1 Overview

Abstract

The "Learning Quality Inventory for In-Company Training in VET" (VET-LQI) is a survey instrument aiming for a valid and comprehensive measurement of vocational training quality in companies. The validated instrument comprises 22 short scales with 76 items, enabling a broad approach to the construct of training quality. The survey instrument was developed by conducting a qualitative meta-synthesis of all existing instruments regarding training quality. 43 test instruments were integrated into the general theoretical framework on workplace learning by Tynjälä (2013). Item and confirmatory factor analyses (CFA) indicate satisfactory reflection of all common workplace characteristics regarding validity and reliability*. Overall, the VET-LQI overcomes the limited applicability of prior instruments by covering all identified content areas of training quality from the literature.

*Except for 'Relevance of Tasks' (Scale 7), which is still included to maintain a broad coverage of the construct vocational training quality.

Keywords

Title: Learning Quality Inventory for In-Company Training in VET (VET-LQI)

Authors: Deutscher, Böhn & Krötz

In ZIS since: 2022

Number of Scales: 22

Number of Items: 76

Survey Mode: PASI

Processing Time: Max. 45 Minutes. Most trainees finish within 30 Minutes.

Reliability: Cronbach's alpha: .68 to .89

Validity: Evidence for structural validity is achieved by CFA; convergent and discriminant validity is satisfactory for most scales; construct validity is given by a qualitative meta-synthesis of existent (validated) instruments.

Construct: In-company training quality

Catchwords: Vocational Training, In-Company Training, Training Quality, Apprentice Questionnaire, VET, Training Quality Survey

Language Documentation: English

Language Items: English, German

Item(s) used in Representative Survey: yes

URL Website: https://link.springer.com/article/10.1007/s12186-020-09251-3#Sec16

URL Data Archive: -

Status of Development: validated

Cachet: -

2 Instrument

Instruction

The 22 scales constitute the survey instrument VET-LQI, which can be used to examine vocational training quality in companies. Before starting the survey, the following instructions were given to trainees:

Dear trainees,

With this questionnaire we would like to find out how you perceive your training. We are particularly interested in your opinion on the company facet of vocational training and less in content related to vocational school.

In the following, only your personal impression counts. There are no right or wrong answers. Please answer honestly – the goal of our research is to improve in-company training. You can contribute to this. There are response options for each question. Please answer the questions by ticking or filling in the appropriate box.

We assure you that you will remain completely anonymous and that the information you provide will only be used to explore training quality. Neither your vocational school nor your training company will receive information about your details at the personal or class level. In turn, we do not receive any information about the training company you belong to.

Completing the questionnaire will take approximately 45 minutes.

Please enter the exact name of your apprenticeship in the following: _____

Thank you for your participation!

Items

Note: In the following, all 22 scales that were part of the original validation of the survey are listed in the required form to enable a transparent review of the validation process. The corresponding tables 1 to 22 comprise all 99 items that were initially tested via item analysis and confirmatory factor analysis.[1] Brackets around the item number and * behind the item indicate that the item was removed during the validation process due to insufficient results and is not considered as being part of the final survey instrument. The list of final items and scales is presented later in table 24.

---

[1] Items 1 to 20 were aspects regarding personal details and were not part of the validation. Also scales on the company or vocational framework were not part of the validation (see section *Item analyses*). Researchers can adapt such framework and personal information as required in their specific research. These excluded items on personal details and framework characteristics can be read in the appendix of Böhn (2020, p. 198 ff.).

Table 1

Scale 1: Items of the Scale Work Climate

No.	Item	Polarity	Subscale
[21]	If necessary, the employees in my company support each other.*	+	
22	There is a personal atmosphere within my company.	+	
23	There is a bad working atmosphere within my company	-	
24	There is strong competition between employees in my company.	-	
25	Employees in my company are rigorously monitored and controlled.	-	

Table 2

Scale 2: Items of the Scale In-Company Learning

No.	Item	Polarity	Subscale
26	Workplace learning in my company is characterized by different teaching methods.	+	
27	Workplace learning in my company is characterized by the usage of different materials and media.	+	
[28]	I have to teach myself concerning the knowledge I need for working tasks in my company.*	-	
[29]	Working tasks in my company cover all aspects of successfully managing them from start to finish.*	+	

Table 3

Scale 3: Items of the Scale Learning Venue Cooperation

No.	Item	Polarity	Subscale
30	What I learn at vocational school is important for the daily work in my company.	+	
31	When managing work tasks in the company, I benefit from knowledge I accumulated during vocational school sessions.	+	
[32]	When managing work tasks in my company, I benefit from knowledge I accumulated during in-company sessions.*	+	
33	The in-company vocational training and the vocational school are well coordinated.	+	

Table 4

Scale 4: Items of the Scale Overload

No.	Item	Polarity	Subscale
[44]	In my company I fear making mistakes at work.*	+	
45	In my company I feel under pressure of time at work.	+	
[46]	In my company I quarrel with supervisors, training personnel, colleagues, or other apprentices.*	+	
[47]	Work tasks in my company are physically demanding.*	+	
48	In my company others interfere my work.	+	
49	I have problems 'recharging batteries' in my spare time after work.	+	
50	Because of the daily demands in my company, I feel totally exhausted, tired, and drained	+	
51	I often think ‚I can't go on any longer'.	+	

Note. Items can be inverted to represent "high" training quality in the case of low overload

Table 5

Scale 5: Items of the Scale Variety of Tasks

No.	Item	Polarity	Subscale
52	In my company I deal with a variety of work tasks.	+	
53	In my company I work on new tasks every now and then.	+	
54	In my company work tasks are highly diversified.	+	

Table 6

Scale 6: Items of the Scale Autonomy

No.	Item	Polarity	Subscale
[55]	In my company I am flexible in the timing of work tasks.*	+	
56	In my company I am often able to make decisions myself.	+	
57	In my company I am able to decide on my own what means to take to reach a goal.	+	
58	In my company I am given an enormous amount of freedom in doing my job.	+	

Table 7

Scale 7: Items of the Scale Relevance of Tasks

No.	Item	Polarity	Subscale
[59]	In my company I am confronted with responsible tasks.*	+	
60	In my company I work on ‚real tasks'.	+	

61	In my company I have to deal with several tasks that are not part of my vocational training program (e.g., make coffee, copying, etc.).	-

Table 8

Scale 8: Items of the Scale Complexity of Tasks

No.	Item	Polarity	Subscale
[62]	In my company work tasks are difficult and complicated.*	+	
63	In my company work tasks are characterized by considering a wide range of information.	+	
64	In my company work tasks are characterized by considering a wide range of objectives.	+	
65	In my company work tasks are characterized by considering changes over time.	+	

Table 9

Scale 9: Items of the Scale Training Requirements and Ability Level

No.	Item	Polarity	Subscale
[66]	In my company I am confronted with tasks that fit my ability level.*	-	
67	In my company I am confronted with tasks that are too complicated.	+	
68	In my company I am confronted with tasks I am insufficiently trained and prepared for.	+	
[69]	In my company I am confronted with tasks that demand too little from me.*	-	
[70]	In my company I am confronted with tasks that are challenging for me.*	+	
71	In my company I am confronted with easy tasks.	-	

Note. Items can be inverted to represent "high" training quality in the case of fitting (rather easy) tasks.

Table 10

Scale 10: Items of the Scale Involvement in Occupational Expert Culture

No.	Item	Polarity	Subscale
72	I am involved in the improvement of work processes in my company.	+	
73	My ideas and proposals are considered in my company.	+	
74	I am involved in the discussion of technical and professional issues in my company.	+	

Table 11

Scale 11: Items of the Scale Functional Involvement

No.	Item	Polarity	Subscale
75	Continuous collaboration is part of the daily work in my company.	+	
76	My work tasks are characterized by a close coordination with other employees in my company.	+	
77	I am involved in all work tasks in my department.	+	
78	Basically, my work tasks play a crucial role for my department.	+	
79	I am well integrated into the operational working procedures.	+	

Table 12

Scale 12: Items of the Scale Social Involvement

No.	Item	Polarity	Subscale
80	Employees in my company are interested in me.	+	
81	Employees in my company are interested in my private wellbeing.	+	
[82]	In my company there is nobody I get in touch with in private.*	-	
83	Employees in my company seem disturbed by my presence.	-	
84	Employees in my company ignore me.	-	

Table 13

Scale 13: Items of the Scale Mentoring

No.	Item	Polarity	Subscale
85	In my company nobody feels responsible for me.	-	
86	In my company I am completely left alone to myself.	-	
[87]	When I perform a task in my company a colleague is present.*	+	
88	When I ask training personnel or colleagues for help they immediately support me.	+	

Table 14

Scale 14: Items of the Scale Curriculum Orientation

No.	Item	Polarity	Subscale
89	I do know my in-company training plan.	+	
90	The arrangements of my in-company training plan are observed.	+	
91	My in-company training program is implemented without a formal training plan.	-	

Table 15

Scale 15: Items of the Scale Feedback

No.	Item	Polarity	Subscale
92	In my company good performances are not praised.	-	
93	Normally I do know whether I perform work tasks satisfactorily or not.	+	
94	I find it hard to figure out whether I perform work tasks satisfactorily or not.	-	
95	The training personnel and my colleagues let me know whether I perform work tasks satisfactorily or not.	+	

Table 16

Scale 16: Items of the Scale Personnel and Instructions

No.	Item	Polarity	Subscale
[96]	Those who train me on the job give explicit work instructions.*	+	
97	Those who train me on the job are able to answer difficult technical questions.	+	
98	Those who train me on the job can explain well.	+	
[99]	There is a lot I can learn from those who train me on the job.*	+	
100	Those who train me on the job show that they enjoy their work.	+	
101	Those who train me on the job are technically competent.	+	

Table 17

Scale 17: Items of the Scale Premature Termination of Contract

No.	Item	Polarity	Subscale
105	I often think about terminating my training program prematurely.	+	
106	I think about terminating my training program prematurely because of operational reasons (e.g., working hours, quarrels with colleagues, etc.).	+	
107	I think about terminating my training program prematurely because of occupational reasons (e.g., dissatisfaction with tasks, lack of interest in field of activity etc.)	+	
108	I think about terminating my training program prematurely because of vocational school reasons (e.g., quarrels with teachers, dissatisfaction with learning content, etc.).	+	
109	I think about terminating my training program prematurely because of personal reasons (e.g., family reasons, health reasons, financial reasons, etc.).	+	
110	I already search for an alternative vocational training.	+	
111	I already search for an alternative training company.	+	

Note. Items can be inverted to represent "high" training quality in the case of low drop-out intention.

Table 18

Scale 18: Items of the Scale Professional Competence

No.	Item	Polarity	Subscale
112	I do have a high level of expertise.	+	
113	I do have a high level of social skills (e.g., strong communication skills, team spirit, etc.).	+	
114	I am able to perform most of the tasks in my company independently and without help.	+	
115	I feel well prepared to cope with future challenges in my job.	+	
116	At the end of my training program, I will be able to perform every job-related task in my domain.	+	

Table 19

Scale 19: Items of the Scale Overall Assessment and Satisfaction

No.	Item	Polarity	Subscale
[124]	I am strongly satisfied with the training program in my vocational school*	+	
125	I am strongly satisfied with the training program in my company.	+	
126	I am strongly satisfied with my vocational training in general.	+	

Table 20

Scale 20: Items of the Scale Vocational Identity

No.	Item	Polarity	Subscale
127	My skilled occupation and me match very well.	+	
[128]	I am proud of my skilled occupation.*	+	
[129]	My skilled occupation is part of my personality.*	+	
130	I totally assimilate to my skilled occupation.	+	
131	I like telling others about my skilled occupation.	+	

Table 21

Scale 21: Items of the Scale Operational Identity

No.	Item	Polarity	Subscale
132	My training company and me match very well.	+	
[133]	I am proud to work for my training company.*	+	
134	My training company is a place like home for me.	+	
[135]	I feel comfortable in my training company.*	+	
136	I like telling others about my training company.	+	

Table 22

Scale 22: Items of the Scale Future Prospects and Career Aspirations

No.	Item	Polarity	Subscale
137	After finishing my vocational training I would like to work in that occupation.	+	
138	After finishing my vocational training I would like to work in my training company.	+	
[139]	After finishing my vocational training I would like to study or get further training.*	-	

Response specifications

1 'strongly disagree'

2 'mostly disagree'

3 'slightly disagree'

4 'partly agree'

5 'slightly agree'

6 'mostly agree'

7 'completely agree'

Alternatively: 'I do not want to or cannot answer this'

Note: Usage also possible via a five-level Likert scale.

Scoring

Table 1 to table 22 each represent one quality criteria from the survey instrument, which can be used by simple mean scores across scales. As indicated above, the response options were scored in ascending order from 1 (strong disagreement) to 7 (complete agreement). The option 'I do not want to or cannot answer this' could be scored by 999 or other values indicating missingness, as it does not become part of the analysis. After data is collected, negatively polarized items (as indicated in tables 1 to 22) are reversed to match their scales' directions. By this, higher values then represent higher training quality for all scales, except for '*Overload*' (higher values indicating higher workload), '*Training Requirements and Ability Level*' (higher values indicating too high requirements), and '*Premature Termination of Contract*' (higher values indicating drop-out intentions).[2]

Application field

The purpose of the VET-LQI is the comprehensive measurement of vocational training quality in companies, which is done across a broad range of quality criteria relevant at the workplace. Initially (as presented here) the survey mode was paper-and-pencil in the form of PASI (paper-and-pencil self-administered interviewing). However, the instrument is easily reproducible as an online survey (CASI: computer-assisted self-administered interviewing), which was already tested in later studies as part of a longitudinal DFG project (see e.g., Krötz & Deutscher, 2022; Ma et al., in preparation).

---

[2] However, if the researchers want to do so, the latter 3 scales can be inverted to correspond to the logic of the majority of quality scales, where higher values indicate higher training quality.

The target population is trainees in the commercial and business management domain. The scales can partly be applied to in-company training in other domains as well, for example technical-industrial or home economics professions. The instrument asks for the individual's subjective perception of the respective training quality criterion, therefore also individual-level analysis is possible. There are no specific professional qualifications required for using the survey instrument.

3 Theory

The following descriptions of theory and the validation process in sections 3 to 5 are largely taken from the published paper by Böhn and Deutscher (2021).

With regard to vocational in-company training, training quality can be defined as the "subjective perception of characteristics of vocational training that are conducive to certain outcomes" (Klotz et al., 2017). This view relies of a widely accepted approach, the 3-p model of workplace learning by Tynjälä (2013), describing learning processes on three dimensions: presage, process, product (e.g., Biggs, 1999; Seyfried et al., 2000; Visser, 1994). Based of Biggs (1999), Tynjälä (2013) emphasises the individual's perspective and considers all learner factors and aspects of learning context within the presage (input) dimension. The process dimension includes characteristics of workplace learning such as the structure and performance of work tasks and the individual's interaction with others. The product (output) dimension covers all learning related outcomes with a focus on the individual's personal and professional development (Tynjälä, 2013, p. 14).

Besides this broad consensus on the processual model, the ideas about specific characteristics in training reality that define these dimensions vary widely. There is a clear lack of conceptual clarity regarding vocational training quality, which is ultimately reflected in the vast amount of test instruments developed in this area that have partly little intersection. Our research objective was to organize all existing test instruments in the VET context by using a qualitative meta-synthesis, to ultimately develop a survey instrument that comprises all relevant characteristics for workplace learning in the context of VET, excluding learning in vocational schools. The survey items identified were then organized and grouped into the model of workplace learning by Tynjäjä (2013) as a theoretical framework. Figure 1 shows the resulting categories (new items were only created when necessary, see section "Item generation and selection" below).
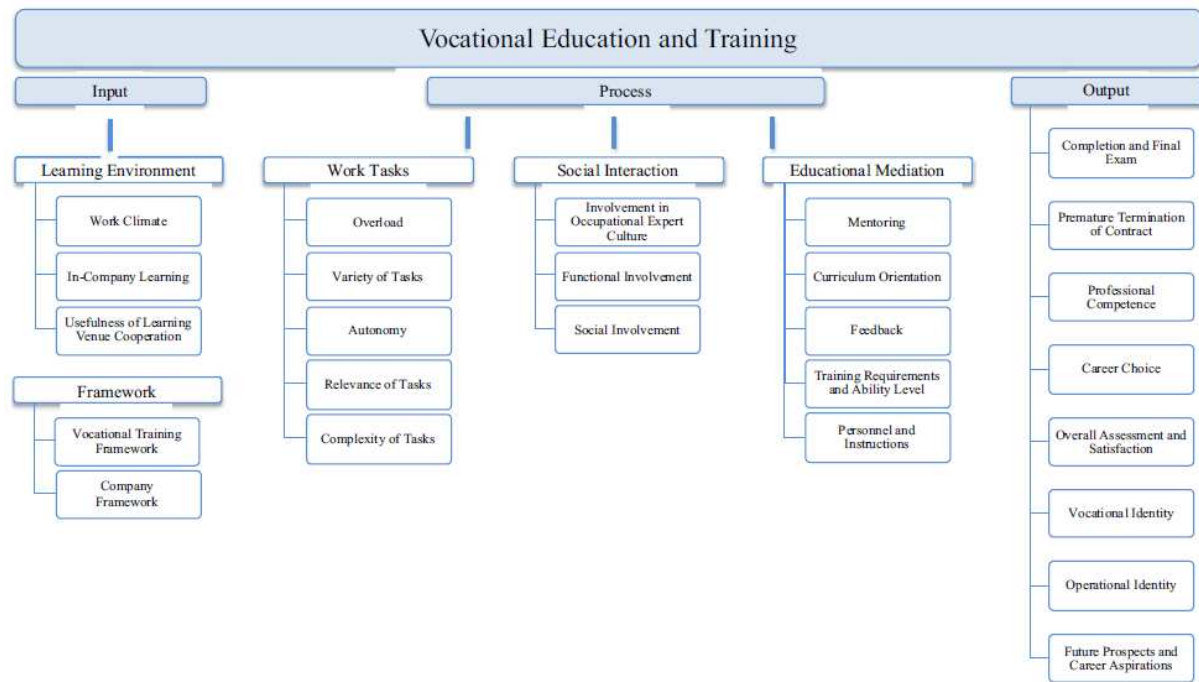
Fig. 1. *Integrative category system based on a qualitative meta-synthesis of test instruments in the context of dual VET (Böhn & Deutscher, 2021, p. 30)*

As the German VET system is special with regard to its duality, consisting of state schooling and private training companies, large amount of research and test instruments focus on related aspects such as learning venue cooperation (e.g., Brooker & Butler, 1997; Dwyer et al., 1999; Ebner, 1997; Feller, 1995; Fink, 2015; Heinemann et al., 2009; Keck et al., 1997; Nickolaus et al., 2015; Prenzel et al., 1996; Ulrich & Tuschke, 1995; Virtanen et al., 2014; Walker et al., 2012). Since the goal of vocational training is the attainment of a formal occupational qualification, trainees' performance needs to be assessed throughout the process. For this reason, work tasks characteristics and the training personnel are of particular relevance for in-company training (Brooker & Butler, 1997; Fuller & Unwin, 2003; Virtanen et al., 2009). Therefore, existing surveys strongly focus on questions related to those aspects (e.g., Beicht et al., 2009; Brooker & Butler, 1997; DGB, 2008; Ebner, 1997; Ernst, 2016; Feller, 1995; Gebhardt et al., 2014; Heinemann et al., 2009; Hofmann et al., 2014; Keck et al., 1997; Koch, 2016; Nickolaus et al., 2015; Prenzel et al., 1996; Velten & Schnitzler, 2012; Virtanen et al., 2014; Zimmermann et al., 1994). We focused on test instruments in the VET context, however, there is contentual overlap to general workplace learning instruments.

4 Scale development

Item generation and selection

A qualitative meta-synthesis was conducted to collect and categorize all existing survey instruments on training and learning quality in the area of VET. Thereby, the instrument integrates the full range of research findings on relevant quality criteria and item formulations (for detailed information on literature research, selection, and categorization see the full papers: Böhn & Deutscher, 2019; 2021). Mayring's (2004) qualitative content analysis (including generalization, selection, and bundling) was conducted to inductively determinate categories for the 3356 items, excluding school-related items. Items with similar contents were grouped (e.g., the items 'Have you previously held a full-time job?' [NCVER, 2000] or 'Do you have previous work experience?' [Virtanen & Tynjälä, 2008] were grouped under the keyword 'personal background'). Subsequently, keywords were merged to categories to achieve a reasonable summary of contents, which was recorded in a codebook. The procedure was repeated by the second researcher, each together with a third person (intracoder reliability = .984, intercoder reliability = .926).

Overall, 43 test instruments (oldest from 1994) with 3356 items were classified into 30 categories. Focusing on workplace characteristics, 2343 items (and their categories) could be arranged within the framework model by Tynjälä (2013). User frequencies of items were analysed. The test instruments showed immense nominal and substantive heterogeneity regarding their items. Within the corresponding studies, validation results were not reported in all cases but indicated satisfactory to excellent levels of internal consistency or model fit.

New scales were constructed by relying on frequently used items and some newly developed items. When existing scales and items showed little validation or occupation-specific formulations, they were adapted to a standard format or new items were created. Overall, 82 items could be largely accepted. All other items were reformulated in an occupation-independent manner and designed based on the underlying category codebook. Additionally, the phrases 'in my company' or 'in my department' were added when necessary to clarify the company focus. The initial German version was translated to English by two researchers and checked by a native English speaker and is thus available in two languages.

A first German version of the VET-LQI, comprising 166 items, was pretested anonymously in three vocational schools in Germany (Baden-Wuerttemberg) in 2017. The sample consisted of 393 trainees in 15 different occupations from the commercial and business management domain. Afterwards, items that caused difficulties in understanding or that did not improve the scale reliability while being dispensable regarding content validity were removed.

The main validation of the German VET-LQI, with 139 items and 31 scales (including 5 person-related and 2 framework scales), then took place anonymously in three other vocational schools in 2018, with 428 trainees from 7 occupations within the commercial and business management domain. For the documentation of the item selection (difficulties etc.), see section "Item Analysis" below and table 24, which is based on confirmatory factor analysis (CFA) due to the theoretical substantiation of the factor structure.

Samples

The sample was recruited via three German vocational schools (PASI mode), where complete school classes of training courses were asked to participate. The sample consisted of 428 apprentices, aged between 16 and 37 (*M* = 20.5). 54.6% of the participants were female, 45.4% were male. Half of the sample was in their first year of training, while 30% were in their second year and 20% in their third year of training. Apprentices were trained in 7 commercial occupations e.g., Retail Salesman (33%), Industrial Management Assistant (24%) and Office Management Assistant (20%). Compared to the statistical population of the analyzed occupations (using Data by the Institute for Vocational Education and Training), t-test and chi²-tests showed no significant differences with respect to age (p = .426), gender (p = .586) and education (p = .100). 41% possessed a higher school certificate, nearly 50% owned at least a secondary school certificate (see Böhn & Deutscher, 2021).

The number of missing values were low and stayed < 5% for the vast majority of items, except for five items (items 26 and 64: 6%; items 91 and 129: 8%; item 90: 9%). Initially, the items were maintained within the following analyses since the missingness was not excessive and important contents of the respective scales would be missing, if the items were deleted (in case of items 90 and 91, the whole scale 14 would no longer exist).

Item analyses

As stated above, the version of the VET-LQI that resulted from the pretesting comprises 31 scales, including 5 person-related and 2 framework scales ('*Vocational Training Framework*', '*Company Framework*'). These person-related and framework scales, as well as 2 additional output scales ('*Completion and Final Exam*', '*Career Choice*') were not part of the following validation process (and thus remain unmentioned in tables 1–22) as they rather represent formative than reflexive theoretical concepts and do not meet the requirements of factor analyses (Bollen & Lennox, 1991; Diamantopoulos & Winklhofer, 2001). Furthermore, the content of these scales can be adjusted and supplemented individually, depending on the context of the survey usage.

As a first step, the internal reliability of the remaining 22 scales (99 items; see tables 1–22) and their item discrimination power were assessed. 15 items could be removed due to low scale correlation (discrimination power < .3; Ebel & Frisbie, 1986). 9 of these items (items 32, 44, 47, 55, 62, 82, 96, 128, 135) also showed missing value ratio above 5%, while 6 items (items 28, 29, 66, 69, 87, 139) improved the respective scale reliabilities.

The resulting scale reliabilities were good or excellent for 19 scales (Cronbach's alpha > .7), while 2 scales '*Training Requirements and Ability Level*' (.684) and '*Overall Assessment and Satisfaction*' (.657) indicated at least acceptable values (Cronbach's alpha > .6; DeVellis, 1991; Nunnally, 1978). The latter scale could be improved by deleting item 124 but in order to keep school-related satisfaction within the instrument, the item was maintained for the factor analysis. Only the scale '*Relevance of Tasks*' showed problematic consistency of .447 (Cronbach's alpha) as it included two items (out of three) with discrimination power < .3 (items 59 and 61).

Additionally, for content-related reasons four items (33, 52, 88, 124) that otherwise improved the scale consistency if deleted and five items above the 5% threshold of missing values (items 26, 64,

90, 91, 129) were initially kept for the next step of factor analysis. As indicated above, in cases of such minor "weaknesses", decisions to initially keep the items were made, in line with Schermelleh-Engel and Werner (2012), in order to maintain the broad approach of the VET-LQI in terms of content validity.

CFA was conducted for the 22 scales (and 84 remaining items) using the free statistics software R and lavaan package, applying the common principal axis extraction and varimax rotation method. As the model structure was based on the extensive meta-synthesis of training quality instruments, the number of factors and the related indicators were already theoretically substantiated. Additionally, three exploratory factor analysis (initial starting point with 99 items, model 1 [84 items] and model 5 [76 items]) were conducted to enable a comparison of how the items could be structured as factors on a mere statistical basis versus our theory-driven approach. In the beginning (99 items), an EFA with principal-axis extraction suggest 25 factors (eigenvalues > 1), when we theoretically substantiated 22 factors. Then, for model 1, the EFA results suggest 17 factors, whereas the final model 5 would suggest 16 factors only. However, as we substantiated our factor solution theoretically by an extensive literature research and synthesis of 43 instruments, we propose our 22-factor model in order to consider more specific quality criteria of in-company training (the EFA factor matrix can be obtained upon request from the authors).

In the following, our proposed model was tested in a CFA with 22 factors. The model fit indices as well as item characteristics (uniqueness, factor loadings, discrimination power) were checked and weak items were deleted, leading to more parsimonious models, as presented in the following.[3] MLM was used as a more robust version of Maximum Likelihood (e.g., Curran et al., 1996).[4]

The local goodness-of-fit indices indicated possible improvements for the initial 22-factor model (table 23). Regarding the factor loadings, loadings should be high and unambiguous directed on one factor.[5] Indicators with low communalities or high values for uniqueness were treated with caution for values > .6 (uniqueness). Lastly, all indicators were significantly related to their factor and there were no high correlations (> .8) between variables, therefore multicollinearity was no issue.

Following the presented thresholds, two items (59, 124) were eliminated from the base model for high uniqueness (> .6), low factor loadings (< .4) and low discriminatory power (< .3).[6] The critical item 61 was kept as it was part of the same scale as item 59, thus only the worst item should be deleted to test the results in the next model (table 23, model 3).

For model 4 (table 23), the items 70 (discrimination power: .291) and 133 were removed due to critical values (correlation with item 132 >.8 and high loadings on two factors). For the final model (table 23, model 5) four items were eliminated: Items 21 and 46 showed high loadings on more than one factor,

---

[3] The assumption and modelling of reflective scales results in highly correlated indicators (items) of one latent construct (factor/scale). For the 22 factors, 84 factor loadings, 22 covariances between the factors, as well as the residual variances of the indicators and variances of the factors (106) had to be estimated. The factor loading of the one indicator with the highest loading per scale was set to one, thus 190 free parameters had to be estimated.

[4] Based on the Satorra-Bentler scaled Chi² statistic, MLM is appropriate especially when indicator variables deviate from normality (e.g., Boomsma & Hoogland, 2001; Schermelleh-Engel et al., 2003).

[5] A threshold for factors loadings was set to a minimum of .4. Cross-loadings were deemed problematic when factor loadings differed less than .2 (Matsunaga, 2010).

[6] Item 59 (uniqueness: .71, factor loading: .372, discrimination power: .289), item 61 (uniqueness: .70, factor loading: .404, discrimination power: .217), item 124 (uniqueness: .39, factor loading: .174, discrimination power: .234).

while item 99 correlated highly with item 98. Item 129 could be excluded for its amount of refused answers (above 5%) without harming the scale reliability.

Table 23.
Global fit indices for CFA.

Estimator: MLM	Exact fit index		Approximate fit index	Residual based fit index	Comparative fit indices	
Model	Chi²/df ratio	p-value (Chi²)	RMSEA	SRMR	CFI	TLI
1. base model: 22 factors with 84 items	4,885.632/3,171 = 1.541	< .001	0.044	0.080	0.844	0.828
2. one factor solution* (84 items)	9,344.624/3,485 = 2.681	< .001	0.078	0.096	0.476	0.463
3. 22 factors with 82 items	4,537.283/3,008 = 1.508	< .001	0.043	0.079	0.859	0.845
4. 22 factors with 80 items	4,293.427/2,849 = 1.507	< .001	0.043	0.074	0.861	0.846
5. 22 factors with 76 items	3,887.370/2,543 = 1.529	< .001	0.044	0.071	0.864	0.847

Models allow covariances between factors. *Correlations between factors set to one.

Comparing the models' fit indices (table 23), the Chi²/df ratio indicated an acceptable fit for model 1 and the following more parsimonious variations (< 2; Byrne 1991). Not surprisingly, the p-value stayed significant for all models due to their complexity. As CFI and TLI deteriorate with model complexity, the indices did not reach an acceptable fit (Hu & Bentler, 1999) but indicated an improvement from model 1 to 5 due to the item exclusions. RMSEA and SRMR indicated a good fit for model 1 and improved further until model 5 (Browne & Cudeck, 1993; Hu & Bentler, 1995). Therefore, model 5 generally shows the best fit indices while being the most parsimonious (see also fig. 2).

Item parameters
Regarding the final item parameters, please see fig. 2 above for item factor loadings and table 24 below for item communalities, means, standard deviations, and discrimination power.
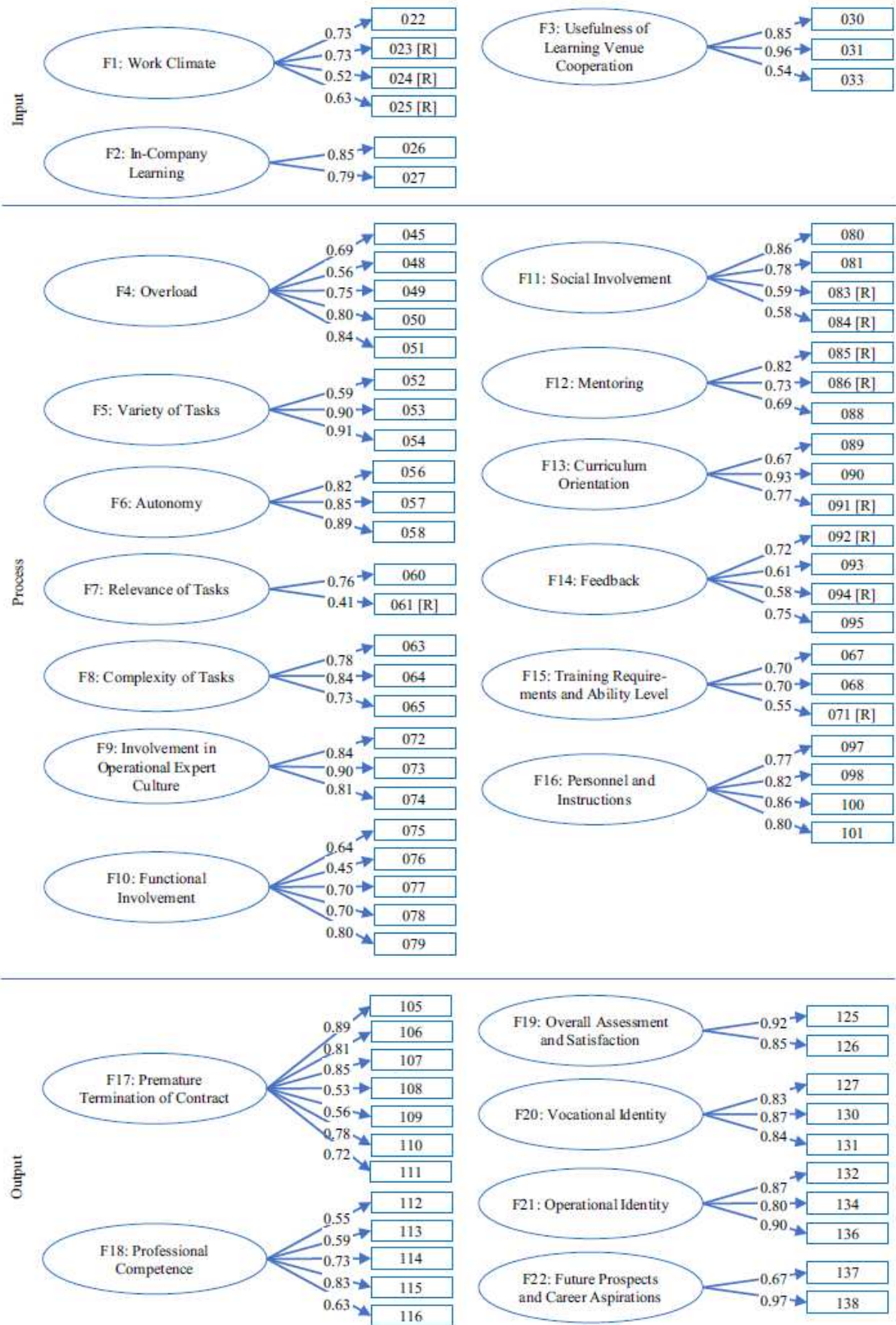
Fig. 2. *Factor loadings based on final model 5 (taken from original paper by Böhn and Deutscher [2021, p. 42])*

Table 24

Final Scale Reliabilities and characteristics of the final Manifest Items

Scale	Cronbach's Alpha	McD's Omega	Item	Communal-ity	Mean	SD	Selectiv-ity
Scale 1	.748	.761	Item 22	.57	5.71	1.23	.611
Work Climate			Item 23	.59	5.33	1.67	.654
			Item 24	.49	5.52	1.38	.520
			Item 25	.56	4.90	1.54	.500
Scale 2	.787	.788	Item 26	.71	3.96	1.63	.534
In-Company Learning			Item 27	.57	4.24	1.63	.469
Scale 3	.812	.825	Item 30	.79	3.75	1.54	.668
Learning Venue			Item 31	.81	3.94	1.44	.758
Cooperation			Item 33	.47	3.40	1.59	.524
Scale 4	.844	.848	Item 45	.53	3.19	1.54	.678
Overload			Item 48	.48	2.59	1.44	.536
			Item 49	.54	3.08	1.75	.635
			Item 50	.74	3.82	1.71	.740
			Item 51	.73	2.88	1.81	.716
Scale 5	.824	.855	Item 52	.60	5.09	1.39	.516
Variety of Tasks			Item 53	.85	4.53	1.60	.781
			Item 54	.78	4.42	1.64	.767
Scale 6	.872	.872	Item 56	.67	4.18	1.53	.774
Autonomy			Item 57	.77	4.32	1.43	.787
			Item 58	.77	4.35	1.43	.773
Scale 7*	.428	.010	Item 60	.54	5.67	1.27	.443
Relevance of Tasks			Item 61	.32	4.85	1.86	.217
Scale 8	.841	.841	Item 63	.66	5.09	1.39	.516
Complexity of			Item 64	.79	4.53	1.60	.781
Tasks			Item 65	.59	4.42	1.64	.767
Scale 9	.677	.690	Item 67	.64	2.73	1.21	.484
Training Re-quirements and			Item 68	.58	3.06	1.42	.375
Ability Level			Item 71	.40	3.46	1.24	.477
Scale 10	.881	.881	Item 72	.71	4.18	1.53	.774
Involvement in occupational			Item 73	.74	4.32	1.43	.787
Expert Culture			Item 74	.75	4.35	1.43	.773

Scale 11	.801	.793	Item 75	.71	5.48	1.29	.568
Functional			Item 76	.55	4.97	1.43	.486
Involvement			Item 77	.61	4.81	1.63	.618
			Item 78	.67	5.32	1.42	.624
			Item 79	.71	5.46	1.30	.646
Scale 12	.786	.779	Item 80	.70	5.45	1.25	.640
Social			Item 81	.61	4.84	1.71	.599
Involvement			Item 83	.57	6.27	1.14	.506
			Item 84	.61	6.37	1.13	.537
Scale 13	.759	.790	Item 85	.64	5.43	1.65	.602
Mentoring			Item 86	.58	5.31	1.56	.552
			Item 88	.59	5.79	1.30	.455
Scale 14	.824	.826	Item 89	.69	5.03	1.76	.650
Curriculum			Item 90	.80	4.77	1.69	.737
Orientation			Item 91	.62	5.18	1.92	.661
Scale 15	.721	.696	Item 92	.59	5.09	1.75	.496
Feedback			Item 93	.57	5.26	1.21	.487
			Item 94	.53	5.25	1.44	.496
			Item 95	.54	4.94	1.48	.600
Scale 16	.872	.872	Item 97	.62	5.46	1.42	.732
Personnel and			Item 98	.69	5.39	1.34	.798
Instructions			Item 100	.72	4.90	1.37	.729
			Item 101	.67	5.70	1.57	.754
Scale 17	.893	.897	Item 105	.87	2.08	1.52	.834
Premature			Item 106	.80	1.93	1.48	.752
Termination of			Item 107	.77	1.95	1.48	.792
Contract			Item 108	.48	1.66	1.13	.548
			Item 109	.52	1.66	1.17	.578
			Item 110	.78	1.86	1.44	.708
			Item 111	.78	1.83	1.43	.620
Scale 18	.777	.775	Item 112	.53	4.87	1.08	.489
Professional			Item 113	.48	5.67	1.06	.479
Competence			Item 114	.63	5.59	1.05	.641
			Item 115	.77	5.52	1.10	.701
			Item 116	.63	5.36	1.40	.493
Scale 19	.850	.855	Item 125	.78	5.13	1.61	.496
Overall Assess-			Item 126	.70	5.18	1.30	.758
ment and Satis-							
faction							

Scale 20	.872	.877	Item 127	.74	5.20	1.30	.709
Vocational			Item 130	.80	4.81	1.45	.699
Identity			Item 131	.68	4.78	1.69	.691
Scale 21	.873	.877	Item 132	.80	5.13	1.54	.719
Operational			Item 134	.67	3.99	1.96	.666
Identity			Item 136	.80	4.92	1.81	.740
Scale 22	.777	.776	Item 137	.60	4.98	1.81	.639
Future Prospects and Career Aspirations			Item 138	.73	4.57	2.04	.639

Note. Scale ranging from 1 (strongly disagree) to 7 (completely agree), *N* = 428.

*Scale 7 ('Relevance of Tasks') did not show sufficient reliability but can be included by researchers who want to check for work task relevance despite the weak values.

5 Quality criteria

Objectivity

The scales objectively represent the most crucial quality criteria of vocational training identified in the literature. During the application of the instrument, the researcher does not have a negative impact on the objectivity as trainees do not need any assistance or guidance in filling out the easy understandable instrument. Concerning the evaluation and interpretation, it has to be considered that trainee responses represent their subjective perception of vocational training within a specific training company and thus do not constitute an objective measurement. A fully objective measurement does not seem suitable for vocational training reality since there are no objective quality standards available for most criteria and scientific observation of large numbers of training processes does not appear to be economically feasible.

Nevertheless, on an aggregated level survey responses can give valuable insights into the state and changes of training quality with regard to specific companies, specific training occupations, or fields of vocational training. The one-sided measurement approach via trainee perception has to be considered. In this regard, the instrument can easily be adjusted to an application for trainers or trainees' colleagues to achieve a multi-perspective picture (as was done in Krötz & Deutscher, 2021a/2021b; 2022).

Validity

Content validity can be assumed by the preceding qualitative meta-synthesis and the theoretical considerations in organizing and reformulating the existent (validated) survey instruments. Besides testing the VET-LQI's structural validity via CFA (see above), the average variance extracted (AVE) criterion yielded good results (> .5) for 16 of 22 factors regarding the amount of variance explained, while following Fornell and Larcker (1981) five factors were just slightly below (> .4). The scale '*Relevance of Tasks*' (scale 7), that already showed low internal consistency (see table 24), only achieved a value of .373. Furthermore, Intercorrelation between the 22 factors and the Fornell-Larcker criterion were assessed to

measure the degree to which the different quality criteria are unrelated. 13 factors met the Fornell-Larcker criterion (Fornell & Larcker, 1981). The factor correlations in table 25 show that nineteen factors had acceptable moderate correlation values. Only factors 19, 20, and 22 indicate problematic correlations of > .8 (e.g., Evans 1996).

Table 25. Correlations between latent factors

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17	F18	F19	F20	F21	F22
F1	1	.450	.171	-.590	.588	.431	.342	.106	-.047	.581	.559	.758	.744	.461	.636	.634	-.624	.245	.660	.724	.470	.744
F2		1	.378	-.320	.601	.221	.276	.260	.108	.376	.262	.345	.407	.551	.398	.476	-.359	.120	.448	.507	.313	.466
F3			1	-.171	.219	.139	.182	.230	-.107	.235	.191	.183	.196	.278	.317	.254	-.145	.127	.265	.302	.308	.255
F4				1	-.472	-.401	-.232	.082	.480	-.391	-.341	-.498	-.619	-.459	-.553	-.430	.627	-.320	-.570	-.628	-.470	-.540
F5					1	.452	.538	.379	.052	.517	.481	.547	.559	.550	.585	.509	-.492	.296	.582	.622	.362	.570
F6						1	.505	.225	-.091	.658	.514	.549	.284	.170	.442	.233	-.268	.388	.372	.405	.335	.436
F7							1	.389	-.044	.562	.642	.651	.468	.350	.592	.392	-.387	.475	.457	.493	.437	.476
F8								1	.302	.262	.273	.199	.154	.159	.188	.148	-.053	.140	.196	.135	.129	.147
F9									1	-.123	-.262	-.086	-.124	-.087	-.170	-.038	.140	-.431	-.054	-.091	-.245	-.100
F10										1	.699	.664	.415	.406	.539	.355	-.369	.413	.505	.546	.486	.601
F11											1	.650	.494	.341	.548	.440	-.405	.544	.472	.523	.547	.585
F12												1	.684	.475	.711	.514	-.537	.429	.652	.677	.556	.689
F13													1	.735	.733	.732	-.589	.191	.671	.777	.411	.600
F14														1	.570	.491	-.488	.180	.616	.682	.418	.529
F15															1	.677	-.534	.386	.725	.694	.487	.634
F16																1	-.512	.275	.625	.714	.448	.660
F17																	1	-.384	-.584	-.738	-.538	-.696
F18																		1	.245	.382	.484	.353
F19																			1	.906	.786	.858
F20																				1	.706	.861
F21																					1	.782
F22																						1

Reliability

Reliability in terms of Cronbach's Alpha of the final scales slightly changed after the item eliminations, in most cases implying slight decreases in Cronbach's Alpha. However, except for scale 7 ('Relevance of Tasks'), the final values achieve a satisfactory level (see table 24 and section 4 'Scale Development').

Descriptive statistics (scaling)

Descriptive information on the scales and items is shown above in table 24. The data is mostly left-skewed for the relatively high agreement on the items, corresponding to relatively high training quality.

Further quality criteria

The maximum processing time for participants amounts to 45 minutes. However, as an experience value from several further applications, most trainees finish within 30 Minutes.

In our experience most trainees tend to give honest responses due to the anonymity in the survey design. Only a small and neglectable share used response behaviours due to laziness as e.g., crossing the same response for the whole survey page. Faking towards socially desirable values was no issue since the majority of trainees appears to be glad for having the opportunity to evaluate the training reality in their company, stating the perceived strengths and weaknesses as opposed to being the ones who usually become evaluated by the training personnel. There are no other biases known. Invariances across different person/group characteristics are potential research purposes of the survey as, for example, relations between specific characteristics (e.g., education) and distinct perceptions of training quality could be revealed.

6 Literature and data sources

Further literature [original validation:]

Böhn, S., & Deutscher, V. (2019). Betriebliche Ausbildungsbedingungen im dualen System – Eine qualitative Meta-Analyse zur Operationalisierung in Auszubildendenbefragungen [Training Conditions in VET – A Qualitative Meta-Synthesis for the Operationalization in Apprentice Questionnaires]. *Zeitschrift für Pädagogische Psychologie, 33*(1), 49–70.

Böhn, S., & Deutscher, V. (2021). Development and Validation of a Learning Quality Inventory for In-Company Training in VET (VET-LQI). *Vocations and Learning 14*, 23–53. https://doi.org/10.1007/s12186-020-09251-3

Further studies the instrument has been (partly) used for:

Krötz, M., & Deutscher, V. (2021a). Differences in perception matter – How differences in the perception of training quality of trainees and trainers affect drop-out in VET. *Vocations and Learning, 14*, 369–409. https://doi.org/10.1007/s12186-021-09263-7

Krötz, M., & Deutscher, V. (2021b). Betriebliche Ausbildungsqualität – Eine Frage der Perspektive? [Quality of In-Company Training – A matter of Perspective?]. *Zeitschrift für Erziehungswissenschaft: ZfE, 24*, 1453–1475. https://doi.org/10.1007/s11618-021-01041-4

Krötz, M., & Deutscher, V. (2022). Drop-out in dual VET: Why we should consider the drop-out direction when analysing drop-out. *Empirical Research in Vocational Education and Training,   14*, 1–26. https://doi.org/10.1186/s40461-021-00127-x

Contact details

Maximilian Krötz, M. Sc.

Research Assistant

Mail: kroetz@uni-mannheim.de

Prof. Dr. Viola Deutscher

Mail: viola.deutscher@uni-mannheim.de

Chair for Economic and Business Education

Competency Development and Training Quality

University of Mannheim

L4, 1

68161 Mannheim

References

Beicht, U., Krewerth, A., Eberhard, V., & Granato, M. (2009). *Viel Licht – aber auch Schatten. Qualität dualer Berufsausbildung in Deutschland aus Sicht der Auszubildenden* (BIBB-Report 09/09: Forschungs- und Arbeitsergebnisse aus dem Bundesinstitut für Berufsbildung) [Many bright spots - but shadows too. The quality of dual vocational training from the trainees' point of view]. Berlin: Bundesinstitut für Berufsbildung (BIBB).

Biggs, J. B. (1999). Teaching for quality learning at university: What the student does. Buckingham: Open University Press.

Böhn S. (2020). Training quality and premature termination of apprenticeship contract –conceptualization, operationalization and measurement. Inauguraldissertation, University of Mannheim.

Böhn, S., & Deutscher, V. (2019). Betriebliche Ausbildungsbedingungen im dualen System – Eine qualitative Meta-Analyse zur Operationalisierung in Auszubildendenbefragungen [Training Conditions in VET – A Qualitative Meta-Synthesis for the Operationalization in Apprentice Questionnaires]. Zeitschrift für Pädagogische Psychologie, 33(1), 49–70.

Böhn, S., & Deutscher, V. (2021). Development and Validation of a Learning Quality Inventory for In-Company Training in VET (VET-LQI). Vocations and Learning 14, 23–53. https://doi.org/10.1007/s12186-020-09251-3

Blom, K. & Meyers, D. (2003). Quality indicators in vocational education and training: International perspectives. Adelaide: NCVER.

Bollen, K. A., & Lennox, R. (1991). Conventional Wisdom in Measurement: A Structural Equation Perspective. *Psychological Bulletin, 110*(2), 305–314.

Boomsma, A., & Hoogland, J. J. (2001). The robustness of LISREL modeling revisited. In R. Cudeck, S. du Toit, & D. Sörbom (Eds.), *Structural equation models: Present and future. A Festschrift in honor of Karl Jöreskog* (139–168). Chicago: Scientific Software International.

Brooker, R., & Butler, J. (1997). The learning context within the workplace: as perceived by apprentices and their workplace trainers. *Journal of Vocational Education & Training, 49*(4), 487–510.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (136−162). Newbury Park: Sage.

Byrne, B. M. (1991). The Maslach burnout inventory: Validating factorial structure and invariance across intermediate, secondary, and university educators. *Multivariate Behavioral Research, 26*(4), 583–605.

Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, 1(1), 16–29.

DeVellis, R. F. (1991). *Scale development: theory and applications*. Thousand Oaks: Sage Publications, Inc.

Deutscher Gewerkschaftsbund (DGB) (2008). *Ausbildungsreport 2008* [Training report 2008]. Berlin: Deutscher Gewerkschaftsbund.

Diamantopoulos, A., & Winklhofer, H. M. (2001). Index Construction with Formative Indicators: An Alternative to Scale Development. *Journal of Marketing Research, 38*(2), 269–277.

Dwyer, P., Harwood, A., Costin, G., Landy, M., Towsty, L., & Wyn, J. (1999). *Combined Study and Work Paths in VET: Policy Implications and Analysis*. Adelaide: NCVER.

Ebel, R. L., & Frisbie, D. A. (1986). *Essentials of education measurement*. Englewood Cliffs: Prentice Hall.

Ebner, H. G. (1997). Die Sicht der Auszubildenden auf die Ausbildung. In D. Euler & P. F. E. Sloane (Eds.), *Duales System im Umbruch. Eine Bestandsaufnahme der Modernisierungsdebatte* [Dual system in transition. A review on modernization debate] (247–262). Pfaffenweiler: Centaurus-Verlagsgesellschaft.

Ernst, C. (2016). *Forschungsprojekt 'Auszubildendenzufriedenheit'. Abschlussbericht* [Research project 'Apprentices satisfaction'. Final report]. Fakultät für Wirtschafts- und Rechtswissenschaften der Technischen Hochschule Köln.

Evans, J. D. (1996). Straightforward statistics for the behavioral sciences. Belmont: Thomson Brooks/Cole Publishing.

Feller, G. (1995). Ansprüche und Wertungen junger Menschen in der Berufsausbildung. Wie und was Auszubildende (nicht) lernen wollen [Claims and evaluations of young people in vocational education and training. How and what apprentices do (not) want to learn]. *Berufsbildung in Wissenschaft und Praxis, 24*(2), 18–23.

Fink, R. (2015). *Strategische Ausbildungsplanung und interne Evaluation: Steuerung pädagogischen Handelns, interne Evaluation, Fragebögen* [Strategic planning of vocational education and training and internal evaluation. Management of educational action, internal evaluation, questionnaires].

Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research, 18*(1), 39–50.

Fuller, A., & Unwin, L. (2003). Fostering Workplace Learning: looking through the lens of apprenticeship. *European Educational Research Journal, 2*(1), 41–55.

Gebhardt, A., Martínez Zaugg, Y., & Metzger, C. (2014). Motivationale, emotionale und selbstwirksamkeitsbezogene Dispositionen von Auszubildenden und deren Wahrnehmung der Lernumgebung und Lernbegleitung im betrieblichen Teil der beruflichen Grundbildung [Motivation, emotional and self-efficacy-related dispositions of trainees and their perceptions of the learning environment and learning support in the company part of initial vocational education and training]. *bwp@ Berufs- und Wirtschaftspädagogik – online, 26: Berufliche Bildungsprozesse aus der Perspektive der Lernenden*, 1–23.

Heinemann, L., Maurer, A., & Rauner, F. (2009). Engagement und Ausbildungsorganisation. Einstellungen Bremerhavener Auszubildender zu ihrem Beruf und ihrer Ausbildung. Eine Studie im Auftrag der Industrie- und Handelskammer Bremerhaven [Engagement and organization of vocational education and training. Attitudes of apprentices in Bremerhaven regarding their occupation and their training. A study of the German Chamber of Industry and Commerce Bremerhaven].

Hofmann, C., Stalder, B. E., Tschan, F., & Häfeli, K. (2014). Support from teachers and trainers in vocational education and training: The pathways to career aspirations and further career development. *International Journal for Research in Vocational Education and Training, 1*(1), 1–20.

Hu, L., & Bentler, P. (1995). Evaluating model fit. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (76−99). London: Sage.

Hu, L., & Bentler, P. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55.

Keck, A., Weymar, B., & Diepold, P. (1997). *Lernen an kaufmännischen Arbeitsplätzen. Berichte zur beruflichen Bildung* [Workplace learning in commercial occupations. Reports on vocational training], 199. Berlin und Bonn: Bundesinstitut für Berufsbildung (BIBB).

Klotz, V. K., Rausch, A., Geigle, S., & Seifried, J. (2017). Ausbildungsqualität – Theoretische Modellierung und Analyse ausgewählter Befragungsinstrumente [Training quality – Theoretical Modeling and Analysis of Selected Test Instruments]. In S. Matthäus, C. Aprea, D. Ifenthaler & J. Seifried (Eds.), *bwp@ Berufs- und Wirtschaftspädagogik online, Profil 5: Entwicklung, Evaluation und Qualitätsmanagement von beruflichem Lehren und Lernen. Digitale Festschrift für Hermann G. Ebner*, 1–16.

Koch, A. F. (2016). Zufriedenheit in der Berufsausbildung. Konstruktion eines Fragebogens zur Erfassung intrinsischer Lern- und Arbeitsmotivation bei Auszubildenden im industriellen und kaufmännischen Sektor [Satisfaction with vocational education and training. Construction of a questionnaire for the acquisition of intrinsic learning and working motivation among apprentices in the industrial and commercial sector]. Hamburg: Verlag Dr. Kovac.

Krötz, M., & Deutscher, V. (2021a). Differences in perception matter – How differences in the perception of training quality of trainees and trainers affect drop-out in VET. Vocations and Learning, 14, 369–409. https://doi.org/10.1007/s12186-021-09263-7

Krötz, M., & Deutscher, V. (2021b). Betriebliche Ausbildungsqualität – Eine Frage der Perspektive? [Quality of In-Company Training – A matter of Perspective?]. Zeitschrift für Erziehungswis-senschaft: ZfE, 24, 1453–1475. https://doi.org/10.1007/s11618-021-01041-4

Krötz, M., & Deutscher, V. (2022). Drop-out in dual VET: Why we should consider the drop-out direction when analysing drop-out. Empirical Research in Vocational Education and Training, 14, 1–26. https://doi.org/10.1186/s40461-021-00127-x

Ma, B., Krötz, M., Deutscher, V., & Winther, E. (in preparation). Same, same but different? Testing measurement invariance of vocational competence in digital versus analog test formats. *Educational Measurement*.

Matsunaga, M. (2010). How to factor-analyze your data right: Do's, don'ts and how-to's. *International Journal of Psychological Research, 3*(1), 97–110.

Mayring, P. (1994). Qualitative Inhaltsanalyse. In A. Böhm, A. Mengel & T. Muhr (Eds.), *Texte verstehen: Konzepte, Methoden, Werkzeuge* [Understanding texts. Concepts, methods, instruments] (159–175). Konstanz: UVK.

NCVER (2000). *Student Outcomes Survey 2000*. National Report. Adelaide: NCVER.

Nickolaus, R., Nitzschke, A., Maier, A., Schnitzler, A., Velten, A., & Dietzen, A. (2015). Einflüsse schulischer und betrieblicher Ausbildungsqualitäten auf die Entwicklung des Fachwissens und die fachspezifische Problemlösekompetenz [The quality of in-company education and training and the development of competence]. *Zeitschrift für Berufs- und Wirtschaftspädagogik, 111*(3), 333–358.

Nunnally, J. C. (1978). *Psychometric theory*. New York: McGraw-Hill.

Prenzel, M., Kristen, A., Dengler, P., Ettle, R., & Beer, T. (1996). Selbstbestimmt motiviertes und interessiertes Lernen in der kaufmännischen Erstausbildung. In K. Beck & H. Heid (Eds.), *Lehr-Lern-Prozesse in der kaufmännischen Erstausbildung: Wissenserwerb, Motivierungsgeschehen und Handlungskompetenzen* [Teaching and learning processes in vocational education and training in the commercial sector. Acquisition of knowledge, motivation and decision-making] (110–127). Stuttgart: Steiner.

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online, 8*(2), 23–74.

Seyfried, E., Kohlmeyer, K., & Furth-Riedesser, R. (2000). *Supporting Quality in Vocational Training through Networking: CEDEFOP Panorama*. Thessaloniki: European Centre for the Development of Vocational Training.

Tynjälä, P. (2013). Toward a 3-p model of workplace learning: a literature review. *Vocations and Learning, 6*, 11–36.

Ulrich, J. G., & Tuschke, H. (1995). Probleme während der Lehre: Kritikpunkte von ostdeutschen Auszubildenden [Problems during apprenticeships: critical points of east German apprentices]. *Sozialwissenschaften und Berufspraxis, 18*(3), 198–212.

Velten, S., & Schnitzler, A. (2012). Inventar zur betrieblichen Ausbildungsqualität (IBAQ) [Inventory of in-company training conditions (IBAQ)]. *Zeitschrift für Berufs- und Wirtschaftspädagogik, 108*(4), 511–527.

Virtanen, A., & Tynjälä, P. (2008). Students' experiences of workplace learning in Finnish VET. *European Journal of Vocational Training, 44*(2), 199–213.

Virtanen, A., Tynjälä, P., & Collin, K. (2009). Characteristics of workplace learning among Finnish vocational students. *Vocations and Learning, 2*, 153–175.

Virtanen, A., Tynjälä, P., & Eteläpelto, A. (2014). Factors promoting vocational students' learning at work: study on student experiences. *Journal of Education and Work, 27*(1), 43–70.

Visser, K. (1994). Systems and procedures of certification of qualifications in the Netherlands. National Report. CEDEFOP Panorama. Berlin: European Centre for the Development of Vocational Training.

Walker, A., Smith, E., & Brennan-Kemmis, R. (2012). The psychological contract in apprenticeships and traineeships: Comparing the perceptions of employees and employers. *International Employment Relations Review, 18*(1), 66–81.

Zimmermann, M., Wild, K.-P., & Müller, W. (1994). Entwicklung und Überprüfung des 'Mannheimer Inventars zur Erfassung betrieblicher Ausbildungssituationen (MIZEBA)'. Forschungsbericht der Universität Mannheim [Development and validation of the 'Mannheim inventory for the acquisition of in-company training conditions (MIZEBA)'. Research report of the University of Mannheim].