Teaching and Teacher Education 122 (2023) 103922

Contents lists available at ScienceDirect

Teaching and Teacher Education

journal homepage: www.elsevier.com/locate/tate

Research paper

Who's included and Who's not? An analysis of instruments that measure teachers' attitudes towards inclusive education☆

Stephan Kielblock [a, *], Stuart Woodcock [b]

[a] DIPF | Leibniz Institute for Research and Information in Education, Germany
[b] Griffith University, Australia

HIGHLIGHTS

- Many different instruments to measure teachers' attitudes to inclusion exist.
- Nearly all of these instruments are exclusive to some extent.
- Some inclusive attitude instruments lack thorough development and validation.
- There was no ideal instrument to measure attitudes towards inclusion for all.
- New attempts are needed to capture teachers' attitudes to inclusion for all.

ARTICLE INFO

Article history:
Received 23 January 2022
Received in revised form
29 September 2022
Accepted 11 October 2022
Available online 29 November 2022

Keywords:
Inclusive education
Attitudes
Measurement
Teachers
Systematic literature review

ABSTRACT

In this paper, inclusive education was understood as the right of all students to access, presence, participation and success in their local school (Slee, 2018). With a focus on teachers' attitudes towards inclusive education, the measurement instruments in 225 relevant empirical studies were examined. The findings showed that the vast majority of instruments utilised exclusionary wording focusing on particular (groups of) learners. There was no ideal instrument which was purely inclusive. It was concluded that there is an urgent need for a new measurement instrument that operationalises the attitude towards inclusive education for all students.

1. Introduction

Since the 1990s, the term 'inclusive education' has gained momentum in research, policies, and practice. Particularly, the Salamanca Statement (UNESCO, 1994) began the movement towards this term in the sense of enabling schools to serve all children. Over the years, a variety of policies were developed to support this process (Burnett, 2008; Kiuppis, 2014; Mundy, 2016; Peters, 2007) and a large number of studies have been conducted on different aspects of inclusive education (for an overview see Van Mieghem et al., 2020).

1.1. Inclusive education

In many studies, inclusive education is defined as the inclusion of particular students or groups of students into regular classes and/or regular schools. In a review of 640 articles published in the International Journal of Inclusive Education between 2005 and 2015, Messiou (2017) found that forty percent of the articles had a specific focus on students with special educational needs and/or disabilities (SEND). The other papers in the study investigated other particular groups of learners, such as children at risk, etc., and only eight percent of the papers utilised a broader focus on all students or on diversity (Messiou, 2017). A study conducted by Nilholm and Göransson (2017) supported this finding. The authors reviewed the 60 most cited articles on inclusive education in Europe and North America. Almost eighty percent of the studies were based on a definition of inclusive education, which focused on the placement of students with SEND in general educational settings or meeting

☆ We have no known conflict of interest to disclose.

* Corresponding author.
E-mail address: s.kielblock@dipf.de (S. Kielblock).

https://doi.org/10.1016/j.tate.2022.103922

their needs (Nilholm & Göransson, 2017). The predominant view on inclusive education tends to be focused on catering *for some* (groups of) students, who are considered to be particularly vulnerable to exclusion.

Aside from the predominant view on inclusive education as catering for *some* students (e.g., students with SEND), many researchers advocate for a broader conceptualisation of inclusive education. According to Thomas (2013), the focus on the 'exceptionality' of a student makes it difficult to challenge implicit beliefs about the students' (dis)ability. In this way, the extensive use of 'labels' to articulate the students' (dis)abilities and needs might actually decelerate progress towards more equity in education and society (Booth, 1995; Graham & Macartney, 2012; Graham & Slee, 2008). That may be a reason why researchers emphasise that inclusive education should focus on *all* children and should primarily enable the community and belonging *for all* (Carrington et al., 2012; Thomas, 2013).

In the present study, inclusive education is understood in line with Slee as "the right of all children to access, presence, participation and success in their local regular school" (2018, p. 8; see also Ainscow et al., 2006; Peters, 2004; Shyman, 2015; Thomas, 2013; UNESCO, 2015). An important aspect of this definition is the inclusion of *all children* in education. The imperative to tackle exclusion and marginalisation that *some* (groups of) individuals face might make it necessary to use labels and categorisations to address these issues. Inclusion comprises the imperative to initiate all necessary steps on all necessary levels that these individuals are again embraced by the notion of 'all'. Inclusive education shifts the focus from labels, diagnosis and deficit of some students to quality education *for all* children. As articulated in the Index for Inclusion, "inclusion is about the education of all children and young people", although "inclusion is often associated with students who have impairments or students seen as 'having special educational needs'" (Booth & Ainscow, 2002, p. 1). We agree with the Incheon Declaration that 'all' means all, "irrespective of sex, age, race, colour, ethnicity, language, religion, political or other opinion, national or social origin, property or birth, as well as persons with disabilities, migrants, indigenous peoples, and children and youth" (UNESCO, 2015, p. 25). Throughout the present study, this kind of thinking about inclusion is referred to as *inclusive education for all*, as opposed to *inclusive education for some*.

1.2. Teachers and their attitudes

It is well-documented in numerous studies that teachers are crucial for positive learning and outcomes for students (Coleman et al., 1966; Goldhaber, 2016; Hattie, 2003, 2009; Mansfield, 2015; OECD, 2005; Rivkin et al., 2005). Accordingly, it is no surprise that the pre-eminent role of teachers in providing quality education for all students was also emphasised in the early years of the inclusion movement (UNESCO, 1990, 1994; UNESCO & UNDP, 1995). More recently, policies specified that teachers must understand diversity in the students' learning and must be able to differentiate in the classroom (UNESCO, 2000, 2005, 2015). In this way, the inclusive mind-sets and attitudes of teachers have been intensively studied (Amor et al., 2019; Van Mieghem et al., 2020).

Attitudes are a tendency of an individual to evaluate an object in a favourable or unfavourable way (Eagly & Chaiken, 1993). According to social psychologists (Rosenberg & Hovland, 1960; Triandis, 1971), attitudes comprise a cognitive, an affective, and a behavioural dimension. Furthermore, sociologists emphasise that the development and enactment of attitudes should be understood in terms of the interrelatedness of the individual and its social environment (Hitlin & Pinkston, 2013; Voas, 2014). Research has shown that there is not a mechanistic relationship between attitudes and the actual behaviour (Chaiklin, 2011; Glasman & Albarracín, 2006). Theory suggests that behaviour is shaped by a complex array of behavioural intents, attitudes, subjective norms, and the perceived behavioural control (Ajzen, 1988, 1991). Taken together, the attitudes of an individual are both, a mirror of its perception of the social environment, and a precursor of subsequent social action. Therefore, teachers' attitudes towards inclusive education can be considered as an important trigger of inclusive teaching.

1.3. Objective

As advocates for seeing inclusive education as catering *for all* students instead of for some students we believe that policy and practice should be guided by high-quality empirical studies on inclusive education *for all*. Yet, according to the literature analysis conducted by Nilholm and Göransson (2017), the understanding of inclusive education *for all* is represented particularly in conceptual papers and position papers, while empirical research papers tend to be based on the idea of inclusive education *for some*. In other words, empirical papers seem to be particularly prone to focus solely on inclusive education *for some.* In another review study, Ruberg and Porsch (2017) found that empirical papers that utilise inclusive education *for all* as a theoretical framework, actually based their methods, results and discussion sections on ideas of inclusive education *for some*. It might be a valid hypothesis that empirical researchers are using long-time established measurement instruments, which were originally developed in the spirit of mainstreaming or integration and, hence, they might actually operationalise inclusion *for some* and not *for all*. A comprehensive analysis of published empirical papers would allow the opportunity to challenge this hypothesis. However, existing reviews only analysed the variety of topics of inclusive education research (Avramidis & Norwich, 2002; de Boer et al., 2011; Van Mieghem et al., 2020), the effects of inclusion on students with and/or without SEND (Ruijs & Peetsma, 2009; Szumski et al., 2017), and inclusion in specific fields (such as physical education, Tant & Watelain, 2016). Not many reviews have focused particularly on the measurement instruments which are used in inclusive education research. The above-mentioned review of Ruberg and Porsch (2017) only focused on German instruments and was very broad in its scope, focusing – besides the measurement instruments – on the definitions of inclusive education, as well as on the findings, as they pertain to inclusive education. Another recent review conducted by Ewing et al. (2018) focused on instruments to measure attitudes towards inclusive education. They attempted to find instruments in English-speaking papers and analyse the scales' quality. Yet, their study did not examine whether these instruments, and the utilised items, were actually capturing the teachers' attitudes towards inclusive education *for all* or *for some*.

The present study attempts to remedy this research gap by conducting a comprehensive analysis of the measurement instruments in published empirical papers. As argued in the introduction of this paper, the teachers' attitudes towards inclusive education seem to be a crucial topic, because a) attitudes trigger inclusive teaching behaviour, b) they are highlighted in policies as being of major importance, and c) they are one of the most commonly researched topics in the field of inclusion. Accordingly, empirical papers on the teachers' attitudes towards inclusive education are targeted by the present systematic review.

The overarching research question for this study is, to what extent do the instruments that measure teachers' attitudes towards inclusive education, which are utilised in published empirical studies, adequately represent inclusive education *for all*? Two subquestions are specified: (1) To what extent do the wording and

terminology used in the instruments adequately represent inclusive education *for all*? (2) Regarding the instruments that utilised the most relevant terminology, have they been developed in line with current recommendations for scale development so that it can be assumed that these instruments are of a high quality?

2. Method

The current study carried out a systematic literature review. Unsystematic, 'traditional' literature reviews tend to be based on an arbitrary selection of references (Petticrew & Roberts, 2006). Hence, they do not represent the existing literature sufficiently. Using a systematic approach to literature searches and literature analyses mitigates this bias. The current study was informed by widely accepted recommendations for carrying out systematic reviews (Evans & Benefield, 2001; Gough et al., 2017; Moher et al., 2009; Petticrew & Roberts, 2006).

2.1. Inclusion and exclusion criteria of the studies

Four criteria have been specified in order to find the most relevant papers on teachers' attitudes towards inclusive education (see Table 1). It can be assumed that the most rigorous studies are published in peer-reviewed journals. In meta-analyses, 'grey literature' is also included in the analysis in order to mitigate the so-called publication bias. Yet, the present study is focused on measurement instruments and not the effects. The incorporation of publications other than journal articles would make the search procedures disproportionately complex and less comprehensible. Hence, relevant studies were those published in peer-reviewed journals. English is one of the most important languages in educational research and on teachers' attitudes towards inclusive education. Thus, relevant records for the present study were thought to be published in English language. In journal articles it can be assumed that the focus of the study is clearly stated in the study's title. Hence, the type of studies, which were defined as being relevant for the present review, had "teachers' attitudes towards inclusive education" (in different variations) in the papers' titles. Studies on the teachers' attitudes towards inclusive education used a variety of methods (e.g., interviews, surveys, systematic literature reviews). However, only those studies that actually operationalised and measured the teachers' attitudes were of particular interest for the present study.

2.2. Search strategies

The systematic literature search was carried out in 2019. No restrictions regarding the time range were specified. Hence, all papers, which were published up to 2019 were potentially eligible to be included in the body of literature. As Petticrew and Roberts (2006) suggested, a bibliographic software (Endnote, Clarivate Analytics, Version X8) which was used to organise and document the search process.

Three different literature search portals were selected to carry out the literature search, namely *ProQuest* (searching databases such as ERIC, Applied Social Sciences Index & Abstracts, and Sociological Abstracts), *EBSCOhost* (searching databases such as PsycINFO and Education Source) and the *Web of Science Core Collection* (searching databases such as the Social Sciences Citation Index and the Arts & Humanities Citation Index). These three literature search portals index the most important databases in educational research and cover a considerable number and significant variety of available records in educational research.

In order to find appropriate search parameters in regard to the teachers' attitudes towards inclusive education, a series of preliminary search trials were carried out. First of all, the populations of interest comprised pre-service and in-service teachers. To be open to both groups the word "teacher*" (the * allows different endings of the word like teachers or teachers') was used as the first part of the search term. As a few relevant studies used educator instead of teacher, "educator*" was added, too. Second, some of the studies on teachers' attitudes used the term beliefs or opinions. Hence, these terms were also used (attitude* OR belief* OR opinion*). Third, the studies should have a focus on inclusive education. In some studies, the brief term "inclusion" was used, although inclusive education was actually meant. Hence, inclusi* was used to cover all variations like inclusion and inclusive. The full expression used for the literature search was "(teacher* OR educator*) AND (attitude* OR belief* OR opinion*) AND (inclusi*)".

After implementation of the retrieved records in Endnote, duplicates, which appeared because of similar coverage of the searched databases, were deleted. The next step carefully re-examined whether they were published in journals and in English language. Then, the titles were reviewed, if the studies' content comprised "teachers' attitudes towards inclusive education". All records that did not fit these criteria were deleted. After this verification, the abstracts were examined whether the studies had utilised a quantitative approach. All records that did not use quantitative methods were deleted from the list. Mixed methods studies, with a clear statement that attitudes were measured retained in the list. For the remaining records, copies of the papers were obtained.

2.3. Procedures of the analysis

Papers that remained after applying the criteria, which were presented in the previous section, were analysed further. As the focus of analysis was on the instruments measuring teachers' attitudes towards inclusive education, each individual measurement instrument was analysed separately. More specifically, if a study utilised more than one relevant attitude instrument, these attitude instruments were analysed separately.

As the attitude measurement instruments were of particular interest for the present study, they were extracted from the obtained studies and it was noted, if a new instrument was developed or if an established instrument was utilised. Established means that an instrument was already published before. If the wording of the items was included in the paper, it was extracted and used for analysis. This was found to be the most robust way to obtain the wording, because changes and adaptations of established

Table 1
Inclusion and exclusion criteria regarding the paper selection.

Inclusion criteria	Exclusion criteria
•Articles published in peer-reviewed journals	•Book chapters, conference papers, dissertations, reports, grey literature
•English language	•Non-English language
•"Teachers' attitudes towards inclusive education" is mentioned in the title	•Teachers' attitudes towards inclusive education are not included in the title in any variation •These terms are included in the title, but with obviously different meaning
•Quantitative methods/measurement	•Narrative interview studies, observational studies, literature review studies, position papers

instruments were captured as part of our data in this way.

If the wording was not included in the paper, and if an established instrument was used, it was examined whether the original instrument had been modified, adapted or changed in any way. If no modifications were indicated in the paper, the original reference of the instrument was searched and used to obtain the wording of the items. The extracted wording of the items to measure teachers' attitudes was utilised in the subsequent analysis as data to answer research question one.

After examination of the empirical studies, the individual items' wording was analysed in more detail. It was anticipated that a large number of items needed to be analysed. Therefore, a quantitative and standardised text analytic approach for analysing the items' wording was utilised. The automatized coding of each single item was carried out in MS Excel. Each item was examined using a complex '= SEARCH()' function, whether particular terms or phrases were present in an item.

This approach made it necessary to develop a list of terms and phrases, which could be considered either as a proxy for exclusionary wording or as a proxy for inclusionary wording of an item. Through an iterative process the completeness and validity of the list was ensured.

The initial list was developed in accordance with the introduction of this paper in that inclusive education is meant to be *for all* children. Accordingly, it was assumed that any item that is clearly addressing particular groups of children (e.g., children with disabilities) cannot be a sufficient indicator for the concept of inclusive education *for all*. On the contrary, not all items without exclusionary language are per se inclusive. For example, they could be indicators for something completely off topic. Hence, items that sufficiently indicated the teachers' attitudes towards inclusive education needed to comprise notions of 'inclusion' or 'for all students', too. Accordingly, the initial list comprised 29 exclusionary terms (such as disorder, handicap, normally achieving) and 14 inclusionary terms (such as inclus* [for variations of inclusive], for all students, every child).

After the software had run the automatized coding, a random selection of 100 items was drawn. The assigned coding for these items was examined by both authors of this study, as to whether all of the selected items were coded correctly. Both authors met to discuss any issues with the automatized coding and made decisions, in what way the list needed to be changed. All items were then automatically coded again by using the updated list. This iteration was repeated until both authors found that no further modification was needed.

The final list comprised 46 exclusionary terms and 15 inclusionary terms (see Table 2). "All students with" and "all students who have" were coded as exclusionary, while "all students" (without "… with" or "… who have") was inclusionary. The former refers clearly to particular groups of students (e.g., "all students with disabilities should be in regular classes"), while only the latter refers to all students. Similarly, "inclus" and "inclusion of all students (pupils, children)" was inclusionary, while "inclusion of", without referring to terms such as "all students" for example, was exclusionary (e.g., "the inclusion of students with intellectual disabilities can be beneficial for students without disabilities").

In order to gain insights into which terms were most common amongst the analysed items, it was examined, how many items comprised certain exclusionary or inclusionary terms. The next step of analysis was to count, how many items comprised at least one exclusionary term or at least one inclusionary term. Using these two variables (at least one exclusionary term in an item on the one hand and at least one inclusionary term on the other hand), an item was categorised as being A) fully exclusionary, B) both exclusionary and inclusionary, C) neither exclusionary nor inclusionary, or D) fully inclusionary. Items, which are categorised A), B) or C) cannot be considered inclusive, while the wording of items in the category D) represent wording which might mirror the ideas of inclusive education *for all*. After all individual items were categorised, the number of 'exclusionary' items, 'neither' items (meaning: items which were neither exclusionary nor inclusionary), 'both' items (meaning: items which were both exclusionary and inclusionary at the same time) and 'inclusionary' items per instrument, relative to the number of items in an instrument, were calculated. This relative number indicated, how exclusive or inclusive a particular instrument was. As the focus of the present study was on the inclusiveness of the instruments, the percentages of inclusionary items per instrument were categorised into not inclusive (0% inclusionary items per instrument), low inclusiveness (1–33%), medium inclusiveness (34–66%), high inclusiveness (67–99%) and fully inclusive (100%).

Those instruments which were medium, high or fully inclusive were analysed regarding the instruments' quality. As a proxy for the quality of the instruments it was examined how many steps of the development of the instruments (Carpenter, 2018; DeVellis, 2011; Hinkin et al., 1997) were documented. In particular the recommendations of Hinkin et al. (1997) for scale development were utilised to structure the comparison of the instruments. As well as the administration of the questionnaire, six important steps of scale development include: (1) item generation, (2) content adequacy, (3) factor analysis, (4) internal consistency, (5) (construct) validation, and (6) replication. It was examined, to what extent the papers that reported the development of the instruments included information in regard to these six steps. The number of steps, for which information were given in the paper, was counted. In this way, an index was developed, ranging from 0 = 'no steps of scale development were reported' up to 6 = 'the whole process of the scale development was documented'.

In a final step, the inclusivity of the instruments and the quality of the development of the instruments were analysed together in a scatter plot. The y-axis represented the inclusiveness of the instrument (percentage of inclusive items per instrument). The x-axis represented the quality of scale development (number of documented steps of scale development). Within the scatterplot, those instruments in the upper right corner were the most inclusive and had the highest quality.

2.4. Study selection

The database queries resulted in 193 records from ProQuest, 269 records from EBSCO and 135 records from Clarivate Analytics' Web of Science Core Collection (see Fig. 1). As the results from the three different search portals had overlaps, duplicates were deleted. The resulting number of potentially relevant records was 309. The search parameters (published in a journal; published in English language; title contains variations of 'teachers' attitudes towards inclusive education'; use of quantitative methods) were used to re-examine the 309 records. As can be seen in Fig. 1, 84 studies did not fit these parameters. One study was not published in a journal and nine studies were not published in English language; these ten records were excluded. Although the search terms were relatively precise, a couple of records appeared in the list that were not relevant for the present study (e.g., "School Nurses and *Teachers*: *Attitudes* Regarding *Inclusion* of Breastfeeding Education in School Curricula"). Twenty-eight such records were deleted from the list. In the following step, the abstracts were carefully examined, and 46 studies were found, which solely used *qualitative methods* such as interviews, group discussions or case studies, or no empirical methods at all. These records were deleted from the list of relevant records, too. In the described way, the original number of 309

Table 2
Indicators for exclusionary and inclusionary items.

Exclusionary items		Inclusionary items
- additional support	- iep	- all children
- all students who have[1]	- impair	- all of the students
- all students with[1]	- inclusion of[2]	- all pupils
- asd	- integra	- all students[1]
- autism	- learning difficulties	- all the children
- behavioural difficult	- learning problems	- all the students
- behavioural problems	- mainstream	- every child
- behavioural difficult	- minorit	- every student
- behavioural problems	- normally achieving	- everyone
- blind	- patholog	- for all children
- braille	- regular achieving	- for all pupils
- deaf	- regular student	- for all students
- deficit	- retard	- for all the students
- developmental need	- scholastic service	- for all
- diabetes	- segregate	- inclus[2]
- disab	- self-contained	
- disorder	- separate	
- disturbed	- severe speech difficulties	
- exceptional	- sign language	
- gifted	- special	
- grade skipping	- talent	
- handicap	- typical	
- high-achieving	- wheelchair	

[1] If an item comprised 'all students who have' or 'all students with', the item was coded as 'exclusionary'. If 'all students' was used in an item in any other variation, the item was coded as 'inclusionary'.
[2] If in an item 'inclus' (e.g. inclusive education) was mentioned, the item was 'inclusionary', while it was 'exclusionary, if 'inclusion of' (except for inclusion of all) was utilised.

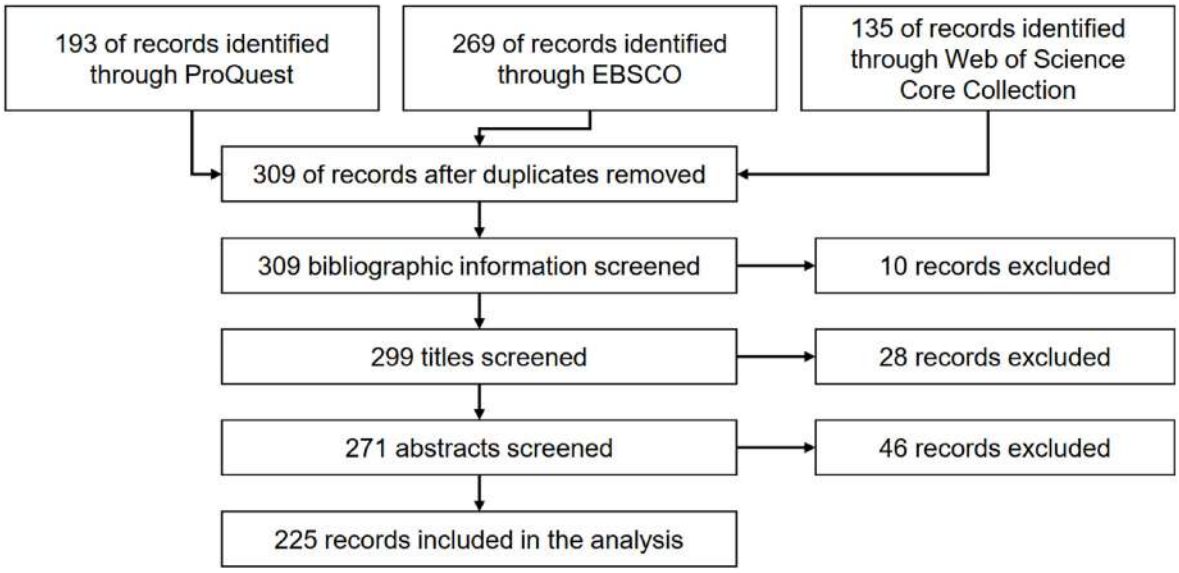

Fig. 1. Deleted and retained records.

records decreased to 225 relevant records.

The analysis of the 225 records revealed that over the years, the number of relevant studies was continuously growing. Although the years of publication were not restricted in the search criteria, the earliest papers that were found to be relevant were Forlin (1995) and Turner (1995). No other papers that fit the criteria had been published before 1995. In the five years before the year 2000, on average, three or four relevant papers were published each year. In the last five-year period of the dataset (2014–2018), between 16 and 17 papers were published on average each year. In 2018, even 24 relevant studies were published, which is the largest number of annually published studies across all years.

The 225 papers were published in 98 different journals. The "International Journal of Inclusive Education" (22 studies; 9.8%) and the "International Journal of Special Education" (21 studies; 9.3%) were the most common journals. Other relevant papers have also been published in the "European Journal of Special Needs Education" (18 studies; 8.0%) or in the "Journal of Research in Special Educational Needs" (16 studies; 7.1%).

3. Results

3.1. Descriptive information regarding the instruments

Three of the 225 studies utilised more than one attitude instrument. Hence, 228 instruments were included in the analysis.

Sixty-six of these instruments (29.1%) were newly developed (see Table 3) and only thirteen of these newly developed instruments were used again in another study. For example, the instrument developed by Cornoldi et al. (1998), was used later by Memisevic and Hodzic (2011) in Bosnia and Herzegovina and by Logan and Wimer (2013) in the United States.

By far most of the instruments in the dataset utilised established scales (70.9%). The Attitudes Towards Inclusive Education Scale (ATIES; Wilczenski, 1992; 1995) was most commonly used, as it was the measurement instrument in 22 studies (which is 9.8% of 228). Second most common was the Opinions Relative to Integration scale (ORI; Antonak & Larrivee, 1995). Thirteen studies (5.7%) utilised this instrument. In ten studies (4.4%), the Sentiments, Attitudes, and Concerns about Inclusive Education Revised scale (SACIE-R; Forlin et al., 2011) was utilised and in nine studies (3.9%), the Opinions Relative to Mainstreaming scale (ORM; Larrivee & Cook, 1979). The original version of the SACIE (Loreman et al., 2007) was used in eight studies (3.5%). In a similar number of studies, the Scale of Teachers' Attitudes Toward Inclusion (STATIC; Cochran, 1997) was used. The My Thinking About Inclusion scale (MTAI; Stoiber et al., 1998) and the Multidimensional Attitudes toward Inclusive Education Scale (MATIES; Mahat, 2008) were the measurement instruments in seven (3.1%) and six studies (2.6%), respectively. The other scales have been used in only five or less studies.

Besides the ORM from 1979, other older scales from the 1970s and 1980s were still in use, too. For example, the Autism Attitude Scale for Teachers (AAST; Olley et al., 1981) was used in recent studies on teachers' attitudes towards inclusive education (Engstrand & Roll-Pettersson, 2014; Garrad et al., 2019; Low et al., 2018). In another recent study (Alghazo et al., 2003), for example, the Attitudes Toward Disabled Persons (Yuker et al., 1966) was utilised.

For 152 instruments in the dataset (66.7%), the wording was available (see Table 3), either directly in the paper (in 79 studies) or it was found in the reference (in 73 studies). For 76 studies (33.3%) no wording was found.

3.2. Analysis of the items' wording

The 152 instruments, for which the wording was available, were analysed further regarding the inclusiveness of the actual wording. Taken together, 2499 items were extracted, and they were analysed using the standardised approach described in the Methods section.

As Table 3 indicates, almost one third of the items (30.3%) included disab*. This term was used in different variations: it was often used to refer to particular students (e.g., "a child with a physical disability") or groups of students (e.g., "students with mild to moderate disabilities"). Examples for items that included the term disab* were "I feel that inclusion provides students with disabilities positive role models" (which was used by Shady et al., 2013) and "I like having children with disabilities in my classroom" (which was used by Tournaki & Samuels, 2016).

The term "special" was relatively common, too. Twenty-eight percent of the items were found using this term in different variations. Special referred to the needs of the child (e.g., "children with

Table 3
Descriptive information.

	Absolute number (n)	Relative number (%)
Descriptive information regarding the 228 instruments		
Utilised instruments		
Established items/scales	161	70.9
New items/scales	66	29.1
(NA)	(1)	–
Wording available		
Not available	76	33.3
Available	152	66.7
Terms utilised in the 2499 items (from 152 scales for which the wording was available)		
Five most common exclusionary terms		
disab	756	30.3
special	698	27.9
integra	116	4.6
typical	75	3.0
mainstream	69	2.8
(…)	(…)	(…)
Five most common inclusionary terms		
inclus	323	12.9
all students	41	1.6
all children	19	0.8
all pupils	6	0.2
for all students	5	0.2
(…)	(…)	(…)
Categorisation of the 2499 items		
Items that comprise exclusive/inclusive terms		
Items with at least one exclusionary term	1769	70.8
Items with at least one inclusionary term	379	15.2
Items' exclusiveness/inclusiveness		
Exclusionary	1574	63.0
Both	195	7.8
Neither	546	21.8
Inclusionary	184	7.4
Inclusiveness of the 152 instruments for which the wording was available		
Fully (100% of inclusive items per instrument)	0	0.0%
High (67–99% of inclusive items per instrument)	2	1.3%
Medium (34–66% of inclusive items per instrument)	9	5.9%
Low (1–33% of inclusive items per instrument)	47	30.9%
Not (0% of inclusive items per instrument)	94	61.8%

special needs"), to the educational setting (e.g., "special physical arrangements", "special education classes", "special schools") or to the teacher (e.g., "special teacher"; "special education teachers"). In less than 5% of the items, terms such as integra*, typical* or mainstream* were used.

Table 3 indicates that inclus* (e.g., the inclusion of all students, but *not* the inclusion of particular students or a group of particular students) was the most common inclusionary term. In 12.9 percent of the items, variations of inclus* were present. Some items referred to the positive and negative effects of inclusion (e.g., "inclusion benefits …"; "inclusion promotes …"; "inclusion is …"). Other items referred to teaching (e.g., "inclusive classroom management"; "inclusive practice"), to the educational setting (e.g., "inclusive settings"; "inclusive classrooms"; "inclusive schools") or to the context (e.g., "inclusive school system"; "inclusive society").

Exclusionary wording occurred in 70.8 percent of the items (see Table 3). Inclusionary wording, as defined by the list of all inclusionary terms, was present in 15.2 percent of the items. As an item can comprise both exclusionary and inclusionary wording at the same time, and as it can comprise neither exclusionary nor inclusionary wording, the combinations of exclusionary and inclusionary wording were examined, too (see Table 3). Solely exclusionary wording can be found in the majority of items (63.0%). An item that was coded as utilising clearly exclusionary wording was for example "students with special needs should be in separate classes" (as used by Everington et al., 1999). A smaller number of items (7.8%) comprised both exclusionary and inclusionary wording at the same time. One third of these combinations stemmed from studies that used the MTAI (Stoiber et al., 1998). The MTAI was based on the definition that inclusion was the "integration of children with and without disabilities" (Stoiber et al., 1998, p. 108). Hence, notions of inclusion and notions regarding disabilities were mixed in the items.

Over one fifth of all items (21.8%) was categorised as neither exclusionary nor inclusionary (abbreviated as 'neither') (see Table 3). Almost half of these items (47.6% relative to the 546 'neither' items) stemmed from studies that utilised the ATIES (Wilczenski, 1992, 1995). Six other items were from the ATMS (Berryman, 1989) and 79 from the SACIE/-R (Forlin et al., 2011; Loreman et al., 2007), which were the precursor and the successor of the ATIES, respectively. All of these items clearly refer to specific groups of students (e.g., "students who need training in self-help skills and activities of daily living"), but the language and terms used in most of these items was not exclusionary per se. Hence, most items of the ATMS, the ATIES and the SACIE(-R) were not exclusionary, but they had no clear inclusionary wording, either. Other items, categorised as 'neither' were very broad, such as "I am ready, to fight for the current school system" (as used by Röhm et al., 2018) or "I enjoyed school and never had any real problem with learning" (as used by Lambe & Bones, 2006).

Table 3 demonstrates that 184 items (7.4%) were coded as being inclusionary. The items stemmed from a range of different studies. Some items were on the effects of inclusion on all students, such as "I believe that inclusion benefits all students academically" (e.g., Miesera et al., 2019) or "I believe that inclusion facilitates socially appropriate behaviour amongst all students." (e.g., Desombre et al., 2018). Other items stated that all students should be able to be part of the regular classroom, such as "the teacher should usually attempt to ensure that all the children in the class, irrespective of levels of difficulty or ability, are able to participate in the class as much as is possible" (e.g., Kraska & Boyle, 2014).

3.3. The inclusiveness of the instruments

While the previous analyses focused solely on single items, the next step analysed how many items per instrument were inclusionary. The analysis of the instruments' inclusiveness demonstrated that the vast majority of the instruments used mostly or completely exclusionary items.

Amongst the 152 instruments, for which the wording was available, 61.8 percent had no inclusive items at all (see Table 3). Thirty-one percent had a low number of inclusive items per instrument. Only 11 instruments (which is less than ten percent of the 152 instruments) had a medium or high proportion of inclusionary items. Nine instruments (5.9%) had a medium percentage of inclusive statements, and two instruments (1.3%) had a high percentage of inclusive items. None of the analysed instruments were fully inclusive.

Amongst these 11 instruments, which were medium or highly inclusive, eight different instruments to measure teachers' attitudes towards inclusive education were present (see upper part of Table 4, first column). Four of these instruments were used in a study for the first time (Hammond & Ingalls, 2003; Hsien et al., 2009; Lambe & Bones, 2006; Shady et al., 2013). The instruments by Lambe and Bones (2006), were utilised in subsequent studies of the authors (Lambe, 2007, 2011; Lambe & Bones, 2007). In the study by Röhm et al. (2018), an instrument, which was developed by Lüke and Grosche (2016), was utilised. Odongo and Davidson (2016) utilised the well-established "School and the Education of All Students Scale" (SEAS; Pearman et al., 1992; Pearman et al., 1997). Two different studies (Miesera et al., 2019; Subban & Mahlo, 2017) utilised the Attitudes to Inclusion Scale (AIS; Sharma & Jacobs, 2016). Yet, both studies utilised a different set of items; hence, in Table 4 "Sharma & Jacobs (1)" and "Sharma & Jacobs (2)" was used in order to differentiate both versions of the AIS.

Table 4 gives an overview on the instruments with a medium and a high proportion of inclusive items per instrument. With 90.0 percent inclusionary statements, Lüke and Grosche (2016; utilised by Röhm et al., 2018) and Hammond and Ingalls (2003) were the most inclusive instruments, followed by Pearman et al. (1992; utilised in Odongo & Davidson, 2016) and Shady et al. (2013) with 60.0 and 58.8 percent inclusive statements, respectively. In the instruments developed by Lambe and Bones (2006; also used in Lambe, 2007, 2011; Lambe & Bones, 2007), Hsien et al. (2009), and Sharma and Jacobs (2016; utilised by Miesera et al., 2019), around half of the items were inclusive. The instrument of Sharma and Jacobs (2016; utilised by Subban & Mahlo, 2017) had 40.0 percent of inclusionary statements.

A closer look at the analyses reveals some differences regarding the philosophical orientation of these instruments. In some of the instruments (Hammond & Ingalls, 2003; Hsien et al., 2009; Pearman et al., 1992; Shady et al., 2013), the terms 'special needs' or 'disabilities' are explicitly used in the non-inclusionary items. According to the present analysis, these instruments focus in part on students with SEND. Thus, these instruments do not fully grasp the philosophy of inclusive education *for all*. One instrument (Sharma & Jacobs, 2016 as utilised by Miesera et al., 2019) comprised notions like students "with a range of abilities", but also statements on particular students "who need assistance with their daily activities" and students "with lower academic ability" (in addition: "with social emotional behaviours" and "with severe disabilities" as utilised in the study by Subban & Mahlo, 2017). This is an attempt not to narrow down the focus on students with SEND, but still to use definite examples of a variety of particular students. In two instruments (Lambe & Bones, 2006; Lüke & Grosche, 2016), the non-inclusionary items were mainly categorised as 'neither' (for example "I enjoyed school and never had any real problem with learning"; Lambe & Bones, 2006). Regarding these items, the instruments tended to be relatively broad in their scope.

As the results demonstrated, the most inclusive instruments

Table 4
Inclusivity of the instruments.

	Analysis of the items' wording			
Instrument(study in which the instrument was utilised)	Exclu. n (%)	Both n (%)	Neither n (%)	Inclu. n (%)
Highly inclusive instruments				
Lüke and Grosche (2016; utilised by Röhm et al., 2018)	0 (0.0%)	0 (0.0%)	2 (10.0%)	18 (90.0%)
Hammond and Ingalls (2003)	0 (0.0%)	1 (10.0%)	0 (0.0%)	9 (90.0%)
Medium inclusive instruments				
Pearman et al. (1992; utilised by Odongo & Davidson, 2016)	3 (30.0%)	0 (0.0%)	1 (10.0%)	6 (60.0%)
Shady et al. (2013)	2 (11.8%)	4 (23.5%)	1 (5.9%)	10 (58.8%)
Sharma and Jacobs (1)* (2016; utilised by Miesera et al., 2019)	0 (0.0%)	0 (0.0%)	4 (50.0%)	4 (50.0%)
Lambe and Bones (2006)	1 (12.5%)	0 (0.0%)	3 (37.5%)	4 (50.0%)
Hsien et al. (2009)	8 (42.1%)	2 (10.5%)	0 (0.0%)	9 (47.4%)
Sharma and Jacobs (2)* (2016; utilised by Subban & Mahlo, 2017)	2 (20.0%)	0 (0.0%)	4 (40.0%)	4 (40.0%)
Most common established instruments; sorted by acronym				
ATIES (Wilczenski, 1992, 1995)	2 (12.5%)	0 (0.0%)	14 (87.5%)	0 (0.0%)
ORI (Antonak & Larrivee, 1995)	25 (100.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
SACIE-R (Forlin et al., 2011)	1 (20.0%)	0 (0.0%)	4 (80.0%)	0 (0.0%)
ORM (Larrivee & Cook, 1979)	29 (96.7%)	0 (0.0%)	1 (3.3%)	0 (0.0%)
SACIE (Loreman et al., 2007)	2 (25.0%)	0 (0.0%)	6 (75.0%)	0 (0.0%)
STATIC (Cochran, 1997)	19 (95.0%)	0 (0.0%)	0 (0.0%)	1 (5.0%)
MTAI (Stoiber et al., 1998)	16 (57.1%)	10 (35.7%)	0 (0.0%)	2 (7.0%)
MATIES (Mahat, 2008)	12 (66.7%)	0 (0.0%)	1 (5.6%)	5 (28.0%)

* The scale by Sharma and Jacobs (2016) was used differently by Miesera et al., (2019; 8 items) and Subban and Mahlo (2017; 10 items).

were not used in many studies. As described above, the following scales were used relatively often: ATIES, ORI, SACIE-R, ORM, SACIE, STATIC, MTAI and MATIES. The lower part of Table 4 indicates how inclusive (*for all*) the scales, which were mostly in use, actually were. First, the large proportion of exclusionary items in the ORI (100.0%), the ORM (96.7%) and the STATIC (95.0%) were noticeable in Table 4. The wording of the items focused on the student or students with SEND. Second, the large proportion of 'neither' items in the ATIES (87.5%), the SACIE-R (80.0%) and the SACIE (75.0%) were noticeable. As mentioned before, these scales were developed from the ATMS (Berryman & Neal, 1980), which used descriptors instead of disability labels, in order to "increase the usefulness of the scale with persons who are not educators of exceptional children" (Berryman et al., 1980, p. 200). That is why the standardised approach to text analysis, which was utilised in the present study, categorised most of these items as being 'neither'. Yet, basically, the items were about the feasibility of particular groups of students being in regular classrooms, which was not in line with the conceptualisation of inclusive education *for all*. The MTAI had some inclusive items (7.0%) and the MATIES even more (28.0%). Yet, the majority of items were exclusionary in both of these scales, too.

The general philosophy of these often-used instruments is deeply rooted in the idea that students with SEND are in the regular classrooms to a greater extent. The ORM, ORI, MTAI, MATIES and the STATIC clearly address the integration of students with SEND in this way. The ATIES, SACIE and the SACIE-R (Forlin et al., 2011) assume that teachers might react differently with respect to different types of SEND. Therefore, these scales differentiate types of SEND. Taken together, the philosophies of the most used instruments were not clearly mirroring a full understanding of inclusive education *for all*.

3.4. The quality of the inclusive instruments

The analysis demonstrated that only a few instruments were available that utilised a medium or high percentage of inclusionary items per instrument. Some of these instruments were new and some of them were already established and replicated. Yet, the question remained to be answered if these instruments were in line with current recommendations for scale development so that it can be assumed that these instruments were of a high quality. Table 5 demonstrated that only the PREIS scale (Lüke & Grosche, 2016, 2018) had fully documented the process of the scale development. After the item generation (using the critical incident technique), content adequacy was established (using experts and think-aloud tests). The analysis utilised up-to-date methodology for factor analysis and construct validation was performed. The study was replicated several times by the authors using a variety of samples (Lüke & Grosche, 2018; Schulze et al., 2019). The SEAS (Pearman et al., 1992) and the AIS (Sharma & Jacobs, 2016) had been developed in a similarly rigorous way. Yet, the content adequacy was not established and/or reported in both cases. The SEAS was correlated to subjects' characteristics in order to validate the measurement, while the AIS was tested within a framework of the Theory of Planned Behaviour (Ajzen, 1991) for validation.

Compared to these three scales (see three top rows in Table 5), the other four instruments seem to lack many steps of the scale development or have at least not reported these steps. A validation was not reported for any of these instruments. Hsien et al. (2009) and Lambe and Bones (2006) established at least content adequacy and the latter study used the items in subsequent studies, yet none of the other steps of scale development were documented. Shady et al. (2013) noted to have adopted the items from a previous qualitative study, yet no other steps of scale development seemed to be carried out. Hammond and Ingalls (2003) have documented not any step of how they developed the instrument. Regarding the item development, Table 5 seems to indicate a clear cut between the scales (PREIS, SEAS and AIS; which have been developed in a particular study and have been used as an established instrument subsequently) and more or less ad-hoc formulated questionnaire items (which were formulated and used in one and the same study).

3.5. Overall valuation of the inclusive instruments

The inclusiveness of the items' wording and the quality of the scale development were two important aspects of attitude measurement instruments in the field of inclusive education. In a last step of analysis these two aspects were analysed together. As can be seen in Fig. 2, the y-axis represents the inclusiveness of an instrument, and the x-axis represents the quality of the scale development. It is clear from the scatterplot that the highly inclusive instrument by Hammond and Ingalls (2003) had not documented the steps of scale development. Three other instruments with a

Table 5
Quality of the inclusive instruments.

Scale reference	Item generation	Content adequacy	Factor analysis	Internal consistency	(Construct) Validation	Replication	No. of steps
Lüke and Grosche (2016)	Critical incident technique	Experts; think-aloud test	CFA	$\alpha = .93$; re-test: $r = .96$	Correlations with other instruments	Several times with different samples by the authors	6
Pearman et al. (1992)	Discussion of themes with relevant stakeholders	NA	PCA	Overall $\alpha = .90$; Subscale 1 $\alpha = .92$; subscale 2 $\alpha = .65$	Correlations with subject characteristics	Replicated several times by others	5
Sharma and Jacobs (2016)	Literature review; themes were written as items	NA	One factor congeneric measurement modelling; CFA	Between $H = .81$ and $H = .90$	SEM using the TPB (Ajzen, 1991) framework	Samples in two countries	5
Lambe and Bones (2006)	NA	Panel of professionals and academics	NA	NA	NA	Items used in subsequent studies	2
Hsien et al. (2009)	NA	Pilot study ($n = 15$)	NA	NA	NA	NA	1
Shady et al. (2013)	Adopted from previous qualitative study	NA	NA	NA	NA	NA	1
Hammond and Ingalls (2003)	NA	NA	NA	NA	NA	NA	0

Note: NA = this information is not available. Lüke and Grosche (2016, 2018): multiple samples (n between 57 and 432); Pearman et al. (1992): primary and secondary in-service teachers and administrative staff ($n = 276$); Sharma and Jacobs (2016): primary and secondary in-service teachers in two countries ($n_{(India)} = 314$, $n_{(Australia)} = 245$); Lambe and Bones (2006): post-primary student teachers ($n = 108$); Hsien et al. (2009): primary and kindergarten in-service teachers ($n = 36$); Shady et al. (2013): primary in-service teachers ($n = 34$); Hammond and Ingalls (2003): elementary school teachers ($n = 343$).

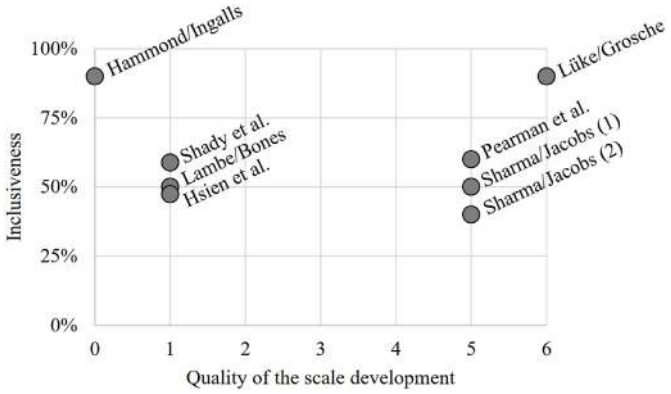

Fig. 2. Scatterplot demonstrating the relationship between inclusiveness and quality of development of the instruments. The inclusiveness (y-axis) represents the percentage of inclusionary items relative to exclusionary items per instrument. The quality of the scale development (x-axis) represents how many steps of the scale development process were documented. In order not to overload the figure, only instruments with medium or high inclusiveness are included in the figure.

medium inclusiveness lacked a documentation of the steps, too (Hsien et al., 2009; Lambe & Bones, 2006; Shady et al., 2013). On the other hand, the SEAS (Pearman et al., 1992, 1997) and the AIS (Sharma & Jacobs, 2016) were based on solid scale development steps, yet these scales had only a medium inclusiveness. Solely the scale by Lüke and Grosche (2016) can be considered both, well-developed and highly inclusive. However, to-date there does not seem to be an attitude scale that is based purely on inclusive education *for all* that is highly developed.

4. Discussion

4.1. Statement of major findings

In the present study, 225 studies with an explicit focus on the measurement of teachers' attitudes towards inclusive education were found through a systematic literature search. Out of these studies, 228 measurement instruments were extracted and for 152 of these instruments, the wording was available. The list of extracted items comprised a total number of 2499 items, which were analysed regarding the inclusiveness or exclusiveness of the items' wording. Taken together, the analyses revealed two major findings.

First of all, the results indicated that a vast majority of instruments that measured the teachers' attitudes towards inclusive education utilised non-inclusive (in accordance to the definition of inclusive education at the beginning of this paper) wording. Only seven percent of all items utilised solely inclusionary wording. In addition, there was not a single instrument in the data set that was 100 percent inclusive, and only eight instruments in eleven studies had a medium or large proportion of inclusionary items per instrument. This finding suggests on the one hand that there are at least some instruments available, which are inclusionary in a rudimentary way. However, on the other hand, the majority of the studies utilised exclusionary instruments to measure teachers' attitudes towards inclusive education *for some*. These results tend to be in line with the finding of Nilholm and Göransson (2017) that empirical research papers often utilise the concept of inclusive education in the sense of integrating particular (groups of) students into the regular classroom. The results of the present study seem to suggest that researchers have not yet overcome the traditional ways of operationalising inclusive attitudes in view of particular students and their placement. The lack of instruments representing inclusive education *for all* leads even researchers, who start from a '*for all*' perspective on inclusive education, to utilise exclusionary instruments and, hence, gain results and conclusions related to inclusive education *for some*. To give an example, Goddard and Evans (2018, p. 122) noted that inclusive education is supposed to mean that schools should cater for the need of all children and that the concept is focused on diversity rather than on disability. In order to capture the teachers' attitudes, the authors developed a survey, including items such as "students with disabilities will improve their social skills when placed in a regular education classroom" or "the best way to meet the needs of the gifted students is to enrol them in special classes and/or schools". In focusing particular groups of students, such as students with SEND or students who are gifted, the items clearly mirror an exclusive

understanding of inclusive education as catering *for some* students who are particularly in need for special attention. Although different exceptional groups of students were addressed in this new instrument (such as students with disabilities or gifted students), the instrument is not representing 'all students'. In the light of the findings of the present study, it makes sense that studies (such as Parasuram, 2006) often find a strong relationship between knowing a person with a disability and the 'inclusive' attitudes. The far majority of studies measured the attitudes towards inclusion in regard to persons with a disability. This gap between a wider conception of inclusive education and an exclusionary operationalisation pertains to most studies, which were analysed in the present study. Hence, this inconsistency seems to underline the conclusion of Ruberg and Porsch (2017), who stated that there is a need for a new instrument which mirrors a wider understanding of inclusion.

A second major finding pertains to the quality of the scales with a medium or high percentage of inclusive items. The results demonstrated that only one scale (with medium or high inclusiveness) has fully documented the process of the scale development. Two other scales have documented most of the scale development process. Yet, the analysis revealed that the development of the other instruments was not documented in a sufficient way. A closer look at these studies revealed that the results sections only comprised percentages of the responses. In these cases, no information was available on how the different items were statistically interrelated and if the items were adequate indicators for one (or more) underlying construct(s). Hence, the quality of the instruments and also the quality of the studies might be of concern, if studies utilised ad-hoc formulated statements and reported only percentages of responses to these statements. However, the use of established scales seemed to be not a guarantee for a high-quality measurement. For example, the PREIS scale was developed clearly to be unidimensional, which was repeatedly tested and documented (see Lüke & Grosche, 2016, 2018; Schulze et al., 2019). Yet, Röhm et al. (2018), who utilised the PREIS, treated the affective, the behavioural, and the cognitive component as separate dimensions. These discrepancies might lead to a flawed measurement and misleading results. Another example for an incautious use of established scales is the utilisation of the AIS (Sharma & Jacobs, 2016) in the studies of Miesera et al. (2019) and Subban and Mahlo (2017). The former study operationalised two latent factors with each four items, which is in line with the text in Sharma and Jacobs (2016). The latter study operationalised one factor of ten items, which is in line with the list of ten items in the appendix of Sharma and Jacobs (2016). Studies that develop new scales might need to be clearer, which exact operationalisation should be tested in further studies.

4.2. Interpretation of the results

Taken together, three issues became apparent through the presented analyses regarding the inclusiveness of instruments to measure teachers' attitudes towards inclusive education for all. First, many scales, which were utilised in recent studies on the teachers' attitudes towards inclusive education, were rooted within a philosophy of mainstreaming or integration. The results showed that, for example, the ORM (Larrivee & Cook, 1979) was utilised in a number of studies. This instrument was developed to fit the discourse in the United States in the 1970s. The instrument's underlying philosophy of mainstreaming tends not to be compatible with conceptualisations of inclusive education *for all*. The utilisation of scales that clearly attempt to measure attitudes towards mainstreaming or integration might not be adequate indicators for attitudes towards inclusive education *for all*.

Second, researchers adapted mainstreaming or integration instruments and updated the wording according to political correctness of the labels; yet the underlying philosophy of the items remained exclusive. Regarding further developments of the attitude measurement, Antonak and Livneh (1988) recommended that scales should be periodically refined rather than newly created. In this way, the ORI (Antonak & Larrivee, 1995), for example, is an update of the ORM. Yet, while the ORI is 100% exclusionary, the original ORM was 96.7% exclusionary. Hence, the update in the 1990s focused the scale even more on the integration of students with SEND, which made the scale even more exclusive in regard to attitudes towards inclusive education *for all*. Another example is the ATMS (Berryman et al., 1980; Berryman & Neal, 1980). The items ask about the feasibility of groups of students being in regular rather than special classrooms, which is not in line with inclusive education *for all*. The ATIES (Wilczenski, 1992, 1995), the SACIE (Loreman et al., 2007), and the SACIE-R (Forlin et al., 2011) are popular further developments of the ATMS. Yet, the underlying philosophy and large parts of the items' wording did not change in the ATIES and the SACIE(-R): Items from all of these scales assume that there are groups of students (e.g., those who require communicative technologies), who might be better taught in regular classrooms or in separate classrooms. The exclusive philosophy, and the idea that there are groups of students that are different compared to what is normal, is deeply inscribed in the statements. Hence, updates of instruments might be less promising to move steps forward towards an adequate operationalisation of attitudes towards inclusive education *for all*.

Third, most of the new scales were also found to be exclusionary. Saloviita (2015) for example mentioned that "positive attitudes towards inclusive education means accepting *all children* in mainstream classrooms […]" (p. 68; emphasis added). Yet, Saloviita (2015) also points out explicitly that children with SEND should be included in mainstream classrooms. The philosophy of this instrument is that exceptional students should be integrated. Accordingly, the results from the present study found that the newly developed instrument by Saloviita (2015) was 100% exclusionary.

There seems to be major confusion amongst empirical researchers regarding the concept and the philosophy of inclusion. Lübke et al. (2019), for example, refer to the difference between a broad and a narrow understanding of inclusion: a narrow definition refers to students with SEND being in the regular classroom, while a broad definition, refers to students with SEND being full members of the classrooms. However, both definitions represent a view towards a particular group of exceptional students. None of these definitions refer to inclusive education as the vision of catering *for all* students. In other words, the confusion regarding the concept of inclusion amongst empirical researchers might lead to decisions to choose or develop scales or items with exclusionary wording, rather than taking the opportunity to make some serious progress regarding the measurement of attitudes towards inclusive education *for all*.

4.3. Limitations

The current study examined a large body of references, instruments, and items, by using a standardised approach. Hence, the major findings were based to a large extent on the specification of exclusive and inclusive terms. Although these terms were derived carefully in cycles of repeated critical reflection of adequate terms by both authors, these terms do not fully replace a qualitative in-depth examination of the items' content. The quantitative approach, as it was used in the present study, was not able to capture the meaning of the items.

Another limitation pertains to the large variety regarding the quality of the analysed papers, which lead to difficulties to extract the relevant information in a systematic and standardised way. Normally, systematic reviews use a quality check relatively early in the procedures, which allows to disregard papers with low quality (see Gough et al., 2017; Petticrew & Roberts, 2006). As the present study attempted to capture the instruments the way they were used in the field and published in journals, none of the studies were excluded because of lack of quality.

Another limitation pertains to an inconsistent use within the analysed studies regarding attitudes towards inclusive education a) being a unidimensional latent construct, b) comprising multiple subdimensions or c) being a sub-dimension itself. The former two are represented by the above-mentioned inconsistency between the unidimensional scale by Lüke and Grosche (2016), which was used by Röhm et al. (2018) as having three dimensions. The latter is, for example, represented by the SACIE scale, which puts sentiments, attitudes, and concerns right next to each other as adjacent sub-dimensions of a diffuse underlying construct, combining sentiments, attitudes and concerns. Forlin et al. (2011) even reported Cronbachs Alpha for the 'combined SACIE scale' suggesting that sentiments, attitudes and concerns are expressions of one underlying latent construct. For the present study, the SACIE(-R) for example was only analysed regarding the attitudes dimension. In other cases, it was difficult to make clear conceivable decisions, too.

The documentation of the scale development steps was taken as a proxy for the scales' quality. There are good reasons to doubt the scientific rigor of a study, if the core information regarding the measurement is missing. Yet, a missing documentation does not necessarily mean that the scale quality is actually weak. Hence, this quality indicator of the instruments might need to be interpreted with some caution.

4.4. Implications

The findings of the current study show that there is a lack of full inclusionary attitude instruments. This has implications for research, policy, and practice.

First, there is an urgent need for *future research*, which conceptualise inclusive education as being *for all* students and which operationalise the attitudes towards inclusive education *for all*. This research might benefit from distinguishing items pertaining to access, presence, participation, and success of all students. An intensified dialogue between researchers who already advocate inclusive education *for all* from a theoretical or conceptual perspective and those researchers who are specialists for conducting empirical studies might help in this regard. In addition, future studies in the field of inclusive education *for all* need to establish high quality standards for conducting empirical studies. The research community will need to be supportive in this regard (e.g., by critical peer reviews).

Second, the findings have implications for *policies*. According to the results of the present study, nearly all empirical studies on teachers' attitudes towards inclusive education are deeply rooted in the framework of inclusive education *for some*. Policy makers need to be aware of this fact. As research shows, policy development often lacks attention to promoting genuinely inclusive educational practices (Hardy & Woodcock, 2015). If policy development was guided by research which re-iterates ideas of inclusive education *for some*, it would seem unlikely that the resulting policy will address *all* students.

Third, the present study also has implications for *practice*. The large number of surveys on inclusion, which teachers are asked to complete, might affect what teachers believe inclusion is about. Questionnaires, that claim to be on inclusion, that ask questions pertaining to teaching students with SEND in regular classrooms might make it difficult for teachers to think beyond inclusion as being the placement of particular (groups of) students (e.g., with SEND) in regular classes. This study is a reminder that teachers in the inclusive classroom need to value each individual student as an individual learner with certain strengths and capabilities. Inclusive education *for all*, which was defined and consistently used in the present study, might support teachers to understand that current thinking about inclusion is likely to even increase exclusiveness in schools and classrooms (Slee, 2013). In this way, the notion of inclusive education *for all* might support school leaders in developing such a vision and mission in their school that is compatible with inclusive values. Therefore, it would be important for school leaders to be able to examine the attitudes of their teachers towards inclusive education *for all* students. Hence, there is an urgent need for the development of a new and accurate attitude measurement instrument that is able to capture how teachers think about inclusive education *for all* students.

5. Conclusion

Taken together, the findings of the present study suggest that the idea needs to be jettisoned that mainstreaming and integration instruments just need refinements regarding politically correct labels for SEND. Even newly developed instruments do not seem to be indicative of a current conceptualisation of inclusive education *for all*. The present study's results indicate that serious new attempts are needed to operationalise and capture the attitudes towards inclusive education *for all*.

Credit author statement

Stephan Kielblock: Conceptualisation, Methodology, Validation, Formal analysis, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization. **Stuart Woodcock**: Conceptualisation, Methodology, Validation, Writing - Review & Editing, Visualization.

Data availability

We have shared our research data as Supplementary Material/ Appendix.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.tate.2022.103922.

References

Ainscow, M., Booth, T., & Dyson, A. (2006). *Improving schools, developing inclusion*. London: Routledge.

Ajzen, I. (1988). *Attitudes, personality, and behavior*. Milton Keynes: Open University Press.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

Alghazo, E. M., Doreen, H., & Alqaryouti, I. A. (2003). Attitudes of pre-service teachers towards persons with disabilities: Predictions for the success of inclusion. *College Student Journal, 37*(4), 515–522.

Amor, A. M., Hagiwara, M., Shogren, K. A., Thompson, J. R., Verdugo, M.Á., Burke, K. M., & Aguayo, V. (2019). International perspectives and trends in research on inclusive education: A systematic review. *International Journal of Inclusive Education, 23*(12), 1277–1295. https://doi.org/10.1080/13603116.2018.1445304

Antonak, R. F., & Larrivee, B. (1995). Psychometric analysis and revision of the opinions relative to mainstreaming scale. *Exceptional Children, 62*(2), 139–149. https://doi.org/10.1177/001440299506200204

Antonak, R. F., & Livneh, H. (1988). *The measurement of attitudes toward people with*

disabilities: Methods, psychometrics, and scales. Springfield, IL: Charles C. Thomas.
Avramidis, E., & Norwich, B. (2002). Teachers' attitudes towards integration/inclusion: A review of the literature. *European Journal of Special Needs Education, 17*(2), 129–147. https://doi.org/10.1080/08856250210129056
Berryman, J. D. (1989). Attitudes of the public toward educational mainstreaming. *Remedial and Special Education, 10*(4), 44–49. https://doi.org/10.1177/074193258901000408
Berryman, J. D., & Neal, W. R. (1980). The cross validation of the attitudes toward mainstreaming scale (ATMS). *Educational and Psychological Measurement, 40*(2), 469–474. https://doi.org/10.1177/001316448004000227
Berryman, J. D., Neal, W. R., & Robinson, J. E. (1980). The validation of a scale to measure attitudes toward the classroom integration of disabled students. *The Journal of Educational Research, 73*(4), 199–203. https://doi.org/10.1080/00220671.1980.10885235
de Boer, A., Pijl, S. J., & Minnaert, A. (2011). Regular primary schoolteachers' attitudes towards inclusive education: A review of the literature. *International Journal of Inclusive Education, 15*(3), 331–353. https://doi.org/10.1080/13603110903030089
Booth, T. (1995). Mapping inclusion and exclusion: Concepts for all? In C. Clark, A. Dyson, & A. Millward (Eds.), *Towards inclusive schools?* (pp. 96–108). London: David Fulton.
Booth, T., & Ainscow, M. (2002). *Index for inclusion. Developing learning and participation in schools*. Bristol: Centre for Studies on Inclusive Education.
Burnett, N. (2008). *What sort of UNESCO for what sort of education: Gaitskell Lecture*. University of Nottingham 22 May 2008. Retrieved from.
Carpenter, S. (2018). Ten steps in scale development and reporting: A guide for researchers. *Communication Methods and Measures, 12*(1), 25–44. https://doi.org/10.1080/19312458.2017.1396583
Carrington, S., MacArthur, J., Kearney, A., Kimber, M., Mercer, L., Morton, M., & Rutherford, G. (2012). Towards an inclusive education for all. In S. Carrington, & J. Macarthur (Eds.), *Teaching in inclusive school communities* (pp. 3–38). Milton: John Wiley & Sons.
Chaiklin, H. (2011). Attitudes, behavior, and social practice. *Journal of Sociology & Social Welfare, 38*(1), 31–54.
Cochran, H. K. (1997). The development and psychometric analysis of the scale of teachers' attitudes toward inclusion (STATIC). Retrieved from http://files.eric.ed.gov/fulltext/ED415259.pdf.
Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966). *Equality of educational opportunity*. Washington, D.C.: U.S. Government Printing Office.
Cornoldi, C., Terreni, A., Scruggs, T. E., & Mastropieri, M. A. (1998). Teacher attitudes in Italy after twenty years of inclusion. *Remedial and Special Education, 19*(6), 350–356. https://doi.org/10.1177/074193259801900605
Desombre, C., Lamotte, M., & Jury, M. (2018). French teachers' general attitude toward inclusion: The indirect effect of teacher efficacy. *Educational Psychology, 39*(1), 38–50. https://doi.org/10.1080/01443410.2018.1472219
DeVellis, R. F. (2011). *Scale development: Theory and applications* (Vol. 3). Thousand Oaks, CA: Sage.
Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Orlando, FL: Harcourt Brace Jovanovich.
Engstrand, R. Z., & Roll-Pettersson, L. (2014). Inclusion of preschool children with autism in Sweden: Attitudes and perceived efficacy of preschool teachers. *Journal of Research in Special Educational Needs, 14*(3), 170–179. https://doi.org/10.1111/j.1471-3802.2012.01252.x
Evans, J., & Benefield, P. (2001). Systematic reviews of education research: Does the medical model fit? *British Educational Research Journal, 27*(5), 527–541. https://doi.org/10.1080/01411920120095717
Everington, C., Stevens, B., & Winters, V. B. (1999). Teachers' attitudes, felt competence, and need of support for implementation of inclusive educational programs. *Psychological Reports, 85*(1), 331–338. https://doi.org/10.2466/pr0.1999.85.1.331
Ewing, D. L., Monsen, J. J., & Kielblock, S. (2018). Teachers' attitudes towards inclusive education: A critical review of published questionnaires. *Educational Psychology in Practice, 34*(2), 150–165. https://doi.org/10.1080/02667363.2017.1417822
Forlin, C. (1995). Educators' beliefs about inclusive practices in Western Australia. *British Journal of Special Education, 22*(4), 179–185. https://doi.org/10.1111/j.1467-8578.1995.tb00932.x
Forlin, C., Earle, C., Loreman, T., & Sharma, U. (2011). The Sentiments, Attitudes, and Concerns about Inclusive Education Revised (SACIE-R) scale for measuring preservice perceptions teachers' perception about inclusion. *Exceptionality Education International, 21*(3), 50–65.
Garrad, T.-A., Rayner, C., & Pedersen, S. (2019). Attitudes of Australian primary school teachers towards the inclusion of students with autism spectrum disorders. *Journal of Research in Special Educational Needs, 19*(1), 58–67. https://doi.org/10.1111/1471-3802.12424
Glasman, L. R., & Albarracín, D. (2006). Forming attitudes that predict future behavior: A meta-analysis of the attitude-behavior relation. *Psychological Bulletin, 132*(5), 778–822. https://doi.org/10.1037/0033-2909.132.5.778
Goddard, C., & Evans, D. (2018). Primary pre-service teachers' attitudes towards inclusion across the training years. *Australian Journal of Teacher Education, 43*(6), 122–142. https://doi.org/10.14221/ajte.2018v43n6.8
Goldhaber, D. (2016). In schools, teacher quality matters most. *Education, 16*(2), 56–62.
Gough, D., Oliver, S., & Thomas, J. (2017). *An introduction to systematic reviews*. Thousand Oaks, CA: Sage.
Graham, L. J., & Macartney, B. (2012). Naming or creating a problem? The mis/use of labels in schools. In S. Carrington, & J. Macarthur (Eds.), *Teaching in inclusive school communities* (pp. 189–208). Milton: John Wiley & Sons.
Graham, L. J., & Slee, R. (2008). An illusory interiority: Interrogating the discourse/s of inclusion. *Educational Philosophy and Theory, 40*(2), 277–293. https://doi.org/10.1111/j.1469-5812.2007.00331.x
Hammond, H., & Ingalls, L. (2003). Teachers' attitudes toward inclusion: Survey results from elementary school teachers in three southwestern rural school districts. *Rural Special Education Quarterly, 22*(2), 24–30. https://doi.org/10.1177/875687050302200204
Hardy, I., & Woodcock, S. (2015). Inclusive education policies: Discourses of difference, diversity and deficit. *International Journal of Inclusive Education, 19*(2), 141–164. https://doi.org/10.1080/13603116.2014.908965
Hattie, J. A. C. (2003). Teachers make a difference, what is the research evidence?. In *Paper presented at the Building teacher quality: What does the research tell us, ACER research conference*. Melbourne, Australia http://research.acer.edu.au/research_conference_2003/4/.
Hattie, J. A. C. (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.
Hinkin, T. R., Tracey, J. B., & Enz, C. A. (1997). Scale construction: Developing reliable and valid measurement instruments. *Journal of Hospitality & Tourism Research, 21*(1), 100–120. https://doi.org/10.1177/109634809702100108
Hitlin, S., & Pinkston, K. (2013). Values, attitudes, and ideologies: Explicit and implicit constructs shaping perception and action. In J. DeLamater, & A. Ward (Eds.), *Handbook of social psychology* (pp. 319–339). Dordrecht: Springer.
Hsien, M., Brown, P. M., & Bortoli, A. (2009). Teacher qualifications and attitudes toward inclusion. *Australasian Journal of Special Education, 33*(1), 26–41. https://doi.org/10.1375/ajse.33.1.26
Kiuppis, F. (2014). Why (not) associate the principle of inclusion with disability? Tracing connections from the start of the 'Salamanca process'. *International Journal of Inclusive Education, 18*(7), 746–761. https://doi.org/10.1080/13603116.2013.826289
Kraska, J., & Boyle, C. (2014). Attitudes of preschool and primary school pre-service teachers towards inclusive education. *Asia-Pacific Journal of Teacher Education, 42*(3), 228–246. https://doi.org/10.1080/1359866X.2014.926307
Lambe, J. (2007). Northern Ireland student teachers' changing attitudes towards inclusive education during initial teacher training. *International Journal of Special Education, 22*(1), 59–71.
Lambe, J. (2011). Pre-service education and attitudes towards inclusion: The role of the teacher educator within a permeated teaching model. *International Journal of Inclusive Education, 15*(9), 975–999. https://doi.org/10.1080/13603110903490705
Lambe, J., & Bones, R. (2006). Student teachers' attitudes to inclusion: Implications for initial teacher education in northern Ireland. *International Journal of Inclusive Education, 10*(6), 511–527. https://doi.org/10.1080/13603110500173225
Lambe, J., & Bones, R. (2007). The effect of school-based practice on student teachers' attitudes towards inclusive education in Northern Ireland. *Journal of Education for Teaching, 33*(1), 99–113. https://doi.org/10.1080/02607470601098369
Larrivee, B., & Cook, L. (1979). Mainstreaming: A study of the variables affecting teacher attitude. *The Journal of Special Education, 13*(3), 315–324. https://doi.org/10.1177/002246697901300310
Logan, B. E., & Wimer, G. (2013). Tracing inclusion: Determining teacher attitudes. *Research in Higher Education Journal, 20*, 1–10.
Loreman, T., Earle, C., Sharma, U., & Forlin, C. (2007). The development of an instrument for measuring pre-service teachers' sentiments, attitudes, and concerns about inclusive education. *International Journal of Special Education, 22*(2), 150–159.
Low, H. M., Lee, L. W., & Che Ahmad, A. (2018). Pre-service teachers' attitude towards inclusive education for students with Autism Spectrum Disorder in Malaysia. *International Journal of Inclusive Education, 22*(3), 235–251. https://doi.org/10.1080/13603116.2017.1362479
Lübke, L., Pinquart, M., & Schwinger, M. (2019). How to measure teachers' attitudes towards inclusion: Evaluation and validation of the differentiated attitudes towards inclusion scale (DATIS). *European Journal of Special Needs Education, 34*(3), 297–311. https://doi.org/10.1080/08856257.2018.1479953
Lüke, T., & Grosche, M. (2016). *Professionsunabhängige einstellungsskala zum inklusiven schulsystem (PREIS)*. https://doi.org/10.6084/m9.figshare.2245630
Lüke, T., & Grosche, M. (2018). Konstruktion und Validierung der Professionsunabhängigen Einstellungsskala zum Inklusiven Schulsystem (PREIS). *Empirische Sonderpädagogik, 10*(1), 3–20.
Mahat, M. (2008). The development of a psychometrically-sound instrument to measure teachers' multidimensional attitudes toward inclusive education. *International Journal of Special Education, 23*(1), 82–92.
Mansfield, R. K. (2015). Teacher quality and student inequality. *Journal of Labor Economics, 33*(3), 751–788. https://doi.org/10.1086/679683
Memisevic, H., & Hodzic, S. (2011). Teachers' attitudes towards inclusion of students with intellectual disability in Bosnia and Herzegovina. *International Journal of Inclusive Education, 15*(7), 699–710. https://doi.org/10.1080/13603110903184001
Messiou, K. (2017). Research in the field of inclusive education: Time for a rethink? *International Journal of Inclusive Education, 21*(2), 146–159. https://doi.org/10.1080/13603116.2016.1223184

Miesera, S., DeVries, J. M., Jungjohann, J., & Gebhardt, M. (2019). Correlation between attitudes, concerns, self-efficacy and teaching intentions in inclusive education evidence from German pre-service teachers using international scales. *Journal of Research in Special Educational Needs, 19*(2), 103–114. https://doi.org/10.1111/1471-3802.12432

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine, 6*(7), Article e1000097. https://doi.org/10.1371/journal.pmed.1000097

Mundy, K. (2016). Leaning in" on education for all. *Comparative Education Review, 60*(1), 1–26. https://doi.org/10.1086/684434

Nilholm, C., & Göransson, K. (2017). What is meant by inclusion? An analysis of European and North American journal articles with high impact. *European Journal of Special Needs Education, 32*(3), 437–451. https://doi.org/10.1080/08856257.2017.1295638

Odongo, G., & Davidson, R. (2016). Examining the attitudes and concerns of the Kenyan teachers toward the inclusion of children with disabilities in the general education classroom: A mixed methods study. *International Journal of Special Education, 31*(2), 1–30.

OECD. (2005). *Teachers matter: Attracting, developing and retaining effective teachers*. Paris: OECD.

Olley, J. G., Devellis, R. F., Devellis, B. M. E., Wall, A. J., & Long, C. E. (1981). The autism attitude scale for teachers. *Exceptional Children, 47*(5), 371–372. https://doi.org/10.1177/001440298104700509

Parasuram, K. (2006). Variables that affect teachers' attitudes towards disability and inclusive education in Mumbai, India. *Disability & Society, 21*(3), 231–242. https://doi.org/10.1080/09687590600617352

Pearman, E. L., Barnhart, M. W., Huang, A. M., & Mellblom, C. (1992). Educating all students in school: Attitudes and beliefs about inclusion. *Education and Training in Mental Retardation, 27*(2), 176–182.

Pearman, E. L., Huang, A. M., & Mellblom, C. I. (1997). The inclusion of all students: Concerns and incentives of educators. *Education and Training in Mental Retardation and Developmental Disabilities, 32*(1), 11–20.

Peters, S. J. (2004). *Inclusive Education: An EFA strategy for all children*. Washington, D.C.: World Bank.

Peters, S. J. (2007). Education for All?" A historical analysis of international inclusive education policy and individuals with disabilities. *Journal of Disability Policy Studies, 18*(2), 98–108. https://doi.org/10.1177/10442073070180020601

Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences: A practical guide*. Malden, MA: Blackwell Publishing.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica, 73*(2), 417–458. https://doi.org/10.1111/j.1468-0262.2005.00584.x

Röhm, A., Schnöring, A., & Hastall, M. R. (2018). Impact of single-case pupil descriptions on student teacher attitudes towards inclusive education. *Learning Disabilities: A Contemporary Journal, 16*(1), 37–58.

Rosenberg, M. J., & Hovland, C. I. (1960). Cognitive, affective, and behavioral components of attitudes. In M. J. Rosenberg, C. I. Hovland, W. J. McGuire, R. P. Abelson, & J. W. Brehm (Eds.), *Attitude organization and change: An analysis of consistency among attitude components* (pp. 1–14). New Haven, CT: Yale University Press.

Ruberg, C., & Porsch, R. (2017). Einstellungen von Lehramtsstudierenden und lehrkräften zur schulischen inklusion: Ein systematisches review deutschsprachiger forschungsarbeiten. [Attitudes of pre-service and in-service teachers towards school inclusion: A systematic review of German studies]. *Zeitschrift für Padagogik, 63*(4), 393–415.

Ruijs, N. M., & Peetsma, T. T. D. (2009). Effects of inclusion on students with and without special educational needs reviewed. *Educational Research Review, 4*(2), 67–79. https://doi.org/10.1016/j.edurev.2009.02.002

Saloviita, T. (2015). Measuring pre-service teachers' attitudes towards inclusive education: Psychometric properties of the TAIS scale. *Teaching and Teacher Education, 52*, 66–72. https://doi.org/10.1016/j.tate.2015.09.003

Schulze, S., Lüke, T., Schröter, A., Krause, K., & Kuhl, J. (2019). Replikationsstudie zur Testgüte der Professionsunabhängigen Einstellungsskala zum Inklusiven Schulsystem – ein Beitrag zur Validierung. *Unterrichtswissenschaft, 47*(2), 201–219. https://doi.org/10.1007/s42010-018-00034-3

Shady, S. A., Luther, V. L., & Richman, L. J. (2013). Teaching the teachers: A study of perceived professional development needs of educators to enhance positive attitudes toward inclusive practices. *Education Research & Perspectives, 40*(1), 169–191.

Sharma, U., & Jacobs, K. (2016). Predicting in-service educators' intentions to teach in inclusive classrooms in India and Australia. *Teaching and Teacher Education, 55*, 13–23. https://doi.org/10.1016/j.tate.2015.12.004

Shyman, E. (2015). Toward a globally sensitive definition of inclusive education based in social justice. *International Journal of Disability, Development and Education, 62*(4), 351–362. https://doi.org/10.1080/1034912X.2015.1025715

Slee, R. (2013). How do we make inclusive education happen when exclusion is a political predisposition? *International Journal of Inclusive Education, 17*(8), 895–907. https://doi.org/10.1080/13603116.2011.602534

Slee, R. (2018). *Defining the scope of inclusive education: Think piece prepared for the 2020 global education monitoring report, inclusion and education*. Retrieved from https://unesdoc.unesco.org/ark:/48223/pf0000265773.

Stoiber, K. C., Gettinger, M., & Goetz, D. (1998). Exploring factors influencing parents' and early childhood practitioners' beliefs about inclusion. *Early Childhood Research Quarterly, 13*(1), 107–124. https://doi.org/10.1016/S0885-2006(99)80028-3

Subban, P., & Mahlo, D. (2017). 'My attitude, my responsibility' Investigating the attitudes and intentions of pre-service teachers toward inclusive education between teacher preparation cohorts in Melbourne and Pretoria. *International Journal of Inclusive Education, 21*(4), 441–461. https://doi.org/10.1080/13603116.2016.1197322

Szumski, G., Smogorzewska, J., & Karwowski, M. (2017). Academic achievement of students without special educational needs in inclusive classrooms: A meta-analysis. *Educational Research Review, 21*, 33–54. https://doi.org/10.1016/j.edurev.2017.02.004

Tant, M., & Watelain, E. (2016). Forty years later, a systematic literature review on inclusion in physical education (1975–2015): A teacher perspective. *Educational Research Review, 19*, 1–17. https://doi.org/10.1016/j.edurev.2016.04.002

Thomas, G. (2013). A review of thinking and research about inclusive education policy, with suggestions for a new kind of inclusive thinking. *British Educational Research Journal, 39*(3), 473–490. https://doi.org/10.1080/01411926.2011.652070

Tournaki, N., & Samuels, W. E. (2016). Do graduate teacher education programs change teachers' attitudes toward inclusion and efficacy beliefs? *Action in Teacher Education, 38*(4), 384–398. https://doi.org/10.1080/01626620.2016.1226200

Triandis, H. C. (1971). *Attitude and attitude change*. New York: John Wiley & Sons.

Turner, N. D. (1995). Two approaches to the field experience and the effects on attitudes of preservice teachers towards inclusion. *The Teacher Educator, 31*(1), 96–105. https://doi.org/10.1080/08878739509555102

UNESCO. (1990). *World declaration on Education for All*. Paris: UNESCO.

UNESCO. (1994). *The Salamanca statement*. Paris: UNESCO.

UNESCO. (2000). *The Dakar Framework for Action*. Paris: UNESCO.

UNESCO. (2005). *Guidelines for inclusion: Ensuring access to education for all*. Paris: UNESCO.

UNESCO. (2015). *Incheon declaration. Education 2030: Towards inclusive and equitable quality education and lifelong learning for all*. Paris: UNESCO.

UNESCO, & UNDP. (1995). *The education for all teacher-training package (vol. 1)*. United Kingdom: UNESCO.

Van Mieghem, A., Verschueren, K., Petry, K., & Struyf, E. (2020). An analysis of research on inclusive education: A systematic search and meta review. *International Journal of Inclusive Education, 24*(6), 675–689. https://doi.org/10.1080/13603116.2018.1482012

Voas, D. (2014). Towards a sociology of attitudes. *Sociological Research Online, 19*(1). https://doi.org/10.5153/sro.3289

Wilczenski, F. L. (1992). Measuring attitudes toward inclusive education. *Psychology in the Schools, 29*(4), 306–312. https://doi.org/10.1002/1520-6807(199210)29:4<306::aid-pits2310290403>3.0.co

Wilczenski, F. L. (1995). Development of a scale to measure attitudes toward inclusive education. *Educational and Psychological Measurement, 55*(2), 291–299. https://doi.org/10.1177/0013164495055002013

Yuker, H. E., Block, J. R., & Younng, J. H. (1966). The measurement of attitudes toward disabled persons. Retrieved from http://files.eric.ed.gov/fulltext/ED044853.pdf.