expected for those studies that did not have a control group for comparison, since it can be assumed that any form of teaching concept and event should contribute to learning and at least impart knowledge, ideally competences. As the review process has shown, only a few studies so far have used control groups to provide insights into the extent to which the use of learning factories is superior to other teaching concepts and events. The results of the studies by Cachay et al. [20], Glass et al. [30], and Roll and Ifenthaler [42] indicate that the evaluated learning factories had a more positive impact on learning and/or competence development of the experimental groups in comparison to the impact of other teaching concepts and events on the respective control groups. At the same time, the study of Roll and Ifenthaler [42] raises the question of whether an increase in the amount of time learners spend at a learning factory also leads linearly to an increase in knowledge and competences or whether this increase is limited. The further results of Adam et al. [40] give a complementary indication that a good balance between theoretical content and practical application at the learning factory can be crucial for the understanding of the intervention content.

4. Discussion

Throughout the review process, it has become apparent that evaluations or feedbacks are frequently mentioned in publications on educational measures at learning factories, as well as the didactic concepts applied there in general. Data were mostly obtained from participants and/or teachers at the end of an intervention and briefly mentioned in the publications in the form of the survey method (primarily questionnaires or interviews). The results of these data collections are summarized in the publications in one or two sentences at most, and conclusions are drawn from them for improvements of the concept. In many cases, however, the actual contents of the assessments, the methods used, the implementation of the evaluation, and specific results remain hidden, so that the findings cannot be verified (cf. i.e., [48–52]) and, as a consequence, the publications could not be included in the review at hand. In one study reviewed in the review process, "typical results" of a learning process were presented without further information on the basis they were generated or selected on [53]. Other publications collected data from 120 [52] to 160 [51] individuals, but did not report further, so there may be great potential for data or at least sampling here, the evaluation or exploitation of which could contribute significantly to the global understanding of learning success and competence development at learning factories.

In summary, the review process has shown that evaluations at learning factories are not uncommon and are frequently conducted in practice. This is certainly due to the fact that evaluations are nowadays a common practice after training and education (cf. [54,55]). Specifically in Germany, evaluations at universities are often mandatory (cf. [56]). However, robust empirical studies on what and how participants in learning factory-related educational measures learn and which competences they develop are rather rare. In the review process, only 24 publications with a total of 22 studies were identified, which dealt with this focus in an empirically comprehensible way.

Many of the studies identified in the review show potential for improvement in survey design, outcome analysis, and/or reporting, although some of these deficits are certainly due to limitations in publication formats (e.g., word count or scope of the paper). If participants' self-assessments were collected exclusively after the intervention, the ability to map learning and competence development through this intervention is limited. The data mainly show whether participants themselves think they have an increase in knowledge/competences as a result of the intervention. Group interviews, on the other hand, provide insight into whether learning success or competence development occurred in general [15–18], but it cannot be reliably captured as to whether knowledge and competence development were possible for all participants of the interview if they were not asked separately. The scales chosen in the questionnaires used were in some cases imbalanced [40]

and sometimes did not appear to be entirely purposeful, e.g., by asking about satisfaction with the level of learning achieved [41] rather than the extent to which it had been achieved. Furthermore, it is notable that some of the data collected were only superficially statistically analyzed, so that the reported results have only limited informational value (cf. i.e., [40]). There are studies in which results were summarized, so that it remains unclear which competences were queried and how they were assessed individually [37], and result reports in which diagrams lack exact values, so that these can only be roughly interpreted [24,28]. While in some publications the interventions are described quite clearly in terms of type and scope (i.e., [15–17,20,25,34,42]), this information is omitted in other publications (i.e., [28,36,37,39]), so that it remains unclear how extensively the participants came into practical contact with a learning factory. Hence, comparing the findings of the publications is difficult, especially in view of the results on the different time span of learning factory use from the study of Roll and Ifenthaler [42].

In the review process, it became evident that the data collection instruments for learning success and competence development at learning factories are very diverse and have not yet been fully exploited. For example, many learning factories store technical data, the analysis of which could be used to complement other data collection methods, such as questionnaires or observational data, especially over a series of experiments and a longer period of time (cf. [57]).

On another note, it should also be kept in mind that the overarching goal of all training measures is to meet trainings needs. Different models on assessing trainings needs can be found in literature (e.g., Tommasi et al. [58], which focusses on industry 4.0 needs). Training needs are the basis to define training objectives which then can be translated into training success parameters and measures. These connections must be considered when deciding on evaluation methods for learning success and competence development at learning factories.

For the present review, potentially relevant scientific databases were researched and used for the literature search. Nevertheless, the possibility remains that other publications exist on studies in the context of learning success and competence development at learning factories which could not be identified based on this search. Reasons for this include the diversity of contexts in which learning factories are used, as well as the multiple disciplines involved in the research and the often-small sample sizes of the studies. Thus, it is evident in practice that results are sometimes only presented at conferences, while conference papers and other contributions such as presentations are not always published in databases.

5. Conclusions

The research community may benefit if evaluations conducted at learning factories in the future are qualitatively improved and comprehensively reported, statistically analyzed, and the results are published in an accessible manner. The review process has shown that there is a lot of data potential in routine or mandatory evaluations that could be analyzed and published, depending on data protection guidelines. However, as these are mostly mere post-intervention questionnaires with self-reports, more comprehensive studies are needed to assess the full potential of learning factories in the context of learning success and competence development.

In this review, it is evident that a variety of assessment instruments are already being used at learning factories to measure learning success and competence development. Nevertheless, few questionnaire studies to date appear to have followed existing scales and psychological methods for question and scale construction. Psychometric testing of the items or scales used have been reported even less frequently. In general, many publications show potential for improvement with regard to study design, implementation, evaluation, and result reports. An increase in study quality should therefore be considered relevant for future research. This could be achieved by basing the studies on defined learning objectives, using established questionnaires, and including professionals in the field of learning and competence assessment into evaluations at learning factories, which would

especially help increase the quality of questionnaires, knowledge tests, and interviews, as well as the statistical analyses. Further, learning factory managers could analyze and interpret changes in data and production parameters (e.g., lead times, produced quantities, scrap rate, energy consumption), especially if participants work at the learning factory over a longer period of time or on a series of comparable experiments. When learning and building competences, participants become more successful in their practical work at the learning factory, which should be measurable using process data relevant to the learning objectives. Again, professionals in the field of learning and competence assessment should be included in this type of assessment to ensure a high quality of the findings.

Knowledge gains are achieved in particular through longitudinal studies, where data are collected both before and after the intervention and, ideally, over the long term. In this context, it is also of great interest to further investigate the application of what has been learned in the work context to ensure a safe transition. Here, the current data situation with two studies [17,18,22] is very limited. Furthermore, more studies with control groups are desirable to determine whether the knowledge and competence developments are a result of an intervention at a learning factory or an intervention in general. Here, the data situation with two studies with few participants [20,42] is also very limited. In general, studies with a higher number of participants are particularly needed to enable comprehensive statistical evaluations and thus reliable results.

The increasing number of studies on the evaluation of learning success and competence development at learning factories, as well as the increasing number of learning factories worldwide, have the potential to contribute to further gains in knowledge.

Author Contributions: Conceptualization, N.R. and S.K.; methodology, N.R. and S.K.; software, N.R. and S.K.; validation, N.R. and S.K.; formal analysis, N.R.; investigation, N.R.; resources, N.R. and S.K.; data curation, N.R.; writing—original draft preparation, N.R.; writing—review and editing, N.R. and S.K.; visualization, N.R.; supervision, S.K.; project administration, N.R. and S.K.; funding acquisition, S.K. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kauffeld, S.; Maier, G.W. Digitalisierte Arbeitswelt. *Gr. Interakt. Organ. Z. Für Angew. Organ.* **2020**, *51*, 1–4. [CrossRef]
2. Kauffeld, S.; Albrecht, A. Kompetenzen Und Ihre Entwicklung in Der Arbeitswelt von Morgen: Branchenunabhängig, Individualisiert, Verbunden, Digitalisiert? *Gr. Interakt. Organ. Z. Für Angew. Organ. (GIO)* **2021**, *52*, 1–6. [CrossRef]
3. Hulla, V.; Herstätter, M.; Moser, P.; Burgsteiner, D.; Ramsauer, H. Competency Models for the Digital Transformation and Digitalization in European SMEs and Implications for Vocational Training in Learning Factories and Makerspaces. *Proc. Eur. Conf. Educ. Res. (ECER)* **2021**, *IV*, 98–107. [CrossRef]
4. Kipper, L.M.; Iepsen, S.; Dal Forno, A.J.; Frozza, R.; Furstenau, L.; Agnes, J.; Cossul, D. Scientific Mapping to Identify Competencies Required by Industry 4.0. *Technol. Soc.* **2021**, *64*, 101454. [CrossRef]
5. Blumberg, V.S.L.; Kauffeld, S. Kompetenzen Und Wege Der Kompetenzentwicklung in Der Industrie 4.0. *Gr. Interakt. Organ. Z. Für Angew. Organ. (GIO)* **2021**, *52*, 203–225. [CrossRef]
6. Abele, E. *CIRP Encyclopedia of Production Engineering*; Chatti, S., Tolio, T., Eds.; Springer: Berlin/Heidelberg, Germany, 2020; ISBN 978-3-642-35950-7. [CrossRef]
7. Abele, E.; Metternich, J.; Tisch, M. *Learning Factories*; Springer International Publishing: Cham, Switzerland, 2019; ISBN 978-3-319-92260-7. [CrossRef]
8. Munn, Z.; Peters, M.D.J.; Stern, C.; Tufanaru, C.; McArthur, A.; Aromataris, E. Systematic Review or Scoping Review? Guidance for Authors When Choosing between a Systematic or Scoping Review Approach. *BMC Med. Res. Methodol.* **2018**, *18*, 143. [CrossRef]

9.   Gusenbauer, M.; Haddaway, N.R. Which Academic Search Systems Are Suitable for Systematic Reviews or Meta-analyses? Evaluating Retrieval Qualities of Google Scholar, PubMed, and 26 Other Resources. *Res. Synth. Methods* **2020**, *11*, 181–217. [CrossRef]

10.   Arksey, H.; O'Malley, L. Scoping Studies: Towards a Methodological Framework. *Int. J. Soc. Res. Methodol. Theory Pract.* **2005**, *8*, 19–32. [CrossRef]

11.   Peters, M.D.J.; Godfrey, C.; McInerney, P.; Munn, Z.; Trico, A.C.; Khalil, H. Chapter 11: Scoping Reviews (2020 Version). In *JBI Manual for Evidence Synthesis*; Aromatis, E., Munn, Z., Eds.; JBI: Adelaide, Australia, 2020. Available online: https://synthesismanual.jbi.global (accessed on 30 September 2022). [CrossRef]

12.   Page, M.J.; Moher, D.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. PRISMA 2020 Explanation and Elaboration: Updated Guidance and Exemplars for Reporting Systematic Reviews. *BMJ* **2021**, *372*, n160. [CrossRef]

13.   Tricco, A.C.; Lillie, E.; Zarin, W.; O'Brien, K.K.; Colquhoun, H.; Levac, D.; Moher, D.; Peters, M.D.J.; Horsley, T.; Weeks, L.; et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Ann. Intern. Med.* **2018**, *169*, 467–473. [CrossRef]

14.   Waffenschmidt, S.; Knelangen, M.; Sieben, W.; Bühn, S.; Pieper, D. Single Screening versus Conventional Double Screening for Study Selection in Systematic Reviews: A Methodological Systematic Review. *BMC Med. Res. Methodol.* **2019**, *19*, 132. [CrossRef] [PubMed]

15.   Ogorodnyk, O. Development of Educational Activity Based on Learning Factory in Order to Enhance Learning Experience. Master's Thesis, Norwegian University of Science and Technology Gjøvik, Gjøvik, Norway, 2016.

16.   Granheim, M.V. Norway's First Learning Factory—A Learning Outcome Case Study. Master's Thesis, Norwegian University of Science and Technology Gjøvik, Gjøvik, Norway, 2016.

17.   Makumbe, S. The Effectiveness of Using Learning Factories to Impart Lean Principles to Develop Business Improvement Skills in Mining Employees. Master's Thesis, University of the Witwatersrand, Johannesburg, South Africa, 2017.

18.   Makumbe, S.; Hattingh, T.; Plint, N.; Esterhuizen, D. Effectiveness of Using Learning Factories to Impart Lean Principles in Mining Employees. *Procedia Manuf.* **2018**, *23*, 69–74. [CrossRef]

19.   Morell de Ramirez, L.; Velez-Arocho, J.I.; Zayas-Castro, J.L.; Torres, M.A. Developing and Assessing Teamwork Skills in a Multi-Disciplinary Course. In Proceedings of the FIE '98. 28th Annual Frontiers in Education Conference. Moving from "Teacher-Centered" to "Learner-Centered" Education. Conference Proceedings (Cat. No.98CH36214), Tempe, AZ, USA, 4–7 November 1998; IEEE: New York, NY, USA, 1998; Volume 1, pp. 432–446. [CrossRef]

20.   Cachay, J.; Wennemer, J.; Abele, E.; Tenberg, R. Study on Action-Oriented Learning with a Learning Factory Approach. *Procedia Soc. Behav. Sci.* **2012**, *55*, 1144–1153. [CrossRef]

21.   Kesavadas, T. V-Learn-Fact: A New Approach for Teaching Manufacturing and Design to Mechanical Engineering Students. In *Volume 5: Education and Globalization*; American Society of Mechanical Engineers: San Diego, CA, USA, 2013; pp. 1–6. [CrossRef]

22.   Riffelmacher, P.P. Konzeption Einer Lernfabrik Für Die Variantenreiche Montage. Doctoral Thesis, Universität Stuttgart, Stuttgart, Germany, 2013.

23.   Ahmad, M.O.; Liukkunen, K.; Markkula, J. Student Perceptions and Attitudes towards the Software Factory as a Learning Environment. In Proceedings of the 2014 IEEE Global Engineering Education Conference (EDUCON), Istanbul, Turkey, 3–5 April 2014; IEEE: New York, NY, USA, 2014; pp. 422–428. [CrossRef]

24.   Plorin, D.; Jentsch, D.; Hopf, H.; Müller, E. Advanced Learning Factory (ALF)—Method, Implementation and Evaluation. *Procedia CIRP* **2015**, *32*, 13–18. [CrossRef]

25.   Zinn, B.; Güzel, E.; Walker, F.; Nickolaus, R.; Sari, D.; Hedrich, M. ServiceLernLab—Ein Lern- Und Transferkonzept Für (Angehende) Servicetechniker Im Maschinen- Und Anlagenbau. *J. Tech. Educ.* **2015**, *3*, 116–149.

26.   Zinke, G.; Schenk, H.; Wasiljew, E. *Berufsfeldanalyse Zu Industriellen Elektroberufen Als Voruntersuchung Zur Bildung Einer Möglichen Berufsgruppe Abschlussbericht. Abschlussbericht. Heft-Nr. 155*; Bundesinstitut für Berufsbildung: Bonn, Germany, 2014.

27.   Abele, S.; Behrendt, S.; Weber, W.; Nickolaus, R. Berufsfachliche Kompetenzen von Kfz-Mechatronikern—Messverfahren, Kompetenzdimensionen Und Erzielte Leistungen (KOKO Kfz). In *Technologiebasierte Kompetenzmessung in der Beruflichen Bildung. Ergebnisse aus der BMBF-Förderinitiative ASCOT*; Oser, F., Landenberger, M., Beck, K., Eds.; WBV: Bielefeld, Germany, 2016.

28.   Henning, M.; Hagedorn-Hansen, D.; von Leipzig, K. Metacognitive Learning: Skills Development through Gamification at the Stellenbosch Learning Factory as a Case Study. *S. Afr. J. Ind. Eng.* **2017**, *28*, 105–112. [CrossRef]

29.   Krathwohl, D.R. A Revision of Bloom's Taxonomy: An Overview. *Theory Pract.* **2002**, *41*, 212–218. [CrossRef]

30.   Glass, R.; Miersch, P.; Metternich, J. Influence of Learning Factories on Students' Success—A Case Study. *Procedia CIRP* **2018**, *78*, 155–160. [CrossRef]

31.   Erler, W.; Gerzer-Sass, A.; Nußhart, C.; Sass, J. Die Kompetenzbilanz. Ein Instrument Zur Selbsteinschätzung Und Beruflichen Entwicklung. In *Handbuch der Kompetenzmessung*; Erpenbeck, J., von Rosenstiel, L., Eds.; Schäffer-Poeschel: Stuttgart, Germany, 2003; pp. 167–179.

32.   Schaper, N. Arbeitsproben Und Situative Fragen Zur Messung Arbeitsplatzbezogener Kompetenzen. In *Handbuch Kompetenzmessung*; Erpenbeck, J., Rosenstiel, L., Grote, S., Sauter, W., Eds.; Schäffer-Poeschel: Stuttgart, Germany, 2017; pp. 523–537.

33.   Balve, P.; Ebert, L. Ex Post Evaluation of a Learning Factory—Competence Development Based on Graduates Feedback. *Procedia Manuf.* **2019**, *31*, 8–13. [CrossRef]

34. Reining, N.; Kauffeld, S.; Herrmann, C. Students' Interactions: Using Video Data as a Mean to Identify Competences Addressed in Learning Factories. *Procedia Manuf.* **2019**, *31*, 1–7. [CrossRef]

35. Richter, T.; Naumann, J.; Groeben, N. The Computer Literacy Inventory (INCOBI): An Instrument for the Assessment of Computer Literacy and Attitudes toward the Computer in University Students of the Humanities and the Social Sciences. *Psychol. Erzieh. Und Unterr.* **2001**, *48*, 1–13.

36. Devika Raj, P.; Venugopal, A.; Thiede, B.; Herrmann, C.; Sangwan, K.S. Development of the Transversal Competencies in Learning Factories. *Procedia Manuf.* **2020**, *45*, 349–354. [CrossRef]

37. Juraschek, M.; Büth, L.; Martin, N.; Pulst, S.; Thiede, S.; Herrmann, C. Event-Based Education and Innovation in Learning Factories—Concept and Evaluation from Hackathon to GameJam. *Procedia Manuf.* **2020**, *45*, 43–48. [CrossRef]

38. Omidvarkarjan, D.; Conrad, J.; Herbst, C.; Klahn, C.; Meboldt, M. Bender—An Educational Game for Teaching Agile Hardware Development. *Procedia Manuf.* **2020**, *45*, 313–318. [CrossRef]

39. Sieckmann, F.; Petrusch, N.; Kohl, H. Effectivity of Learning Factories to Convey Problem Solving Competencies. *Procedia Manuf.* **2020**, *45*, 228–233. [CrossRef]

40. Adam, M.; Hofbauer, M.; Stehling, M. Effectiveness of a Lean Simulation Training: Challenges, Measures and Recommendations. *Prod. Plan. Control* **2021**, *32*, 443–453. [CrossRef]

41. Mahmood, K.; Otto, T.; Kuts, V.; Terkaj, W.; Modoni, G.; Urgo, M.; Colombo, G.; Haidegger, G.; Kovacs, P.; Stahre, J. Advancement in Production Engineering Education through Virtual Learning Factory Toolkit Concept. *Proc. Est. Acad. Sci.* **2021**, *70*, 374. [CrossRef]

42. Roll, M.; Ifenthaler, D. Learning Factories 4.0 in Technical Vocational Schools: Can They Foster Competence Development? *Empir. Res. Vocat. Educ. Train.* **2021**, *13*, 20. [CrossRef]

43. Kleppe, P.S.; Bjelland, O.; Hansen, I.E.; Mork, O.-J. Idea Lab: Bridging Product Design and Automatic Manufacturing in Engineering Education 4.0. In Proceedings of the 2022 IEEE Global Engineering Education Conference (EDUCON), Tunis, Tunisia, 28–31 March 2022; IEEE: New York, NY, USA, 2022; pp. 195–200. [CrossRef]

44. Urgo, M.; Terkaj, W.; Mondellini, M.; Colombo, G. Design of Serious Games in Engineering Education: An Application to the Configuration and Analysis of Manufacturing Systems. *CIRP J. Manuf. Sci. Technol.* **2022**, *36*, 172–184. [CrossRef]

45. Erpenbeck, J.; von Rosenstiel, L.; Grote, S.; Sauter, W. (Eds.) *Handbuch Kompetenzmessung: Erkennen, Verstehen Und Bewerten von Kompetenzen in Der Betrieblichen, Pädagogischen Und Psychologischen Praxis*, 3rd ed.; Schäffer-Poeschel: Stuttgart, Germany, 2017; ISBN 978-3-7910-3511-6.

46. Kauffeld, S. *Kompetenzen Messen, Bewerten, Entwickeln—Ein Prozessanalytischer Ansatz Für Gruppen*; Schäffer-Poeschel: Stuttgart, Germany, 2006; ISBN 3-7910-2508-2/978-3-7910-2508-7.

47. Kauffeld, S. Das Kompetenz-Reflexions-Inventar (KRI)—Konstruktion Und Erste Psychometrische Überprüfung Eines Messin-strumentes. *Gr. Interakt. Organ. Z. Für Angew. Organ. (GIO)* **2021**, *52*, 289–310. [CrossRef]

48. Ugurlu, S.; Gerhard, D. Integrative Product Creation—Results from a New Course in a Learning Factory. In Proceedings of the 16th International Conference on Engineering and Product Design Education: Design Education and Human Technology Relations, E and PDE 2014, Enschede, The Netherlands, 4–5 September 2014; pp. 543–548, ISBN 978-1-904670-56-8.

49. Andersen, A.-L.; Brunoe, T.D.; Nielsen, K. Engineering Education in Changeable and Reconfigurable Manufacturing: Using Problem-Based Learning in a Learning Factory Environment. *Procedia CIRP* **2019**, *81*, 7–12. [CrossRef]

50. Balve, P.; Albert, M. Project-Based Learning in Production Engineering at the Heilbronn Learning Factory. *Procedia CIRP* **2015**, *32*, 104–108. [CrossRef]

51. Ogorodnyk, O.; Granheim, M.; Holtskog, H.; Ogorodnyk, I. Roller Skis Assembly Line Learning Factory—Development and Learning Outcomes. *Procedia Manuf.* **2017**, *9*, 121–126. [CrossRef]

52. Grøn, H.G.; Lindgren, K.; Nielsen, I.H. Presenting the UCN Industrial Playground for Teaching and Researching Industry 4.0. *Procedia Manuf.* **2020**, *45*, 196–201. [CrossRef]

53. Gento, A.M.; Pimentel, C.; Pascual, J.A. Lean School: An Example of Industry-University Collaboration. *Prod. Plan. Control* **2021**, *32*, 473–488. [CrossRef]

54. Kauffeld, S.; Zorn, V. Evaluationen Nutzen—Ergebnis-, Prozess- Und Entwicklungsbezogen. *Handb. Qual. Stud. Lehre Und Forsch.* **2019**, *67*, 37–62.

55. Kauffeld, S. *Nachhaltige Personalentwicklung Und Weiterbildung*; Springer: Berlin/Heidelberg, Germany, 2016; ISBN 978-3-662-48129-5. [CrossRef]

56. Section 5 of the Lower Saxony (Germany) Higher Education Act [Niedersächsisches Hochschulgesetz (NHG)] on Evaluation of Research and Teaching. 2007. Available online: https://www.nds-voris.de/jportal/?quelle=jlink&query=HSchulG+ND+%C2%A7+5&psml=bsvorisprod.psml&max=true (accessed on 30 September 2022).

57. Teizer, J.; Chronopoulos, C. Learning Factory for Construction to Provide Future Engineering Skills beyond Technical Education and Training. In Proceedings of the Construction Research Congress 2022, Arlington, VA, USA, 9–12 March 2022; American Society of Civil Engineers: Reston, VA, USA, 2022; pp. 224–233. [CrossRef]

58. Tommasi, F.; Perini, M.; Sartori, R. Multilevel Comprehension for Labor Market Inclusion: A Qualitative Study on Experts' Perspectives on Industry 4.0 Competences. *Educ. Train.* **2022**, *64*, 177–189. [CrossRef]