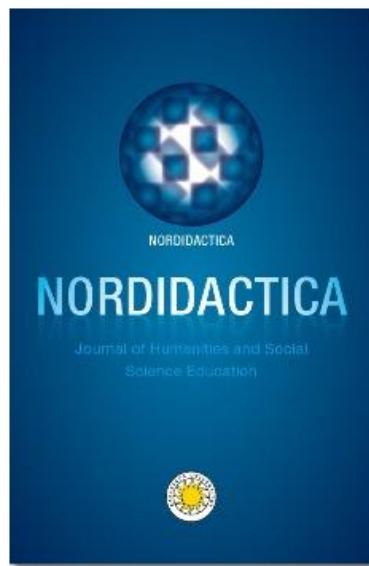


Assessing Intercultural Competence through Swedish National Tests of Social Science: A Validation Study

Markus Al-Afifi



Nordidactica

- Journal of Humanities and Social Science Education

2025:2

Nordidactica – Journal of Humanities and Social Science Education

Nordidactica 2025:2

ISSN 2000-9879

The online version of this paper can be found at: www.kau.se/nordidactica

Assessing Intercultural Competence through Swedish National Tests of Social Science: A Validation Study

<https://doi.org/10.62902/nordidactica.v15i2025:2.26748>

Markus Al-Afifi

Department of Education, Uppsala University

Abstract: Despite concerted efforts of international organisations to advocate for the integration of intercultural competence (IC) into all facets of education, IC remains under-theorised in large-scale and social science assessments. By utilising the Swedish National Tests of Social Science (NTSS) as a resource for evaluating students' IC, this study examines all available NTSS test items to provide an exhaustive review of the IC aspects addressed in NTSS content over its first decade of existence. In total, 210 NTSS test items were used from 2013 to 2022, of which 46 were deemed interculturally relevant. The results indicate that the NTSS capture several aspects of global knowledge, social knowledge, and perspective-taking. In contrast, there has been limited testing of cultural knowledge, reflexivity, and respect, suggesting variability in the NTSS' coverage of the IC construct. The findings demonstrate NTSS's utility beyond its primary functions as an external grading instrument of social science, offering insights into its assessment capacity, which transcends the assessment parameters drawn by the Swedish National Education Agency.

KEYWORDS: INTERCULTURAL COMPETENCIES, NATIONAL TEST, SOCIAL SCIENCE, VALIDITY, CONSTRUCT

About the author: Markus Al-Afifi is a PhD Candidate in Social Science Education Didactics at the Department of Education, Uppsala University. His research focuses on interculturalism in social science education, students' attitudes towards immigrants, and international large-scale assessments.

Various IGOs and NGOs consider intercultural competence (IC) a desirable learning objective (e.g., CofE, 2008; UNESCO, 2018; OECD, 2018). IC refers to a multidimensional construct encompassing a range of knowledge, skills, and attitudes required to orient oneself effectively in an increasingly globalised local context (Fantini, 2007; Perry & Southwell, 2011). The Council of Europe (2016) highlights its democratic importance, citing intercultural dialogue—a concept closely related to IC—as intrinsic to safeguarding democracy. Because of its societal and democratic utility, international policymakers have recommended that IC be included in national curricula at all levels of education (Leeds-Herwitz, 2013; Schauer, 2020).

However, the goal of IC to become fully institutionalised in national educational systems is far from realised. Due to its amorphous nature, the theoretical composition of IC has developed to be extensive, causing a lack of consensus on its meaning (Zhang & Zhou, 2019). As a result, the increasing multidimensionality of IC poses challenges in its assessment since the construct does not lend itself easily to standardised assessment tools. More precisely, a fundamental issue of IC assessments in education is the lack of subject-specific conceptual frameworks. Without such a framework, the ability to monitor the development of students' learning is inhibited. Compounding this issue, countries like Sweden have chosen not to participate in previous large-scale IC assessments¹, perhaps reflecting the conceptual ambiguity surrounding this concept. Deardorff and colleagues (2006) encourage researchers to develop IC frameworks specified for the given educational context to strengthen IC assessments. While language-subject researchers have responded to these calls for IC frameworks (Borghetti, 2017; Luo & Chan, 2022), frameworks for other interculturally relevant school subjects, including social science education², are underdeveloped. Consequently, the assessment possibilities for students' IC in social science education (SSE) remain limited, both preventing concordance with international policy initiatives for promoting IC and limiting insight into Swedish students' IC.

These challenges highlight the need for a subject-specific assessment to strengthen the evaluation of students' IC in SSE, which would bridge the gap in understanding students' abilities to consider intercultural issues. Given Sweden's decision not to participate in the PISA 2018 Global Competence Study, which aimed to measure IC, there is a pressing need to explore alternative ways to assess IC within the national context. One alternative pathway to solving this problem is repurposing existing assessments embedded in the Swedish educational system to align Swedish and international educational objectives and allow insight into Swedish students' IC, which is the key ambition of this study.

¹According to Cobb and Couch (2021), the PISA 2018 Global Competence Study aimed to assess key aspects of IC. While Sweden participated in PISA 2018, they were one of the few participating countries that opted out of partaking in the global competency test's cognitive and attitude assessments.

² In their country report, Sandahl et al. (2022) suggested that the subject of *Samhällskunskap* should be translated into *social science education*. This translation is employed in this study to ensure consistency in domain terminology.

In the current paper, I sought to examine how an existing large-scale assessment of SSE can be used as a supplementary resource to assess students' IC. In particular, the Swedish National Test of Social Science (NTSS) for grade 9 was examined, as this assessment serves as a pivotal tool for evaluating students' understanding of interculturally relevant issues related to equality, globalisation, migration, discrimination, segregation, national minorities, and human rights. Whereas the purpose of NTSS is broader than assessing IC, I investigate the possibility of evaluating IC through the NTSS. By reframing the construct the test intends to measure, I display NTSS utility beyond its primary functions as an external grading instrument of SSE. Thus, in response to international calls for intercultural education, this study contributes to understanding how national tests in Sweden, specifically the NTSS, afford to assess IC.

Aims and Research Questions

This research aims to enrich the current understanding of how large-scale assessments in social science can be used to assess students' IC. To achieve this aim, the study examines the NTSS test items in relation to a specification of the IC construct through a validation approach. By eliciting and evaluating validity evidence of *construct comprehensiveness*, the analysis provides an overview of the strengths and limitations of NTSS for assessing students' IC regarding the studied psychometric property. To enable this analysis, I developed a new conceptual framework derived from intercultural theory as an analytical instrument for what IC can be in SSE. The following research question has been formulated to examine the study objective: What aspects of the IC construct are addressed in the NTSS test items from 2013 to 2022?

Literature Review

Previous research has shed light on the intercultural framing of Swedish education (e.g., Lorentz, 2018; Lahdenperä, 2018; Sharif, 2008). Johansson (2023) argues an intercultural mission is embedded in the Swedish curriculum (Lgr22), comprising knowledge and attitudinal learning objectives. It is stated in Lgr22 that Swedish schooling should develop students' knowledge of similarities and differences in living conditions and culture, cultural diversity within the country, and societal problems and possible solutions (Swedish National Agency for Education [NAE], 2022). It also stipulates that teachers should foster an awareness of cultural heritage as a part of one's identity, empathy towards others, and view the cultural perspectives of others (*ibid.*).

The Swedish curriculum's overarching focus on understanding culture allows for different approaches to IC in school subjects, and research indicates that some subjects align more closely with IC than others. A substantial body of educational research has framed IC as a concept associated with language and communication, studying the interaction between culturally distinct groups (e.g., Fredriksson & Wahlström, 1997; Lundgren, 2002; Stier, 2021). In parallel to this 'language approach' to interpreting IC, the research field in education has expanded to include other disciplinary foci, such as

history education (Eliasson & Nordgren, 2016; Johansson, 2012; Nordgren & Johansson, 2015) and religious education (Jackson & O’Grady, 2019; Johannessen & Skeie, 2019). In this emerging wave of IC research taking a ‘social studies approach’ to IC research, a few researchers have argued that IC is an integral part of civic education and SSE (e.g., Linkova, 2014; Nestian Sandu, 2015; Perry & Southwell, 2011). While these researchers have provided valuable insights into the affinity between IC and SSE, they do not explicitly demonstrate how IC overlap with the contents of SSE. Thus, the existing research presents an important yet understudied opportunity for further exploration in the field of IC research.

The forthcoming section briefly summarises the literature on assessment research in a Swedish context, outlining trends and shortcomings in the field. While acknowledging the important research previously conducted on SSE assessment in a Swedish context (e.g., Jansson, 2011; Karlsson, 2012; Odenstad, 2011), this review will focus on previous research on national assessments.

Prior Research on National Assessments

A substantial body of Swedish national test research in the social studies subjects (Samhällsorienterande ämnen) has studied samples of student responses (e.g., Andersson et al., 2023; Arrhenius, 2024; Nolgård, 2021). Compared to studies on national tests in other social studies subjects, such as history education (Alvén, 2021a, 2021b; Ammert & Eliasson; 2019; Eliasson, 2013; Nolgård, 2021; Nolgård & Nygren, 2019; Rosenlund, 2011; Samuelsson & Wendell, 2016), religious education (Andersson et al., 2023; Sporre et al., 2023; Tykesson, 2017), and geography education (Alm Fjellborg & Kramming, 2021; Arrhenius et al., 2024, 2020; Molin & Alm Fjellborg, 2019), there is a relatively small body of literature that is concerned with SSE. Exceptions include Löfstedt’s (2018) study of interrater reliability and Andersson and Larsson’s (2023) study of students’ perspective-taking ability in the context of the pros and cons of referendums. Both Löfstedt (2018) and Andersson and Larsson (2023) were limited to investigating data obtained from the 2013 test, leaving a decade of subsequent NTSS cycles unexplored. Accordingly, little research has investigated the NTSS content and construction, and no studies have examined IC in the context of the NTSS.

Although the field has generally focused on other objectives than the present study, there are certain similarities between the present study and other assessment research on national tests. For instance, Ammert and Eliasson (2019) employed a validation approach to the National Test in History for grade 9, gathering construct validity evidence. Drawing on the Standards for Education and Psychological Testing (Standards for brevity) (AERA, APA, & NCME, 2014) validation framework, they examined the alignment between the target construct, i.e., the syllabus and assessment criteria in history education, and the contents of the National Test in History. Additionally, like the present study, previous research has examined how the national tests implicitly assessed constructs not explicitly included in a syllabus or knowledge criteria. For instance, Andersson et al. (2023) elucidated how powerful knowledge was assessed in the context of the 2017 National Test of Religious Education. Comparably,

Sporre (2019) studied how ethical competence was addressed in the curricula, syllabuses, test items, and assessment instructions for the National Test of Religious Education. The present study shares key similarities with these studies but also introduces a novel approach to national assessment research regarding the study object (IC) and material (all accessible NTSS test items).

National Testing of Social Science

As a product of a government decision from the Swedish Ministry of Education (U2011/6543/S), NTSS for grade 9 was first introduced in 2013 and has been implemented annually to assess knowledge and skills following the educational objectives defined in the SSE syllabus. The NAE formulated the purpose and function of the tests as follows:

The primary purpose of the national tests is to support equal and fair assessment and grading and to provide a basis for analysing the extent to which the knowledge requirements are met at the school, principal, and national levels. [...] The national tests can also help to concretise the syllabuses (cited in Hirsh, 2016, p. 7, my translation).

It is important to note that in response to updated regulations for national tests (Prop. 2017/18:14), the NTSS development has increased emphasis on test equivalence. Yet, its foundational purpose of supporting equal and fair assessment and concretising the syllabuses has been sustained throughout 2013–2022.

NAE outsources the construction of the National Test to different university departments, and the Department of Pedagogical, Curricular, and Professional Studies at the University of Gothenburg constructs NTSS on behalf of NAE. Different NTSS aims to assess particular parts of the syllabus and knowledge requirements, only targeting specific sections of its content per test cycle (University of Gothenburg, 2022). Accordingly, each test item can be considered an operationalisation of a particular SSE syllabus and knowledge requirements sample specified in the test's assessment instructions.

The National Test can redirect learning priorities, as its content signals what subject matter is important to teachers and students. This, in turn, influences teachers' prioritisation of certain teaching content (Arensmeier & Lennqvist Lindén, 2017; Lundahl, 2009). Importantly, the contents of these state-mandated tests also inform teachers how to assess subject concepts in SSE.

Specifying the NTSS Target Construct

The underlying construct(s) assessed in the NTSS is SSE's syllabus and knowledge requirements. The SSE syllabus for grade 9, the focus of this study, consists of the subject's frame, aim, core contents, and knowledge requirements. The frame comprises a short introductory paragraph that contextualises the subject. The syllabus explicitly references "intercultural relations" in the frame (NAE, 2022, p. 218), acknowledging the subject's IC relevance. The aim comprises a concise description of the learning objectives of SSE in compulsory school. Particularly, the SSE aim states that the subject

should provide students with “the prerequisites to analyse social issues from different perspectives” (NAE, 2022, pp. 218–219), including the views of different social groups. The aim is followed by the core contents for years 7–9, which focuses on terms such as identity, ethnicity, power, justice, social equality, migration, integration, segregation, minorities, global issues, social issues, and interdependence (NAE, 2022, pp. 222–223). Lastly, the syllabus includes the knowledge requirements, detailing the criteria required to obtain a particular grade in SSE. These criteria include knowledge about democratic values, human rights, societal structures, and social issues (NAE, 2022, p. 225). Identical and similar formulations regarding SSE appear in previous curricula (NAE, 2019).

Given the content of the SSE syllabus and knowledge criteria, I argue that the NTSS assessment scope can capture core aspects of IC and, by extension, that the NTSS score interpretations can be used to draw inferences about students’ IC. This argument is theoretically supported by the *Standards* (AERA, APA & NCME, 2014), which assert that a single test can be interpreted differently depending on its use. Importantly, this argument is consistent with claims in previous research about the content overlap between IC and SSE (Linkova, 2014; Nestian Sandu, 2015; Perry & Southwell, 2011). The above review of the contents of the Swedish syllabus and knowledge criteria for SSE further strengthens this argument. These claims, coupled with the need for alternative assessments of IC, justify the proposed investigation of the NTSS. In the following section, I contextualise what IC can be in SSE and propose a conceptual framework that will serve as the analytical instrument for understanding the NTSS as an IC assessment.

A Theoretical Framework of IC for SSE

IC is considered a multidimensional construct encompassing a range of competencies, commonly categorised into three dimensions: knowledge, skills, and attitudes (Deardorff, 2006, 2009; Malazonia et al., 2021). Researchers have posited that these dimensions can be further divided into subcomponents, whose contents vary depending on the research context (Luo & Chan, 2022; Zhang & Zhou, 2019).

The conceptual framework I propose in this paper is a collection of IC subcomponents extracted from the definitions of the ‘social studies approach’ mentioned in the literature review section. Hence, I have purposefully omitted the language-related IC components from this conceptualisation, excluding associations to communicative competence, cross-cultural communication, intercultural interaction, international communication, and multicultural communication. The conceptual framework is further substantiated by incorporating theoretical insights from a recent systematic literature review of the topic (Al-Afifi et al., 2025), which examined 4904 articles on IC. By isolating subcomponents of the social studies perspective, I address a critical gap identified in the research literature: the lack of a subject-specific conceptual framework for IC applicable to SSE.

Intercultural Knowledge

To organise the contents of intercultural knowledge, I order them into three categories: *cultural*, *social*, and *global*. Below, I elucidate the key contents of each category.

Cultural knowledge entails understanding the norms, beliefs, customs, traditions, and social practices of diverse cultural groups (Byram, 1997; Piątkowska, 2015; Ubani, 2013) and how they shape one's cultural worldview (Bennett, 2009). Culture-specific information about the similarities and differences between foreign and home cultures is posited to raise cultural self-awareness, which is suggested to be a necessary precursor of intercultural learning (*ibid.*).

Social knowledge encompasses understanding social systems within diverse cultural environments, with an emphasis on concepts of power and justice (Gorski, 2008). Regarding power, knowledge of hegemonic structures, power dynamics, social systems, bias, subjectivity, and discriminatory practices are considered important for acquiring IC. In previous IC definitions, expressions of power have been exemplified in the context of sociopolitics, immigration, and poverty (Sue, 2001). In addition, being able to affect social change to rectify inequalities has been described as a critical feature of IC social knowledge (Banks, 1999; Nestian Sandu, 2015). To recognise why societal change is desirable, knowledge of justice is considered important to gain IC (Gorski, 2009). Justice concerns an egalitarian view of values central to social life, such as fairness and equality (Jackson & Cohen, 1988). According to IC research, understanding justice comprises perspectives on equal opportunity, equal legal protection, and fair treatment (Aguado-Odina et al., 2017; Pittman, 2009).

Lastly, *global knowledge* comprises issues of concern to the international community, aspiring to bring a global perspective to the local environment (Huber et al., 2014). Bringing awareness of the world's interconnectedness highlights the implications of global events (Deardorff, 2006), such as migration, poverty, conflicts, and climate change (Budiharso et al., 2024). Similarly, expressions of global interdependence and international understandings, such as human rights, are deemed a part of this knowledge category (Deardorff, 2009).

Intercultural Skills

Various skills have been proposed to be associated with IC, mainly related to critical thinking (Esen, 2021). Two subskills of critical thinking are prominent in IC theory, which emphasises the ability to self-critically examine culturally conditioned assumptions: perspective-taking and reflexivity.

Johnson and Johnson (2015) suggest that *perspective-taking* is a fundamental characteristic of IC. Perspective-taking has been defined in IC research (e.g., Mateo et al., 2016) according to Epley and Caruso's (2008) definition, stating that it is a mental ability "to get beyond one's point of view to consider the world from another's perspective" (p. 299). This depiction aligns with Bennett's (1998) definition, in which perspective-taking is an intentional effort to experience and comprehend someone else's culture to overcome ethnocentrism. Thus, developing perspective-taking in intercultural

research requires overcoming barriers of egocentric views (cf. Gehlbach, 2011; Mateo et al., 2016).

Furthermore, *reflexivity* is promoted in IC theories, emphasising self-contemplation to analyse cultural awareness (Blasco, 2012). Matthews (2020) suggests that reflexive thought brings individual dispositions to the fore, compelling oneself to reflect upon the structures of selfhood, including how one's assumptions, biases, and social position influence one's understanding of the world. As such, reflexivity highlights one's positionality in a social context through internal conversation (Archer, 2010). In contrast to self-reflection, reflexivity transcends the individual level, considering the relation between the self and social context.

Intercultural Attitudes

Several attitudes and values have been studied in relation to IC, such as tolerance, compassion, open-mindedness, and inclusivity (Deardorff, 2015; Huber & Brotto, 2012). Much of this research consistently links these attitudes and values to establishing *respect* between cultural groups. Trompenaars and Wooliams (2009) define respect in the context of IC research as being “a basic moral principle and human right that is accountable to the values of human dignity, worthiness, uniqueness of persons, and self-determination” (p. 446). Several objects to what respect should be given have been suggested, such as respect for cultural value-based differences (Deardorff, 2009).

An Integrative View of the Conceptual Framework

Taken together, the concept of IC comprises components that can be divided into an extrospective aspect, which involves understanding others' cultural, social, and global perspectives, and an introspective aspect, which involves considering one's own cultural, social, and global positions. At its core, the interplay between internal and external awareness is a key notion emphasised in the IC literature related to SSE. In addition, IC definitions assume a normative outlook on adopting certain attitudes conjoined with respect for people's differences (Deardorff, 2006). This conceptualisation of IC, containing three categories of knowledge (cultural, social, and global knowledge), two categories of skills (perspective-taking and reflexivity), and one principal attitude (respect), constructs a fundamental framework for understanding this competence's meaning in SSE.

Method

The present study utilises a concept derived from validity theory to anchor the analysis of the NTSS's aptitude to assess IC (Cizek et al., 2018). Validity is “the relationship between the content of a test and the construct it is intended to measure” (AERA, APA & NCME, 2014, p. 14), and validity is substantiated by establishing multiple forms of relevant validity evidence (Messick, 1989). This validation process requires constructing arguments for and against an intended interpretation of test use (Kane, 2006). The arguments are, in turn, an integration of “various strands of evidence

into a coherent account of the degree to which existing evidence and theory support the intended interpretation of test scores for specific uses" (AERA, APA & NCME, 2014, p. 21).

This study constructs a validity argument by eliciting and evaluating validity evidence of *construct comprehensiveness*, one of the four sources of validity proposed in Cizek et al.'s (2018) Generalised Assessment and Alignment Tool (GAAT) validation approach. Conceptually, construct comprehensiveness is the degree to which a content domain of a test covers the breadth of the construct of interest (*ibid.*). Such content-oriented evidence is not exclusionary to Cizek et al. (2018) but is considered a significant source of validity evidence in various validation frameworks (e.g., AERA, APA & NCME, 2014; Kane, 2006; Messick, 1989), which all fundamentally agree that the content of a test should be consistent with the specification of the construct. Applying this validation concept to the context of this study, construct comprehensiveness denotes the content representativeness of the test domain (NTSS test items) in relation to the specification of the IC construct (the proposed conceptual framework).

This study deviates from Cizek et al.'s (2018) approach to construct comprehensiveness. They take the analysis further by computing a validation score based on the interpretations of the degree of construct comprehensiveness. Such a quantitative extension is beyond this study's scope. Instead, the degree of fit is determined by gathering an evidential basis for the IC assessment capacities afforded by the NTSS, enabled by a qualitative analysis of what aspects of the IC construct overlap with the NTSS test items.

To evaluate the construct comprehensiveness of IC in the NTSS, a *directed content analysis* was deemed appropriate to qualitatively assess the IC aspects addressed in the NTSS test items. Content analysis is a flexible method for coding and categorising data, allowing pattern recognition in text-based materials (Cavanagh, 1997; Kondracki & Wellman, 2002). What distinguishes directed content analysis from conventional content analysis is that the former uses existing theory to code the data and aims to extend the theoretical framework (Hsiu-Fang & Shannon, 2005). Hence, directed content analysis was considered a suitable method due to its deductive approach (Potter & Levine-Donnerstein, 1999). The theoretical understanding of IC outlined in the theory section guided the interpretation of what IC aspects are constructed in the NTSS test items. According to the *Standards* (AERA, APA & NCME, 2014), the role of the conceptual framework in validation is to point to the "kinds of evidence that might be collected to evaluate the proposed interpretation in light of the purposes of testing" (pp. 11–12). As such, the central categories of the analysis were defined *a priori*, comprising cultural knowledge, social knowledge, global knowledge, perspective-taking, reflexivity, and respect. This initial layer of analysis was then further elaborated by generating subcategories based on what specific aspects of the knowledge and skills the interculturally relevant test items addressed. The secondary categorisation is intended to better understand the IC aspects addressed in the NTSS.

Material and Data Collection

The empirical material in this study is the NTSS test items. The present study is delineated to NTSS up to 2022, all constructed under the former curriculum Lgr11. Due to the COVID-19 pandemic, the 2020 test was cancelled, and the following 2021 test was neither mandatory nor contained new test items as it reused a preceding test. The remaining eight NTSS constructed in its first decade of existence between 2013 and 2022 constitute the material for this study.

National test data are confidential five years after publication. Consequently, the two most recent included NTSS (2019, 2022) are still confidential, implying that their test items cannot be disclosed in this study. Nevertheless, these NTSS cycles may be included in the analyses for research purposes, which they were in this study. To uphold the confidentiality agreement between the researcher, test developers, and NAE, the following analysis will not include explicit information linking this study's results to the confidential test items. This study only supports its analysis' inferences by exemplifying non-confidential test items that are representative of the particular finding (cf. Alvén, 2021a). The exemplified test items were originally written in Swedish and have been translated by the author.

Data Analysis and Procedure

This study uses the proposed conceptual framework as an analytical backdrop to investigate the construct comprehensiveness of IC aspects in NTSS test items. The directed content analysis of NTSS encompassed (a) identification of which of the studied test items addressed IC and (b) theory integration that sought to interpret what specific aspects of IC are addressed in the test item.

The test item identification defined prespecified eligibility criteria to distinguish between interculturally relevant and non-relevant test items. This distinction is necessary, as a test item might test a generic skill, such as perspective-taking, but in a non-intercultural context. As such, the study's analytical interest lies not in the knowledge, skills, and attitudes tested in the NTSS *per se* but in the intercultural frame paired with a particular knowledge, skill, or attitude.

TABLE 1

Eligibility criteria for NTSS test items intercultural relevance.

Eligibility criteria for intercultural relevance
1. Test item prompts cultural norms, values, beliefs, customs, traditions, or social practices
2. Test item prompts cultural, social, ethnic, religious, or minority groups
3. Test item prompts subjects related to globalisation or internationalisation, such as global interdependence
4. Test item prompts global issues, such as global migration, human rights, conflicts, famine, and poverty
5. Test item prompts cultural-social issues, such as immigration, segregation, discrimination, prejudice, and equality
6. Test item prompts students to consider countries other than Sweden

ASSESSING INTERCULTURAL COMPETENCE THROUGH SWEDISH NATIONAL TESTS OF SOCIAL SCIENCE: A VALIDATION STUDY

Markus Al-Afifi

A test item's intercultural relevance was determined by comparing its content with the eligibility criteria stated in Table 1. If the test item included one or more of these prompts, it was considered IC-relevant.

TABLE 2

Coding Scheme for IC aspects.

Code name	Code definition	Example
Cultural knowledge	Reference to: <ul style="list-style-type: none">diverse cultural groups' norms, beliefs, customs, traditions, and social practices, including home and foreign culture.	Item prompts consideration of cultural identity.
Social knowledge	Reference to: <ul style="list-style-type: none">influences of social structures on power dynamics within a society.the distribution, maintenance, and exercise of power among individuals and groups, such as discrimination, oppression, marginalisation, suppression, and exploitation in social systems.principles of fairness, such as social justice or equity and equality.	Item prompts consideration of access to resources circumscribed by a social context.
Global knowledge	Reference to: <ul style="list-style-type: none">transnational issues and their possible solutions, such as poverty and conflicts.global events and their effects, such as migration.expressions of global interdependence, such as Sustainable Development Goals and human rights.	Item prompts consideration of interconnectedness between people across different countries.
Perspective-taking	Reference to: <ul style="list-style-type: none">social perspective(s) assumed in the test item, such as national minorities or a specific cultural group.	Item prompts consideration of a particular person, group, or country.
Reflexivity	Reference to: <ul style="list-style-type: none">individual positionality in relation to a social context.cultural self-awareness.	Item prompts consideration of introspection of one's own beliefs or centrisms.
Respect	Reference to: <ul style="list-style-type: none">respect between cultural groups.	Item prompting consideration of cultural similarities or differences.

After identifying the interculturally relevant test items, the theory integration process commenced. A coding scheme was developed to classify which specific IC aspects an interculturally relevant test item prompted students to address (Table 2). The scheme is framed according to the theoretical underpinnings of intercultural knowledge, intercultural skills, and intercultural attitudes specified in the conceptual framework. By

using this coding scheme, it is possible to systematically analyse what IC aspects are addressed in the NTSS test items.

Results

This study examined the aspects of IC addressed in NTSS test items from 2013 to 2022 to evaluate the construct comprehensiveness of IC in the NTSS. The subsequent sections present results from the directed content analysis of the NTSS test items. First, a brief overview of the NTSS test items is introduced to provide a contextual backdrop for the following analysis. The subsequent sections elucidate the interculturally relevant test items, accounting for the specific NTSS contents associated with IC.

NTSS Intercultural Relevance

The included test cycles contain varying amounts of test items, ranging from 24 to 30. In total, the NTSS test items from 2013 to 2022 consisted of 210, of which 46 were deemed interculturally relevant (Appendix 1). 11 interculturally relevant test items were identified in the two most recent NTSS (2019, 2022).

Significant individual variation was observed in what IC aspects an NTSS captured on a test cycle level (Figure 1). While the analysed NTSS cycles might have covered certain IC aspects collectively to some extent over ten years, there is substantial variability in the IC content covered per test. During its first decade, the NTSS assessed circa six interculturally relevant test items on average per test ($M = 5.75$). However, the range between the lowest and highest observed number of interculturally relevant test items varied extensively ($Range = 1-10$), indicating less consistency in NTSS's reliability as an IC assessment.

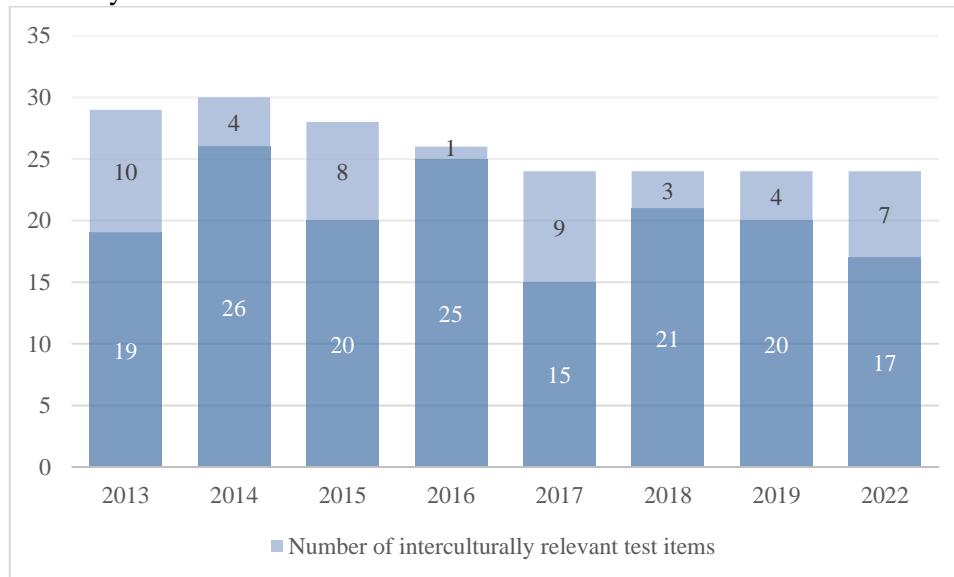


FIGURE 1

Frequency distribution of the number of interculturally relevant test items in each test cycle.

ASSESSING INTERCULTURAL COMPETENCE THROUGH SWEDISH NATIONAL TESTS OF SOCIAL SCIENCE: A VALIDATION STUDY

Markus Al-Afifi

Regarding the eligibility criteria, most of the interculturally relevant test items were included because they prompted global issues (20 items), such as global migration, human rights, conflicts, and poverty, or social issues (19 items), such as immigration, segregation, discrimination, prejudice, and equality. In contrast, only a handful of items were included due to their relevance to cultural norms, values, beliefs, customs, traditions, and social practices or cultural, social, ethnic, religious, or minority groups. Accordingly, this discrepancy indicates how IC aspects have been contextualised in the NTSS, significantly more so in the context of global and social issues than in culture.

TABLE 3

Aspects of IC addressed in NTSS test items.

Intercultural dimension	IC aspect	Test item addresses
Intercultural knowledge	Cultural knowledge	Advantages and disadvantages of globalised culture, cultural background as an influencing factor of choices, global culture, defining culture and culture clash.
	Social knowledge	Power, equality, equity, discrimination, prejudice, racism, exploitation, living conditions, and segregation. Equal opportunity and access to rights. Democratic rights and responsibilities. Equal opportunity to education, housing, and employment.
	Global knowledge	Transnational issues, views on globalisation, common goods of international cooperation, global trade.
Intercultural skills	Perspective-taking	Perspectives of socially, politically, and economically vulnerable groups. Perspectives of national minorities, migrants, child labourers, and disenfranchised girls in other countries. Individual, group, organisational, societal, and global perspectives.
	Reflexivity	National self in relation to a global context. Individual dispositions in relation to the global market economy, segregation, and globalisation.
Intercultural attitudes	Respect	Respect for human rights and lack of respect (discrimination and inequality)

IC Aspects Addressed in NTSS Test Items

What aspects of IC are addressed in NTSS test items from 2013 to 2022? As Table 3 indicates, there is considerable variation in what IC knowledge, skills, and attitudes have been captured in NTSS test items. Table 3 illustrates the different aspects accounted for in six separate categories structured according to the subject-specific IC framework, each subdivided into emergent themes of NTSS content focus. One interculturally relevant test item typically captured more than one IC aspect within the same dimension (e.g., global knowledge and social knowledge) or across dimensions (e.g., intercultural knowledge and intercultural skills). These findings suggest that the NTSS test item construction allows students to demonstrate a multidimensional view of IC. In the following sections, I will analyse each subcategory of intercultural knowledge, skills, and attitudes. Original item formulations are in Appendix 1, referenced in the numbers below.

Cultural Knowledge

In contrast to the global and social knowledge dimension, hardly any of the 46 interculturally relevant test items concerned cultural knowledge. One item referenced culture in the introductory sentence of the question, directing students' attention to using cultural knowledge in their answers of advantages and disadvantages of globalisation (No. 28). Similarly, *cultural background* and *ethnicity* are stated as key terms to consider when responding to two distinct test items (No. 1, 7). Additionally, students were supposed to write the right letter in the right box, in which *culture* and *culture clash* was supposed to be connected with their definitions (No. 6). Apart from these instances, aspects of cultural knowledge, such as cultural norms, values, beliefs, customs, traditions, or social practices, have not been a leading objective of the NTSS assessment foci.

Some items implicitly addressed cultural knowledge, allowing students to answer using cultural knowledge but not explicitly instructing them to account for it. For example, the query below about girls' disenchantment in other countries could be partially answered by detailing cultural barriers to girls' education, unequal gender norms, and oppressive values (No. 18).

Describe what might prevent girls in some countries from being allowed to attend school.

A notable observation from this data category was that items indicating relevance for cultural knowledge emphasised foreign culture, diverting students' attention towards the cultural other rather than the cultural self. While test items such as the girls' disenchantment query might allow for some comparison between home and foreign culture, the test item's purpose is anchored in an out-of-Swedish context, as the NTSS respondents are enrolled in Swedish compulsory education, hence not included in the item's intended population.

Social Knowledge

The content of the NTSS test items addressed a range of IC aspects related to social knowledge, including power, equality, equity, integration, discrimination, prejudice, racism, exploitation, living conditions, and segregation. The most prominent social knowledge aspect of IC was *discrimination*, which students were prompted to consider across several NTSS cycles. For instance, items targeting discrimination included the following:

Which of the following situations could be an example of discrimination? (No. 2).

In what ways can discrimination be a social problem? (No. 3).

Discrimination is not allowed under Swedish law. Tick whether the options are examples of discrimination under Swedish law (Yes) or not (No)! If you put two crosses on any sub-task, your answer is considered invalid for that sub-task. (No. 24).

These examples illustrate how discrimination was probed with different assessment objectives, situating it as a social problem and legal offence.

A further aspect of social knowledge identified was *prejudice* (No. 11, 42). Specifically, one test item prompted students to consider bias in media coverage about bomb attacks in the U.S. and Iraq, inciting students to probe societal prejudice about why bombing might be considered more or less ordinary in different parts of the world (No. 11).

One test item on *exploitation* encouraged students to consider why child labour exists, its consequences, and possible solutions (No. 29). Regarding *racism*, one interculturally relevant test item incited students to argue for and against banning racist organisations (No. 23).

Several items implicitly addressed social aspects of *power*, *equality*, and *equity*. For instance, in the following test item, students are prompted to consider social, cultural, legal, economic, and political structures, integrating aspects of power and equality, without being given direction to consider these aspects explicitly (No. 18)

Describe what might prevent girls in some countries from being allowed to attend school.

Similarly, the exemplified test item indirectly addresses power, equality, and equity. For example, discrimination presumes an asymmetrical power relationship between two parties, where one can exert power by performing discriminatory acts.

One item regarding *living conditions* asked students to consider individuals' ability to influence their life situation and that of others by asking them to consider what causes people to flee (No. 25).

Aspects of *justice* were also accounted for in the interculturally relevant test items, primarily through test items that addressed equal opportunity and access to rights, such

as international access to human rights. This set of interculturally relevant test items was reflected in questions such as the following:

Which of the following statements violates UN human rights? Tick the 3 statements that violate human rights. If you tick more than 3, the answer is invalid (No. 15).

In the box you will find examples of human rights from the UN Declaration of Human Rights. Select three of these and briefly describe what the right means. (No. 21).

Human rights are said to be a prerequisite for democracy. Discuss why respect for human rights is so important in a democracy. Use some different human rights in your reasoning. (No. 22).

As the last examples show, justice as access to human rights was also framed within the context of democracy.

Apart from rights and responsibilities, students were also prompted to consider justice in terms of equal opportunity to education between genders, which allows students to account for fair treatment of individuals based on gender (No. 18). Equal opportunity for access to the labour and housing market was also reflected in response alternatives to multiple-choice questions (No. 2). In this query about situations that could be viewed as discrimination, one of the listed response alternatives stated, “Hassan has been waiting in the housing queue since he received his residence permit. His friend Stefan is in the same queue. One day, Hassan discovers that Stefan has been given an apartment, even though Stefan has been in the queue for less time than he has” (NAE, 2013, p. 11, my translation). In this scenario, the character with a non-typically Swedish first name is treated unfairly based on presumed ethnicity and denied equal housing opportunities to a native Swede. A similar scenario related to unequal chances of employment stated that “an employer hiring someone with poorer qualifications (e.g., grades) and a Swedish name instead of someone with better qualifications and a foreign name” (NAE, 2017, p. 4, my translation). In these examples of response alternatives, the NTSS alludes to students’ sense of fairness, considering that the selection process for jobs and housing should treat all individuals equally regardless of irrelevant attributes. Lastly, students were incited to reflect on how to rectify inequalities in the context of child labour (No. 29). The exemplified test items show that all central aspects of justice promoted in IC theory, including equal opportunity, equal legal protection, fair treatment, and rectifying inequalities, have been accounted for in the NTSS. In summary, the NTSS test items addressed several central aspects of the IC social knowledge dimension.

Global Knowledge

The NTSS content has addressed global knowledge from four separate standpoints: transnational issues, views on globalisation, common goods of international cooperation, and global trade. While these standpoints all shared a common outlook for a globally outreaching perspective, their central assessment objective has different points of departure.

The test items targeting *transnational issues* involved migration, poverty, war, and conflict (e.g., No. 5, 8, 9, 10, 12, 13, 16, 17, 25, 26, 27, 29, 45). The test items that addressed *views on globalisation* asked students to account for why human living conditions have improved globally and the causes of globalisation (No. 17, 26, 43, 44). One test item addressed both the positive and negative aspects of globalisation, prompting students to consider its merits and flaws (No. 28).

A third category of interculturally relevant test items addressing global knowledge addresses *common goods of international cooperation*, including the Universal Declaration of Human Rights, the Millennium Development Goals, and foreign aid (No 8, 10, 15, 21, 22, 36). Lastly, some test items targeted *global trade*, including exports and imports to and from foreign sources and offshoring industries' means of production to foreign countries (No. 14, 20, 30, 35, 46). One item also prompted consideration of the Human Development Index correlation compared to BNP in various countries (No. 33).

Perspective-Taking

Students were tasked with taking a variety of perspectives across the 46 interculturally relevant test items. Positioning the item formulations from the standpoint of socially, politically, and economically vulnerable groups was a reoccurring theme in the NTSS, including perspectives of national minorities, migrants, child labourers, and disenfranchised girls in other countries (No. 4, 18, 25, 29, 31). From these perspectives, students are incited to reflect upon these matters from another's perspective, going beyond one's own point of view. As is perhaps best indicated by the test item about child labour (No. 29), several perspectives were often interwoven into one test item, allowing students to apply an intersectional perspective to the task at hand. In this item, students are first asked to describe reasons for child labour, its consequences, and solutions, which allows them to consider these aspects from individual, societal, and global perspectives. Intersectional perspectives were also present in less detailed formulations, such as the query about girls' disenfranchisement (No. 18), which requires considering gender and nationality.

The interculturally relevant test item encompassed not only the perspectives of disadvantaged groups but also more impersonal perspectives, such as those of media consumers (No. 11, 37, 42). Moreover, students had to assume the perspectives of organisations and governments, such as legislative power (No. 10, 23, 39). For instance, in test item No. 23, students were assigned to argue for and against prohibiting anti-democratic organisations, advocating for democratic rights and liberal democratic values.

Reflexivity

The interculturally relevant test items addressed some aspects of reflexivity. Students were prompted to reflect upon their national self in relation to a global context, inciting reflexive thought about Sweden's position in international relations (No. 5, 32, 34). In the following example, reflexivity was also addressed by assigning students to

consider an interculturally relevant issue on an individual level. This allowed students to reflect upon how certain individual dispositions might be affected to a greater or lesser degree pertaining to the global market economy.

Sweden, like most countries in the world, exports and imports goods to and from all over the world. How are countries and individuals affected by global trade? Describe 1-3 consequences. (No. 32).

Similar formulations were present in test items targeting other IC content, such as segregation and globalisation (No. 28, 38, 43, 44).

However, no NTSS test item addressed self-contemplation to analyse cultural awareness, which is the most central aspect of reflexivity according to the IC framework. No evidence was found for how the NTSS addressed the cultural self in relation to a social context. These results indicate that the NTSS addressed some aspects of reflexivity, but no test item prompted students to anchor their responses in cultural reflexivity.

Respect

One of the interculturally relevant test items explicitly asked students to discuss respect for human rights (No. 22). Other interculturally relevant test items indicated a more distant association with respect (e.g., No. 40). The formulations do not foreground respect in the following test items. Nevertheless, respect, or the lack thereof, is central to understanding the test item's target content, such as discrimination and inequality (No. 3, 18). Thus, a test item's focus on IC aspects may not be immediately obvious based solely on the explicit test item formulation but may emerge implicitly.

Construct Comprehensiveness of IC in NTSS

The analysis of the NTSS test items from 2013 to 2022 yielded key insights into the coverage and limitations of IC assessment. Notably, approximately one-fifth of all NTSS test items over the observed ten-year period were found to address interculturally relevant content. Considering that the NTSS content domain is not constructed to make score interpretations of IC, it still adequately represents a range of IC aspects. The results of IC knowledge and skills dimensions in the NTSS test items indicate content coverage of several aspects of global knowledge, social knowledge, and perspective-taking. Regarding construct comprehensiveness, this result highlights the NTSS degree of fit to adequately capture the IC construct. In contrast, the limited testing of cultural knowledge and reflexivity suggests variability in the NTSS' coverage to measure different IC aspects within the knowledge and skills dimensions. These results show that although the NTSS can be a resourceful measure of some IC knowledge and skills, the extent of their assessment retains considerable restrictions.

Regarding attitudes, the analysis revealed that the NTSS addressed respect. This result is noteworthy, considering that students' opinions, beliefs, and attitudes are not assessment criteria in the NTSS. However, the attitudinal aspects were limited to only a few instances, limiting the significance of this result.

Overall, the findings of this research both provide evidence in favour of the NTSS's propensity to assess IC as well as some glaring gaps in its test content of the IC construct. All gauged construct aspects were somewhat represented, but construct underrepresentation of, for instance, cultural knowledge was a systematic issue, which undercuts the NTSS' IC inference possibilities. The NTSS degree of fit as an IC assessment based on existing evidence provides support for and against the reframed interpretation of test use. Thus, it can be concluded that the NTSS covers some breadth of the construct of interest, but the test content lacks consistency in addressing the entire construct. This conclusion is sustained when examining individual test cycles, where intercultural relevance varied extensively.

A Concluding Discussion

This study brought an intercultural frame to the national tests to explore the possibility of using an existing assessment as a supplementary resource to assess students' IC. By eliciting and evaluating validity evidence of construct comprehensiveness, the analysis provided insights into the strengths and limitations of NTSS coverage of IC. Concerning the research question, the results showed that the NTSS addresses several aspects of IC, providing evidence supporting the integrality of IC in the NTSS content. These findings display the NTSS utility beyond its primary functions as an external grading instrument.

Although the NTSS has tapped into key IC content, this finding implies some important caveats. The analysis revealed that the IC aspects addressed in the NTSS test items differed substantially in coverage. For instance, both the IC knowledge and skills dimensions lacked representation of focal aspects. In addition, the assessment opportunities afforded by the NTSS for the IC attitudes dimension were generally more limited than those for the knowledge and skills dimensions. The study's conclusions about the shortcomings of the NTSS as an IC assessment should not be interpreted as a critique of the true validity of the NTSS, as this investigation has not assessed the testing validity of its target construct. Dictated by its design, the NTSS test format biases the distribution of IC towards disproportionality between knowledge, skills, and attitudes.

Moreover, the results of this study can be used for several purposes, such as informing the construction of IC assessments for research and teaching purposes. From a research perspective, the 46 interculturally relevant test items listed in Appendix 1 can be used as a point of departure directing the next course of action to assess Swedish students' IC. The study's focus on large-scale assessments was intended to enable insight into students' proficiencies of IC from a nationally representative sample in later research. For instance, using Appendix 1 as a resource, researchers can identify which test items have addressed which IC aspect and then collect the archived responses to evaluate students' knowledge, skills, or attitudes. Accordingly, the identified interculturally relevant test items can be a reference point for scholarly inquiry and practical application in education.

Implications

The implications of this study's findings are threefold. First, the results provide insights into how the national tests, particularly the NTSS, can be used as an assessment instrument for a broader range of constructs than initially intended by the NAE. Similar to Andersson et al.'s (2023) study of powerful knowledge in the National Test of Religious Education, this study showed how the national tests' assessment capability can be interpreted differently by reframing the content of the intended construct. Furthermore, the NTSS has contributed to the field of intercultural assessment with salient and meticulously constructed test items for measuring several IC aspects, which can inform the construction of future IC assessments.

Secondly, the analysis raises awareness of the NTSS's intercultural profile. By extension, the study strengthens the perception of SSE as a subject with intercultural relevance, as one of the ancillary purposes of the national tests is to concretise the syllabus. By elucidating the interculturally relevant contents of the NTSS, the current understanding of SSE as an IC-relevant subject is strengthened. The present study not only contributes to a newfound understanding of the relevance of IC in SSE but also builds upon claims made in previous research about its relevance for Swedish education (Johansson, 2023; Lorentz, 2018; Lahdenperä, 2018; Sharif, 2008).

Thirdly, the study developed a new subject-specific conceptual framework in response to insufficient definitions of an SSE interpretation of IC. This framework explicitly theorises the association between IC and SSE, recognised in previous research (Linkova, 2014; Nestian Sandu, 2015; Perry & Southwell, 2011). As I have elaborated, IC occupies a broader disciplinary range that goes beyond the limits of language and communication, and is disciplinary aligned with SSE. Although IC has been studied from multiple orientations, this research has mainly been American-derived and language-centred. In contrast, the framework developed in this study has been applied to the Swedish subject of SSE. Its potential applications extend beyond those explored in this study, such as informing teaching content, assessments, and learning objectives for intercultural education. Accordingly, this framework can be considered progress towards understanding what subject-specific intercultural education can be in SSE. Still, additional research is needed to expand the proposed framework.

Limitations and Further Research

The study's findings are subject to at least two limitations. First, this study only provided validity evidence from one source of validity. Theorists are adamant that validation requires multiple forms of validity evidence (Kane, 2006). In this regard, this study falls short of providing insight into the degree of fit of the NTSS as an IC measurement, as it only examined construct comprehensiveness. Yet, the scope of this study necessitated a narrower validation approach due to the novelty of this research, whose objectives had to be explained and justified to a greater degree than, for instance, the validation presented in Ammert and Eliasson (2019). Future research is needed to comprehensively validate the NTSS as an IC assessment.

Moreover, the analysis had certain interpretative challenges. Some test items explicitly prompted students to consider a particular IC aspect, thereby strengthening the inferences made about what IC aspect the item addressed. Other test items implicitly addressed IC aspects, entering an ambivalent space of interpretability. To address this issue, I adopted a systematically inclusive approach in assessing which items were categorised as IC relevant, offering a generous interpretation of the inclusion criteria for intercultural relevance. Still, having two reviewers in this process would have strengthened the reliability of the results.

As a closing remark, although this study has contributed to our understanding of the NTSS as an assessment of IC, the international policy objective of IC-adapted education still faces several challenges in a Swedish educational context. In some respects, Sweden is obligated to provide students with IC education as a member of the UN, Council of Europe, and OECD. Nevertheless, the concept has yet to be mainstreamed in educational practice. The discrepancy between the importance of IC in international and national policy discourse causes an implementation gap between policy and practice. As this study has demonstrated, the NAE can contribute to closing this gap by taking advantage of the intercultural relevancy already embedded in the NTSS. The NAE's recognition of IC's applicability to Swedish national tests is not only an essential step towards adhering to international guidelines on intercultural education, but also communicates the importance of IC to Swedish teachers and students.

References

Aguado-Odina, T., Mata-Benito, P., & Gil-Jaurena, I. (2017). Mobilizing intercultural education for equity and social justice. Time to react against the intolerable: A proposal from Spain. *Intercultural Education*, 28(4), 408–423. Education Source. <https://doi.org/10.1080/14675986.2017.1333874>

Al-Afifi, M., Nygren, T., Rasmusson, M., & Tväråna, M. (2025). Intercultural Competence in the Classroom: A Systematic Review of Teaching Practices for Developing Intercultural Competence in Different Educational Contexts. Manuscript submitted for publication.

Alm Fjellborg, A., & Kramming, K. (2021). Sustainable development: Exploring gender differences in the Swedish national test in geography for grade 9. *International Research in Geographical and Environmental Education*, 31(3), 172–187. <https://doi.org/10.1080/10382046.2021.1927366>

Alm Fjellborg, A., & Molin, L. (2018). Vilka typer av uppgifter gynnar elever som följer kursplanen i svenska som andraspråk? En undersökning med data från de nationella proven i geografi. *Acta Didactica Norge*, 12(4), 5. <https://doi.org/10.5617/adno.6286>

Alvén, F. (2021a). Immigrant Students and the Swedish National Test in History. *Frontiers in Education*, 6, 774900. <https://doi.org/10.3389/feduc.2021.774900>

Alvén, F. (2021b). Pojkars och flickors resultat på det nationella provet i historia: Bestämmer uppgifterna historiska innehåll och form pojkar respektive flickors resultat? In Att bedöma i historieämnet.: Perspektiv på historisk kunskap, bedömning och prov / (Eds.) Ludvigsson, D.; Andersson Hult, L. (pp. 205–226). Stockholm: Historielärarnas Förening, 2021.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (Eds.). (2014). *Standards for educational and psychological testing*. American Educational Research Association.

Ammert, N., & Eliasson, P. (2019). Externa historieprov: En studie av validitetsfrågor. In *Skrifter med historiska perspektiv*. Malmö University, Fakulteten för lärande och samhälle. <https://doi.org/10.24834/978-91-7104-986-5>

Andersson, K., & Larsson, K. (2023). Elevers perspektivbyten i svar på uppgifter från det nationella provet i samhällskunskap. *Nordidactica: Journal of Humanities and Social Science Education*, 13(2023: 3), 66-84.

Andersson, K., Börjesson, M., & Larsson, K. (2023). Powerful Knowledge and the 2017 Swedish National Test in Religious Education. *Social Sciences*, 12(10), 533. <https://doi.org/10.3390/socsci12100533>

Archer, M.S. (2010). *Conversations about reflexivity*. London: Routledge.

Arensmeier, C., & Lennqvist Lindén, A.-S. (2017). Bemyndigande eller granskning—Nationella prov som styrinstrument. *Utbildning & Demokrati – tidskrift för didaktik och utbildningspolitik*, 26(2), 49–74. <https://doi.org/10.48059/uod.v26i2.1080>

Arrhenius, M., Lundholm, C., & Bladh, G. (2020). Swedish 12–13-year-old students’ conceptions of the causes and processes forming eskers and erratics. *Journal of Geoscience Education*, 69(1), 43–54. <https://doi.org/10.1080/10899995.2020.1820838>

Arrhenius, M., Bladh, G., & Lundholm, C. (2024). Causes, processes and consequences of earthquakes. Investigating Swedish 12-13-year old students’ geographical understanding. *International Research in Geographical and Environmental Education*, 1–17. <https://doi.org/10.1080/10382046.2023.2298555>

Banks, J.A. (1999). An Introduction to Multicultural Education (2nd ed.). Boston: Allyn and Bacon.

Bennett, M. J. (1998). Overcoming the golden rule: Sympathy and empathy. In M. J. Bennett (Ed.), *Basic concepts of intercultural communication: Selected readings* (pp. 191–214). Yarmouth, ME: Intercultural Press.

Bennett, M. J. (2009). Cultivating intercultural competence: A process perspective. In D. K. Deardorff (Ed.), *The SAGE Handbook of Intercultural Competence* (pp. 121–140). Thousand Oaks, CA: Sage.

Blasco, M. (2012). On reflection: Is reflexivity necessarily beneficial in intercultural education? 1. *Intercultural Education*, 23(6), 475–489.

<https://doi.org/10.1080/14675986.2012.736750>

Borghetti, C. (2017). Is there really a need for assessing intercultural competence? Some ethical issues. *Journal of Intercultural Communication*, 17(2), 1–15. <https://doi.org/10.36923/jicc.v17i2.736>

Budiharso, T., Widodo, T. W., Sintesa, N., Irawati, I., & Moybeka, A. M. S. (2024). Promoting Global Citizenship Education: Strategies for Developing Intercultural Competence in Students. *Global International Journal of Innovative Research*, 2(2). <https://doi.org/10.59613/global.v2i2.78>

Byram, M. (1997). *Teaching and Assessing Intercultural Communicative Competence*. New York: Multilingual Matters. <https://doi.org/10.21832/byram0244>

Cavanagh, S. (1997). Content analysis: concepts, methods and applications. *Nurse Researcher*, 4(3), 5–16. <https://doi.org/10.7748/nr.4.3.5.s2>

Cizek, G. J., Kosh, A. E., & Toutkoushian, E. K. (2018). Gathering and Evaluating Validity Evidence: The Generalized Assessment Alignment Tool. *Journal of Educational Measurement*, 55(4), 477–512. <https://doi.org/10.1111/jedm.12189>

Cobb, D., & Couch, D. (2022). Locating inclusion within the OECD's assessment of global competence: An inclusive future through PISA 2018? *Policy Futures in Education*, 20(1), 56–72. <https://doi.org/10.1177/14782103211006636>

Council of Europe. (2008). *White Paper on Intercultural Dialogue: Living Together as Equals in Dignity*. Council of Europe.

Council of Europe. (2016). *Competences for democratic culture: Living together as equals in culturally diverse democratic societies*. Council of Europe.

Deardorff, D. K. (2006). Identification and Assessment of Intercultural Competence as a Student Outcome of Internationalization. *Journal of Studies in International Education*, 10(3), 241–266. <https://doi.org/10.1177/1028315306287002>

Deardorff, D. (2009). *The SAGE handbook of intercultural competence*. Thousand Oaks, CA: Sage.

Deardorff, D. K. (2015). Definitions: Knowledge, Skills, Attitudes. In Bennett, J. M. (Ed.). (2015). *The SAGE Encyclopedia of Intercultural Competence*. In *SAGE Publications, Inc.* <https://doi.org/10.4135/9781483346267>

Eliasson, P. (2013). Historiemedvetande på prov – nationella ämnesprov i historia i Sverige. *Kasvatus & Aika (Verkkolehti)*, 7(3). Retrieved from <https://journal.fi/kasvatusjaika/article/view/68651>

Eliasson, P. & K. Nordgren. (2016). Vilka är förutsättningarna i svensk grundskola för en interkulturell historieundervisning?. *Nordidactica: Journal of Humanities and Social Science Education* 6:2 (2016), s. 47–68.

Epley, N., & Caruso, E. M. (2009). Perspective taking: Misstepping into others' shoes. In K. D. Markman, W. M. P. Klein, & J. A. Suhr (Eds.), *Handbook of imagination and mental simulation* (pp. 295–309). Psychology Press.

Esen, M. (2021). Critical Thinking and Intercultural Competence: Review, use, and classification of the 21st century skills in intercultural EFL education. *International Journal of English Language Teaching*, 8(1), 23. <https://doi.org/10.5430/ijelt.v8n1p23>

Fantini, A. (2007). Exploring and assessing intercultural competence (CSD Research Paper No. 07-01). St. Louis, MO: Washington University, Center for Social Development. doi: <https://doi.org/10.7936/K7TB16CX>

Fredriksson, U. & Wahlström, P. (1997). Skola för mångfald eller enfald. Stockholm: Legus.

Gehlbach, H. (2011). Making Social Studies Social: Engaging Students Through Different Forms of Social Perspective Taking. *Theory Into Practice*, 50(4), 311–318. <https://doi.org/10.1080/00405841.2011.607394>

Gorski, P. C. (2008). Good intentions are not enough: a decolonizing intercultural education. *Intercultural Education*, 19(6), 515–525. <https://doi.org/10.1080/14675980802568319>

Gorski, P. C. (2009). Intercultural education as social justice. *Intercultural Education*, 20(2), 890. <https://doi.org/10.1080/14675980902922135>
Government decision

Hirsh, Å. (2016). *Nationella prov i grundskolan: En studie av hur lärare och rektorer uppfattar och hanterar prov och provresultat*. Stockholm: Skolverket.

Hsieh, H., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), 1277–1288. <https://doi.org/10.1177/1049732305276687>

Huber, J. Byram, M., Barrett, M., Lázár, I., Mompoint-Gaillard, P., & Philippou, S. (2014). Developing intercultural competence through education. In *Council of Europe Publishing eBooks*. Retrieved from <https://ci.nii.ac.jp/ncid/BB17122692>

Huber, J., & Brotto, F. (2012). Intercultural competence for all: preparation for living in a heterogeneous world. In *Council of Europe Pub. eBooks*. Retrieved from <https://ci.nii.ac.jp/ncid/BB17117003>

Jackson, P. G., & Cohen, R. (1988). Justice: Views from the Social Sciences. *Contemporary Sociology*, 17(1), 68. <https://doi.org/10.2307/2069436>

Jackson, R., & O'Grady, K. (2018). The religious and worldview dimension of intercultural education: the Council of Europe's contribution. *Intercultural Education*, 30(3), 247–259. <https://doi.org/10.1080/14675986.2018.1539306>

Jansson, T. (2011). *Vad kommer på provet? Gymnasielärares provpraxis i samhällskunskap*. Licentiate. Karlstads University.

Johannessen, Y. L., & Skeie, G. (2018). The relationship between religious education and intercultural education. *Intercultural Education*, 30(3), 260–274. <https://doi.org/10.1080/14675986.2018.1540142>

ASSESSING INTERCULTURAL COMPETENCE THROUGH SWEDISH NATIONAL TESTS OF SOCIAL SCIENCE: A VALIDATION STUDY

Markus Al-Afifi

Johansson, M. (2023). *Interkulturalitet och historia historieundervisningens teori och praktik i en mångkulturell värld*. Diss. Karlstad: Karlstads University, Karlstad.

Johnson D. W. & Johnson, R., T. (2015). Competition and Cooperation. In Bennett, J. M. (Ed.). (2015). *The SAGE Encyclopedia of Intercultural Competence*. In SAGE Publications, Inc. eBooks. <https://doi.org/10.4135/9781483346267>

Kane, M. T. (2006). Validation. In R. L. Brennan (Ed.), *Educational Measurement* (4th ed., pp. 17–64). Westport, CT: Praeger Publishers.

Karlsson, A. (2012). Samhällsguide, individualist och moderator. Samhällskunskapslärares professionella förhållningssätt i betygsättningsrelaterat arbete. In H. Lödén (Ed.) *Forskning av och för lärare* (pp. 145–164). Karlstad University Press.

Kondracki, N. L., & Wellman, N. S. (2002). Content analysis: Review of methods and their applications in nutrition education. *Journal of Nutrition Education and Behavior*, 34, 224-230. [http://doi.org/10.1016/s1499-4046\(06\)60097-3](http://doi.org/10.1016/s1499-4046(06)60097-3)

Lahdenperä, P. (2018). *Interkulturalitet i undervisning och skolutveckling*. Skolverket.

Leeds-Hurwitz, W. (2013). Intercultural competences: A conceptual and operational framework. Paris: UNESCO. Retrieved from <http://unesdoc.unesco.org/images/0021/002197/219768e.pdf>

Linkova, M. G. (2014). Civic Education as an Intercultural Phenomenon. <https://doi.org/10.13140/RG.2.1.2260.6326>

Lorentz, H. (2018). *Interkulturell pedagogisk kompetens: integration i dagens skola*. (2nd ed.). Lund: Studentlitteratur.

Lundahl, C. (2009). *Varför nationella prov? – framväxt, dilemman, möjligheter*. Lund: Studentlitteratur.

Lundgren, U. (2002). *Interkulturell förståelse i engelskundervisning: en möjlighet*. Diss. Lund University. 2002. Malmö.

Luo, J., & Chan, C. K. Y. (2022). Qualitative methods to assess intercultural competence in higher education research: A systematic review with practical implications. *Educational Research Review*, 37, 100476. <https://doi.org/10.1016/j.edurev.2022.100476>

Löfstedt, A. (2018). Det första nationella provet i samhällskunskap—En studie i bedömarsamstämmighet. *Acta Didactica Norge*, 12(4), 13. <https://doi.org/10.5617/adno.6283>

Malazonia, D., Lobzhanidze, S., Maglakelidze, S., Chiabrishvili, N., & Giunashvili, Z. (2021). Intercultural competencies of students vs. Their civic activities (Case of Georgia). *Cogent Education*, 8(1), 1918852. <https://doi.org/10.1080/2331186X.2021.1918852>

ASSESSING INTERCULTURAL COMPETENCE THROUGH SWEDISH NATIONAL TESTS OF SOCIAL SCIENCE: A VALIDATION STUDY

Markus Al-Afifi

Mateo, J. C., McCloskey, M. J., Grome, A. P., Abbe, A., & Behymer, K. J. (2016). *Framework for Understanding Intercultural Perspective Taking in Operational Settings*. Retrieved from <https://apps.dtic.mil/sti/citations/AD1006861>

Matthews, B. (2020). The role of reflexive thought in the achievement of intercultural competence. *Intercultural Education*, 31(3), 330–344.
<https://doi.org/10.1080/14675986.2020.1728093>

Messick, S. (1989). Validity of Psychological Assessment. *American Psychologist*.

Molin, L., & Fjellborg, A. A. (2019). Geographical variations in the relation between final course grades and results on the national tests in social sciences, 2015-2017. *Educational Review*, 73(4), 451–469. <https://doi.org/10.1080/00131911.2019.1642303>

National Agency for Education. (2019). *Curriculum for Compulsory School, Preschool Class and School-Age Educare – LGR11*. Skolverket. Retrieved from <https://www.skolverket.se/download/18.35e3960816b708a596c3965/1567674229968/pdf4206.pdf>

National Agency for Education. (2022). *Curriculum for Compulsory School, Preschool Class and School-Age Educare – LGR22*. Skolverket. Retrieved from <https://www.skolverket.se/publikationer?id=12435>

Nestian Sandu, O. (2015). Civic and Intercultural Education: A Means for Community Development and Attitude Change. *SAGE Open*, 5(2), 215824401558037. <https://doi.org/10.1177/2158244015580371>

Nolgård, O. (2021). Realising Roma rights through a high-stakes history test: Secondary school students narrating the Swedish Roma past, present and future. *History Education Research Journal*, 18(2). <https://doi.org/10.14324/HERJ.18.2.05>

Nolgård, O., & Nygren, T. (2019). Considering the past and present of Romani in Sweden: Secondary school pupils' thinking and caring about the history of the Romani in national tests. *Education Inquiry*, 10(4), 344–367. <https://doi.org/10.1080/20004508.2019.1607708>

Nordgren, K., & Johansson, M. (2014). Intercultural historical learning: a conceptual framework. *Journal of Curriculum Studies*, 47(1), 1–25. <https://doi.org/10.1080/00220272.2014.956795>

Odenstad, C. (2011). *Prov och bedömning i samhällskunskap - En analys av gymnasielärares skriftliga prov*. Licentiate. Karlstad University Press.

OECD. (2018). Preparing our Youth for an Inclusive and Sustainable World: The OECD PISA global competence framework. Retrieved from <https://www.oecd.org/education/Global-competency-foran-inclusive-world.pdf>

Perry, L. B., & Southwell, L. (2011). Developing intercultural understanding and skills: Models and approaches. *Intercultural Education*, 22(6), 453–466. <https://doi.org/10.1080/14675986.2011.644948>

Piątkowska, K. (2015). From cultural knowledge to intercultural communicative competence: Changing perspectives on the role of culture in foreign language teaching. *Intercultural Education*, 26(5), 397–408.
<https://doi.org/10.1080/14675986.2015.1092674>

Pittman, C. T. (2009). Multicultural education and social justice actions. *Intercultural Education*, 20(2), 173–186. <https://doi.org/10.1080/14675980902975349>

Potter, W. J., & Levine-Donnerstein, D. (1999). Rethinking validity and reliability in content analysis. *Journal of Applied Communication Research*, 27(3), 258–284.
<https://doi.org/10.1080/00909889909365539>

Proposition 2017/18:14, Nationella prov – rättsvisa, likvärdiga, digitala

Rosenlund, D. (2011). *Att hantera historia med ett öga stängt: samstämmighet mellan historia A och lärares prov och uppgifter*. Licentiate. Lund University, 2011.

Sandahl, J., Tväråna, M., & Jakobsson, M. (2022). Samhällskunskap (social science education) in Sweden. *JSSE - Journal of Social Science Education*, Bd. 21 Nr. 3 (2022). <https://doi.org/10.11576/JSSE-5339>

Samuelsson, J., & Wendell, J. (2016). Historical thinking about sources in the context of a standards-based curriculum: a Swedish case. *The Curriculum Journal*, 27(4), 479–499. <https://doi.org/10.1080/09585176.2016.1195275>

Schauer, G. A. (2020). Measuring Intercultural Competence. In P. Winke & T. Brunfaut (Eds.), *The Routledge Handbook of Second Language Acquisition and Language Testing* (1st ed., pp. 359–370). Routledge.
<https://doi.org/10.4324/9781351034784-38>

Sharif, H. (2008). Mångkultur, skola och interkulturell pedagogik. I Molin, M., Gustavsson, A. & Hermansson, H. (Eds.). *Meningsskapande och delaktighet: om vår tids socialpedagogik* (pp.105–107). Göteborg: Daidalos

Sporre, K., Osbeck, C., Lilja, A., Lifmark, D., Franck, O., & Lyngfelt, A. (2023). ‘Constructing the test to the teaching’ – and the complexities of assessing learning in ethics education. *Journal of Beliefs & Values*, 1–19. <https://doi.org/10.1080/13617672.2023.2271269>

Stier, J. (2021). Interkulturalitet i en skolkontext: Begrepp och definitioner. Skolverket. Retrieved from <https://larpportalen.skolverket.se/api/resource/P03WCPLAR160044>

Sue, D. W. (2001). Multidimensional facets of cultural competence. *The Counseling Psychologist*, 29(6), 790–821. <https://doi.org/10.1177/0011000001296002>

Trompenaars, F., & P. Woolliams. (2009). Research Application *Toward a General Framework of Competence for Today's Global Village* In D. K. Deardorff (Ed.), *The SAGE Handbook of Intercultural Competence* (pp. 438–455). Thousand Oaks, CA: Sage.

ASSESSING INTERCULTURAL COMPETENCE THROUGH SWEDISH NATIONAL TESTS OF SOCIAL SCIENCE: A VALIDATION STUDY

Markus Al-Afifi

Tykesson, J. (2017). Differential Item Functioning in the National Tests in Religious Education in Sweden. In O. Franck (Ed.), *Assessment in Ethics Education* (pp. 145–162). Springer International Publishing. https://doi.org/10.1007/978-3-319-50770-5_7

Ubani, M. (2013). Threats and solutions: Multiculturalism, religion and educational policy. *Intercultural Education*, 24(3), 195–210.
<https://doi.org/10.1080/14675986.2013.797701>

UNESCO. (2018). UNESCO Survey on Intercultural Dialogue 2017. Analysis of Findings. Paris: UNESCO

University of Gothenburg. (2022). *Nationellt prov I grundskolans årskurs 9. Provrappport Samhällskunskap Årskurs 9 Prov vårterminen 2022*.

U2011/6543/S, Uppdrag om nationella prov.

Zhang, X., & Zhou, M. (2019). Interventions to promote learners' intercultural competence: A meta-analysis. *International Journal of Intercultural Relations*, 71, 31–47. <https://doi.org/10.1016/j.ijintrel.2019.04.006>