Handbook of Research on Science Education

Volume II

Edited by

Norman G. Lederman

Sandra K. Abell

Routledge
Taylor & Francis Group
NEW YORK AND LONDON

2014

31

The Evolving Landscape Related to Assessment of Nature of Science

Fouad Abd-El-Khalick

Introduction

The centrality, ubiquity, and longevity of nature of science (NOS) in precollege science education cannot be overstated. The goal of helping precollege students internalize informed NOS understandings has been the subject of continuous and intensive research and development efforts around the globe since the early 1950s (see Lederman & Lederman, Chapter 30, this volume). This longstanding focus will endure into the future: NOS continues to be explicitly emphasized as a prominent curricular component and instructional goal in, for example, the most recent science education reform efforts in the United States embodied in the *Next Generation Science Standards* (NGSS Lead States, 2013). Syntheses of the research literature on NOS in science education provide for a robust narrative (e.g., Abd-El-Khalick & Lederman, 2000a; Lederman, 1992, 2007). This narrative speaks to gains in student NOS understandings in response to particular instructional interventions but asserts continued frustration with the prevalence of naïve NOS conceptions among a majority of precollege students. The lack of substantial progress is attributed to science teachers' naïve NOS conceptions; systemic issues with science teacher education, including the nature of teachers' scientific education in the academy; the predominance of a culture of school science instruction that is incommensurate with scientific practice, even among teachers who seem to have internalized informed NOS understandings; a host of situational and contextual factors that mediate the translation of teachers' NOS understandings into their practice; and naïve representations of NOS in commercial science textbooks and instructional resources (Abd-El-Khalick, Waters, & Le, 2008; see Lederman & Lederman, Chapter 30, this volume). Research efforts continue to be dedicated to understanding the relative importance of the aforementioned factors and how best to mitigate their effects, and to delineating the nature and development among teachers of pedagogical content knowledge specific to teaching about NOS (Wahbeh & Abd-El-Khalick, 2013). Additionally, albeit drawing on incongruent empirical bases, debates continue about the differential impacts of instructional interventions (e.g., explicit versus implicit; integrated versus nonintegrated) on students' and teachers' NOS understandings; research also is being directed toward gauging the most effective contexts (e.g., argumentation, authentic research apprenticeships, historical case studies, inquiry-oriented experiences, socioscientific issues) and those best suited for addressing varying sets of NOS–related objectives; and vigorous discussions persist in relation to the sources that inform the construct of NOS, in addition to (re)emergent discussions about the very nature of the construct, this time from the perspective of domain-general versus domain-specific NOS understandings (see Abd-El-Khalick, 2012a, 2013). The aforementioned narrative also speaks to a multitude of assessments, which have been developed and used to gauge learners' understandings: As many as two dozen NOS–specific instruments have come into being over the past 60 years (Lederman, 2007). The latter instruments and associated assessment approaches are the *empirical* content and object of investigation of the present chapter.

A focus on NOS assessments is crucial. First, it is well known that the nature and underlying dimensions of the construct of NOS in science education have been and continue to be contested (e.g., Abd-El-Khalick, 2012a; Alters, 1997; Bianchini & Solomon, 2003; Good & Shymansky, 2001; Irzik & Nola, 2011; Smith, Lederman, Bell, McComas, & Clough, 1997; Wong & Hodson, 2010). However, in a "real" and practical sense, the *only* NOS construct (or constructs) in currency in the field of science education is the construct (or are those constructs) being assessed. Research claims and judgments related to the status of students' and teachers' NOS understandings (naïve, informed, etc.), relative impact of differing instructional interventions on learners' understandings, and differential appropriateness of one instructional context compared to another for teaching about specific NOS aspects, all rest on the construct as embodied in assessments used by science education researchers. Thus, by scrutinizing the embodiment of NOS in various assessment instruments and approaches, one aim of this chapter is to assess the extent to which the NOS research domain is "strained" given the seemingly contested nature of the target of assessment (i.e., the construct of NOS itself) and the varied perspectives regarding the ideal context(s) within which these assessments are most meaningful, as well as the goal(s) of such assessments. Second, prior reviews of NOS–related assessments (e.g., Aikenhead, 1973, 1988; Doran, Guerin, & Cavalieri, 1974; Guerra-Ramos, 2012; Lederman, 2007; Lederman, Wade, & Bell, 1998; Pearl, 1974) have compiled increasingly longer lists of NOS instruments over the decades, outlined the general domains targeted by these instruments, examined the approaches and types of items used (Likert, multiple-choice, agree/disagree, open-ended, etc.) and the implications of such use for the trustworthiness of the assessments, and/or—where applicable—gauged the instruments' psychometric properties. All these aspects are quite important; this chapter draws on the literature to update such aspects and defers to the aforementioned reviews in several respects throughout its discussions. However, the dominant narrative we are left with is that the NOS research domain is inundated by the use of a large number of different instruments, which tend to focus on differing dimensions of NOS. Thus, the narrative continues; attempts to generalize across empirical studies or undertake robust comparisons across contexts and interventions will always be undermined by instrument variance. Again, in a "real" and practical sense, the extent to which any one instrument impacts discourse within a research domain and contributes to shaping knowledge claims within that domain is determined by the extent to which the instrument is *used* in empirical investigations. Thus, another aim of this chapter is to scrutinize the *extent of use* of various NOS instruments to generate an empirical account of the landscape related to NOS assessments and the evolution of this landscape.

The Landscape of NOS Assessments

The terrain related to NOS assessment is, to say the least, unwieldy. The following section outlines the difficulties associated with taking accurate stock of NOS–specific assessments and presents an explicit set of inclusion criteria, which were used to admit NOS instruments into the present review. The next section lists these instruments, describes the frequency and timeline for their development, and characterizes their types and general features.

Difficulties With Mapping the Field and Inclusion Criteria

Accurately depicting the NOS assessment landscape is a major challenge in and of itself. To start with, NOS in science education has been the subject of wide-ranging and continuous research efforts for close to six decades. A conservative search of assessments related to NOS using Google Scholar would return several thousand entries. Additionally, a plethora of NOS assessment instruments have been developed since the early 1950s (e.g., Wilson, 1954) and continue to be developed to the present day (e.g., Hacıeminoğlu, Yılmaz-Tüzün, & Ertepınar, 2012). Nonetheless, for a number of reasons, taking accurate stock of these instruments and assessments is not as straightforward as one might assume. First, the NOS construct itself has evolved over time. For instance, in the 1950s and early 1960s, several instruments bundled the assessment of cognitive, affective, and attitudinal outcomes related to the nature of the scientific enterprise, such as the Science Attitude Questionnaire (SAQ; Wilson, 1954), *Attitudes Toward Science and Scientific Careers* (ASSC; Allen, 1959), and *Inventory of Science Attitudes, Interests, and Appreciations* (ISAIA; Swan, 1966). However, by the early 1970s, the constructs of "scientific attitudes" (e.g., skepticism, open-mindedness) and "attitudes toward science" (e.g., interest in science, attitudes toward science and scientists) were being carefully delineated as distinct from outcomes associated with covering terms, such as "understandings about science" (see Gardner, 1975; Mackay & White, 1974). Thus, depending on their authors' perspectives, literature reviews of "assessments" or "measures" of NOS–related objectives end up with differing lists of specific instruments (see Aikenhead, 1973, 1988; Doran et al., 1974; Guerra-Ramos, 2012; Hacıeminoğlu et al., 2012; Lederman, 2007; Lederman et al., 1998; Mayer & Richmond, 1982; Pearl, 1974). Second, over the years, many researchers have adapted, modified, and recombined extant NOS instruments in a variety of ways and to varying extents to examine learners' NOS understandings. For example, Abd-El-Khalick and BouJaoude (1997) and Dogan and Abd-El-Khalick (2008) slightly modified the 14 items they selected from the *Views of Science-Technology-Society* (VOSTS) instrument (Aikenhead & Ryan, 1992; Aikenhead, Ryan, & Fleming, 1989) by deleting the last three standard choices for each *VOSTS* item ("I don't understand," "I don't know enough about this subject to make a choice," and "None of these choices fits my basic viewpoint") and replacing these with a choice that allowed respondents to articulate whatever ideas they deemed were representative of their views on the target issue. The authors, however, left item stems and positions virtually intact. In comparison, Haidar (2000) developed a questionnaire that drew heavily on but substantially modified *VOSTS* items by regrouping the multiple positions associated with each item into only three positions (for similar modifications, see Kang, Scharmann, & Noh, 2005). In other cases, researchers used one subset or another of items from extant instruments (or even combined item subsets from different instruments), sometimes with slight modifications and, on occasion, with major modifications coupled with renaming the resulting "instrument." For instance, Wong, Hodson, Kwan, and Yung (2008) used a slightly modified version of the *Views of Nature of Science–Form C* (VNOS–C) Questionnaire (Lederman, Abd-El-Khalick, Bell, & Schwartz, 2002), while Liu and Lederman (2007) adapted seven items from the *VNOS–C,* added another item, and recast the resulting instrument as the "Nature of Science Questionnaire" (p. 1287). Other instrument modifications were more involved. For instance, Rampal (1992) adapted a number of items from the *Test on Understanding Science* (TOUS; Cooley & Klopfer, 1961) "both linguistically and culturally to incorporate the popular positions teachers . . . [in the context of India] tend to take on such issues" (p. 418). Next, modified items were coupled with a set of eight open-ended questions, which focused on perceptions of scientists' areas of work, "minimum educational requisites to be a scientist," "women in science, the role of science in removing poverty and unemployment, [and] the role of the computer in current scientific research" (pp. 418–419), and the two-part instrument was named the "Views about Science and Scientists" questionnaire. The result of such modifications, both modest and elaborate, is a rather large number of instrument *variants* (see also Buffler, Lubben, & Ibrahim, 2009).

A third reason underlying the difficulty of taking accurate stock of NOS–specific assessments is that a number of researchers have developed NOS–related measures and assessments, which were specific to their purposes in the context of particular studies (e.g., among others, Abd-El-Khalick & BouJaoude, 2003; Bianchini & Solomon, 2003; Driver, Leach, Miller, & Scott, 1996; Flores, Lopez, Gallegos, & Barojas, 2000; Gess-Newsome, 2002; Guerra-Ramos, Ryder, & Leach, 2010; Lakin & Wellington, 1994; Lombrozo, Thanukos, & Weisberg, 2008; Mellado, 1997; Murcia & Schibeci, 1999; Rubba, Horner, & Smith, 1981; Ryder, Leach, Driver, 1999; Turgut, 2011; Vesterinen & Aksela, 2012; Zimmermann & Gilbert, 1998). These assessments mostly were of the semistructured or open-ended individual interview types but also included scenario-based items or interviews, convergent and open-ended questionnaires, repertory girds, learner-generated artifacts (e.g., reflective journals, diaries), and combinations of these variants. Nonetheless, most of these assessments were not developed as or intended to be formal "instruments." In many cases, efforts to establish the validity and/or reliability of these measures were limited to establishing face validity or securing feedback from reviewer panels (see Munby, 1982, for issues associated with the overreliance on such approaches in instrument validation). Additionally, other researchers did not "take up" these measures or put them to use in empirical investigations. Fourth, even though referenced in some reviews or studies, a few instruments could not be retrieved. For instance, Trembath (1972) reported on the development and use of an 18-multiple-choice-item questionnaire to assess prospective elementary teachers' NOS views. Mayer and Richmond (1982) noted that "Richardson and Showalter (1967) authored *The Abridged Scientific Literacy Instrument*" (p. 54), which they listed among instruments designed to measure understandings of the nature of the scientific enterprise. Similarly, Walls (2012) listed the "Science Inventory" (Hungerford & Walding, 1974) in his review of NOS instruments. However, the instruments were not included in the original reports, could not be located and retrieved, were not cited in other literature, and/or were not used in other empirical studies that examined NOS–related objectives. Finally, tracking the use of some NOS measures in empirical studies was challenging because the developers did not title their instruments (e.g., Wilson, 1954), which were given slightly varied titles or labels throughout the years, while other instruments were given labels or acronyms that differed from those put forth by their originators. For example, Doran and colleagues (1974) labeled the instrument originally developed by Schwirian (1968) the "Science Support Scale" (p. 321) and used the acronym "SSS" to refer to it (p. 322). However, Schwirian (1969) had named the instrument the *Schwirian Science Support Scale* and abbreviated it as the *TRI-S* (p. 204). To provide as accurate an account as possible, this review referred to original publications of instruments for names, acronyms, domains, and other information. In case the developer(s) of an instrument did not provide an original name and/or acronym, the name provided was kept as close to the developer(s)' characterization of the measure as possible. Additionally, as evident in Table 31.1, original instrument names and acronyms were clearly distinguished—and appear italicized in the table—from those introduced at a later time by other researchers.

It could be seen that mapping the terrain related to NOS assessments in a meaningful and rigorous manner is challenging, to say the least. Thus, an explicit set of inclusion criteria was developed and applied both to admit instruments into the present examination and to count "instances of their use." First, albeit a serious limitation, only literature with full text (as compared, for example, to an abstract) available in English was consulted. Second, empirical studies published in refereed journals were admitted: Dissertations and theses, books and book chapters, conference papers and proceedings, ERIC documents, monographs, and other unpublished papers were excluded from the domain of review. The exception to this second criterion was applied to cases in which the development of a NOS instrument or the single instance of its use was not available as a refereed journal article (e.g., Korth, 1968, 1969; Lederman & Ko, 2002; Welch, 1966a). Drawing on peer-reviewed journal publications ensures the rigor of the research consulted and robustness of the conclusions drawn from the present review. Third, as noted, while it is well understood that early instruments lumped cognitive and affective NOS–related outcomes together, evidence suggests that by the early 1970s, outcomes related to scientific attitudes and attitudes toward science and scientists (including interests and appreciations) were clearly demarcated from understandings about science, scientific

TABLE 31.1
(An Incomplete) List of Nature of Science Instruments (1954–2012)

Date	Author(s)	Instrument	Abbreviation	Dimensions/categories/aspects	Items
1954	Wilson	Science Attitude Questionnaire[1]	SAQ[1]	Attitudes/perceptions toward scientists, science, and scientific products; ideas on methods, social aspects, tentativeness, and objectivity in science	26 agree/disagree
1957	Mead & Métraux	Images of Science, Scientists, and Scientific Careers[1]	ISSSC[1]	Attitudes toward scientists and scientific careers (vis-à-vis respondents' sex)	3 incomplete sentences (varied by sex) prompt essay responses
1958	Stice	*Facts About Science Test*	*FAS*	Attitudes/perceptions of scientists, science as an institution, and societal impacts of science	80 3-alternative multiple choice
1959	Allen	*Attitudes Toward Science and Scientific Careers*	ASSC[1]	Reciprocal impacts of science and society, the scientist, scientific work, and nature of science	95 5-point Likert (CA, A, N, D, TD)[2]
1961	Cooley & Klopfer	*Test on Understanding Science (From W)*	*TOUS*	Scientific enterprise, the scientist, and methods and aims of science	53 4-alternative multiple choice + 7 2-part (statement/reason)
1966	Swan	*Inventory of Science Attitudes, Interests, and Appreciations*	*ISAIA*	Affective outcomes related to scientific attitude, appreciations, interest, and skills	50 agree/disagree + 21 skill achievement experiences
1966	Welch	*Welch Science Process Inventory (Form D)*	*SPI*	Activities, assumptions, products, and ethics of science	135 agree/disagree
1967	Scientific Literacy Research Center	*Wisconsin Inventory of Science Processes*	*WISP*	Assumptions, activities, objectives, and products of science	93 3-alternative: Accurate, Inaccurate, Don't know/understand
1967–1968	Kimball	*Nature of Science Scale*	*NOSS*	Science is: driven by curiosity, empirical, reproducible, process oriented, open, dynamic, parsimonious, operational, tentative/uncertain, imposes human constructs on nature, and has no one "scientific method"	29 3-point Likert (A, N, D)[2]
1968	Schwirian	*Schwirian Science Support Scale*	*TRI-S*	Attitudes toward rationality, utility-arianism, universalism, individualism, and belief in progress and meliorism	40 5-point Likert (SA, A, U, D, SD)[2]
1968	Korth	*Test of the Social Aspects of Science*	*TSAS*	Interactions among science, technology, and society; social nature of scientific enterprise; social responsibilities of science and scientists	52 5-point Likert (SD, D, U, A, SD)[2]
1974	Aikenhead	*A Measurement of Knowledge About Science and Scientists (Project Physics: Form 1)*	*KASSPP*[1]	(See TOUS and SPI)	101 items (derived from *TOUS* and *SPI*)
1975	Billeh & Hasan	*Nature of Science Test*[3]	*NOST*	Assumptions, products, processes, and ethics of science	60, 4-alternative multiple choice
1975	Hillis	*Views of Science*	VOS[1]	Tentative NOS	40 5-point Likert (SA, A, N, D, SD)[2]
1976	Rubba	*Nature of Scientific Knowledge Scale*	*NSKS*	Science is amoral, creative, developmental, parsimonious, testable, and unified	48 5-point Likert (SA, A, N, D, SD)[2]
1977	Billeh & Malik	*Test on Understanding the Nature of Science*[3]	*TUNS*	Assumptions, processes, and ethics of science, and scientific enterprise	55 4-alternative multiple choice
1981	Cotham & Smith	*Conceptions of Scientific Theories Test*	*COST*	Ontological implications of theories, testing of theories, generation of theories, and theory choice	40 4-point Likert (SA, A, D, SD)[2] related to 4 theories
1982	Ogunniyi	*Language of Science*	*LOS*	Definitions, characteristics, formation, and function of concepts, laws, and theories	68 3-point Likert (A, D, Don't Know)[2]
1987	Johnson & Peeples	Methods and Nature of Science[1]	MaNS[1]	Theories and laws; limitations of science and observation; tentativeness; experimental methodology	15 5-point Likert + 5 items on acceptance of evolutionary theory) (SA to SD)[2]
1987	Aikenhead et al.	*Views of Science-Technology-Society*	*VOSTS*	Definitions (science and technology); external sociology of science (reciprocal influences of technology/science on society, and school science on society); internal sociology of science (characteristics of scientists, social construction of scientific knowledge and technology); epistemology (nature of scientific knowledge)	114 multiple choice, with 3 standard alternatives under each item: "I don't understand," "I don't know enough . . . to make a choice," and "None of these choices fits my basic viewpoint"

1988	Koulaidis & Ogborn	Views about Philosophy of Science[1]	VaPS[1]	Scientific method, criteria of demarcation, patterns of growth, and status of scientific knowledge	6 2-alternative multiple choice + 23 agree/disagree statements (grouped as 10 items)
1990	Lederman & O'Malley	Views of Nature of Science–Form A[1,5]	*VNOS-A*	Tentative NOS	7 open-ended items + follow-up interviews
1992	Meichtry	*Modified Nature of Scientific Knowledge Scale*	*MNSKS*	Creative, developmental, testable, and unified NOS	32 5-point Likert (SA, A, N, D, SD)[2]
1993	Pomeroy	Beliefs about Nature of Science and Science Education[1]	BNSSE[1]	Creativity, intuition, and cultural-embeddedness in science, myth of the "scientific method" and determinism, and limitations of observation	30 5-point Likert (SA to SD)2 (+ 20 items related to beliefs about science education)
1995	Nott & Wellington	*Critical Incidents*	CI[1]	Experimental and social procedures of science; moral, ethical, and political issues related to science and scientists; limits and tentativeness of science; theory-laden NOS; and explanatory function of theories	13 classroom scenarios or events; ask teachers what would, could, and should they do in response to incidents
1997	Aldridge et al.	*Beliefs About Science and School Science Questionnaire*	*BASSSQ*	Processes of scientific inquiry, and certainty of scientific knowledge	20 5-point Likert (AN, S, ST, O, AA[4]) (+ 21 items related to beliefs about school science)
1998	Abd-El-Khalick et al.	*Views of Nature of Science–Form B*[5]	*VNOS-B*	Empirical, inferential, creative, tentative, and theory-laden NOS; and nature and relationship between theories and laws	7 (2 contextual) open-ended + follow-up interviews
2000	Abd-El-Khalick & Lederman	*Views of Nature of Science–Form C*[5]	*VNOS-C*	Empirical, inferential, creative, tentative, and theory-laden NOS; nature and relationship between theories and laws; lack of a single "scientific method"; social and cultural embeddedness of science	10 (3 contextual) open-ended + structured follow-up interviews
2002	Abd-El-Khalick	*Perspectives on Scientific Epistemology*	*POSE*	Empirical, inferential, creative, tentative, and theory-laden NOS; nature and relationship between theories and laws; generation of scientific knowledge; lack of a single "scientific method"	10 (4 contextual) open-ended + follow-up interviews
2002/ 2004	Lederman & Khishfe Lederman & Ko	*Views of Nature of Science–Form D/Form E*[5]	*VNOS-D/ VNOS-E*	Empirical, inferential, creative, theory-laden, and tentative NOS	7 (3 contextual) open-ended, administered as interview or survey
2005	Tsai & Liu	*Scientific Epistemological Views*[6]	*SEVs*	Social negotiation; invented/creative, theory-laden, and changing/tentative NOS; cultural impacts	19 5-point Likert (SA, A, N, D, SD)[2]
2006	Chen	*Views on Science and Education Questionnaire*	*VOSE*	Tentativeness; nature of observation; scientific methods; hypotheses, theories, and laws; imagination; validation of scientific knowledge; objectivity and subjectivity in science	69 5-point Likert (SD, D, U/NC, A, SA)[2] statements grouped in 13 items (+ 16 statement under 2 items related to science education)
2006	Liang et al.	*Student Understanding of Science and Scientific Inquiry*	*SUSSI*	Observation and inference; change of scientific theories; laws versus theories; social-cultural influences; and imagination and creativity in and methodology of investigation	24 5-point Likert (SD, D, U, A, SA)[2] grouped in 6 items + invitation to explain responses with examples per item
2009	Buaraphan	*Myths of Science Questionnaire*	*MOSQ*	Tentative, creative, theory-laden, social, and social-cultural NOS; relationship between hypotheses, theories and laws; lack of single "scientific method" and limitations of science; science and technology	14 3-point Likert (A, U, D)[2]
2012	Hacıeminoğlu et al.	*Nature of Science Instrument*	*NOSI*	Empirical, inferential, tentative, creative, and theory-laden NOS	13 3-point Likert (Wrong/ Do not know/Right)

[1] Instrument name or abbreviation introduced later and not presented in the original publication. Original instrument names and abbreviations are italicized.
[2] Likert scales: Completely/Strongly Agree (CA/SA), Agree (A), Neutral/Undecided (N/U), Disagree (D), Completely/Strongly Disagree (CD/SD).
[3] Overlap between the *NOST* and *TUNS* could not be determined as the instruments could not be located. Available sample items suggest some differences.
[4] This particular scale corresponds to: Almost never, Seldom, Sometimes, Often, Almost always.
[5] All forms of the *VNOS* were counted as a single instrument. Also note that the *VNOS–D* and *VNOS–E* are virtually identical instruments. Lederman and Ko (2002) slightly modified and relabeled the *VNOS–D* as *VNOS–E*, whereby the "E" is meant to signal the instrument's intended audience (i.e., elementary students) than to indicate the development of a new version of the *VNOS*.
[6] Tsai (2013, personal communication, January 13, 2013) advised that the *Pupils' Nature of Science Scale* (Huang, Tsai, & Chang, 2005) is subsumed under the *Scientific Epistemological Views* (Tsai & Liu, 2005).

knowledge, and the scientific enterprise. For instance, Gardner (1975) included the *Scientific Attitudes Inventory* (Moore & Sutman, 1970) in his literature review of studies about "attitudes to science." Akindehin (1988) used the *Nature of Science Scale* (NOSS; Kimball, 1967–1968) and *Test of Science-Related Attitudes* (TOSRA; Fraser, 1978), both of which appear as NOS–assessment instruments in some prior reviews. Akindehin, nonetheless, used the *NOSS* to measure his participants' "understandings of the nature of science" in clear contradistinction of using the *TOSRA* to assess their "science-related attitudes" (1988, p. 77). For a similar treatment of the *TOSRA* and the *Tests of Perception of Scientists and Self* (TOPOSS; Mackay & White, 1974) as instruments designed to assess affective outcomes related to attitudes toward science, the reader is referred to Laforgia (1988; also see Osborne, Simon, & Collins, 2003). To be sure, this review does *not* adopt a revisionist perspective on the history of the field. Nonetheless, there was need to strike a balance between accounting for the historical development of NOS assessments while remaining faithful to contemporaneous developments in conceptualizations of the construct of NOS in science education. Thus, instruments that purported to measure science-related attitudes and interests and other affective outcomes published after 1970 were not included as NOS instruments. In this regard, it should be noted that an examination of the literature indicates that a conflation between NOS and science process skills, while evident in a few early cases, has not been as prevalent in the literature of the 1960s and early 1970s as with the conflation between NOS and attitudinal-related outcomes. For instance, while Doran and colleagues (1974) included the *Processes of Science Test* (POST; Biological Sciences Curriculum Study, 1962) in their analysis of NOS instruments, contemporaneous reviews (e.g., Aikenhead, 1973; Mackay & White, 1974) clearly excluded the *POST* from the domain of NOS measures. A content analysis of the *POST* justifies this exclusion: Only 6 of the 40 *POST* items touch on issues invoked in contemporaneous instruments, such as the *TOUS* (Cooley & Klopfer, 1961), *Welch Science Process Inventory* (SPI; Welch, 1966b) and *Wisconsin Inventory of Science Processes* (WISP; Scientific Literacy Research Center, 1967), which clearly addressed NOS–related objectives. Additionally, an examination of the purposes for which researchers had used the *POST* (i.e., Anderson & Callaway, 1986; Anderson, DeMelo, Szabo, & Toth, 1975; Faryniarz, 1992; Orgren & Doran, 1975; Riban, 1976; Riban & Koval, 1971; Rivers & Vockell, 1987; Starr, 1972) indicates that science process skills were being measured as distinct from understandings about science. For example, the *POST* was used to assess students' "general problem solving" abilities (Rivers & Vockell, 1987, p. 408), and "ability to recognize and apply the processes of science" (Orgren & Doran, 1975, p. 17), which included recognizing appropriate tabular representations of a specific dataset and drawing conclusions from descriptions of specific experimental setups. Thus, instruments that purported to measure instructional outcomes related to science process skills were excluded from the domain of the present review.

Fourth, as noted above, various modifications of many extant instruments abound, including the *TOUS* (e.g., Aikenhead, 1974a; Fisher & Fraser, 1980; Rampal, 1992; Welch, 1972), *VOSTS* (e.g., Bennett & Hogarth, 2009; Kang et al., 2005; Mbajiorgu & Ali, 2003; Yalvac, Tekkaya, Cakiroglu, & Kahyaoglu; 2007), *Nature of Scientific Knowledge Scale* (NSKS; Rubba, 1976, 1977; i.e., Meichtry, 1992), and *VNOS–C* (e.g., Akerson, Hanson, & Cullen, 2007; Rudge, Cassidy, Fulford, & Howe, 2013; Wong & Hodson, 2009, 2010). Nonetheless, modifications rarely entailed a substantial reconsideration or reconceptualization of the NOS models underlying the modified instruments. Additionally, in most cases, modifications of extant instruments were not coupled with rigorous efforts to re-establish the validity and/or reliability of the reconstituted measure once the integrity of original instruments was breached by selecting subsets of their items, rewording or translating these items, or altering responses for convergent items in one fashion or another. Exceptions included Aikenhead's (1974a) *A Measurement of Knowledge About Science and Scientists (Project Physics: Form 1*; KASSPP1) and Meichtry's (1992) *Modified Nature of Scientific Knowledge Scale* (MNSKS). Thus, in the overwhelming majority of cases, instrument modifications (including translating and contextualizing instruments, e.g., Dogan & Abd-El-Khalick, 2008; Haidar, 2000) were *not* considered distinct or new instruments. Instead, modifying and using an existing instrument in an empirical study was considered another instance of use of the original instrument. Finally, a few instruments have been developed to gauge relationships between learners' understandings of specific elements of NOS and other dimensions. For example, Cobern and Loving (2002) developed the "Thinking about Science" survey instrument, which aimed to assess the relationship of respondents' views of science to the economy, environment, public policy, religion, aesthetics, race and gender, and science for all, as well as epistemology. The latter dimension focused on views of the objectivity and certainty of scientific knowledge, as well as reliability of scientific methods. Using two questionnaires, Liu and Lederman (2007) aimed to describe their participants' "views of NOS and worldviews, and to explore relationships, if any, between these two domains of beliefs" (p. 1286). Similarly, Windschitl (2004) used interviews, among other data sources, to examine his participants' folk theories of inquiry, which entailed elucidating their understandings of some NOS aspects. Guerra-Ramos (2012) included the latter three instruments and approaches in her review of measures used to elicit teachers' NOS ideas. However, these studies assessed conceptions of widely agreed-upon NOS aspects targeted by other instruments with the aim of exploring relationships with conceptions of other constructs, such as inquiry and worldviews, which are clearly distinct from—albeit related—to NOS (see Lederman, 2007). These and similar instruments, thus, were not considered distinct or new instruments when it comes to assessing NOS understandings.

(An Incomplete) List of NOS Assessments: Characterizing the Landscape

The application of the aforementioned inclusion criteria to the NOS assessment landscape resulted in the list of 32 NOS–specific instruments shown in Table 31.1. In this table, it should be noted, the different *VNOS* forms (Lederman et al., 2002) are counted as a single instrument. Given the arguments presented in the preceding section, the list is justifiably and understandably incomplete: Some researchers who had developed a NOS instrument in a language other than English, translated and contextualized an existing instrument to fit their educational and/or cultural milieu, substantially modified an existing instrument to reflect certain nuances in their perspectives on NOS or educational research, or reworked one or more subcomponents of an existing instrument or instruments into a new measure or one with broader goals might find that the list in Table 31.1 is missing one or a few instruments. Nonetheless, to the extent that the explicit inclusion criteria presented are sensible, based on reasonable assumptions, and derived from a careful and critical examination of both the empirical literature and prior literature reviews, I contend that the present list is both robust and accurate. To be sure, it is difficult to imagine that including a few more instruments in the present review would alter the historical narrative, analyses, and resulting conclusions in any substantial manner.

Table 31.1 indicates that exploring and assessing NOS understandings among precollege and college students (e.g., Kang et al., 2005; Mead & Métraux, 1957; Walls, 2012; Welch, 1969; Wood, 1972) and preservice and inservice teachers (e.g., Billeh & Hasan, 1975; Yalçinoğlu & Anagün, 2012; Zoller, Donn, Wild, & Beckett, 1991b), as well as science teacher educators and scientists (e.g., İrez, 2006; Kimball, 1967–1968; Wong & Hodson, 2010), has been a consistent and ongoing domain of interest throughout the past six decades. Each of these decades has featured the development of four to seven measures, with an average of about five NOS instruments developed per decade. Indeed, between 1954 (the year the first NOS instrument, the SAQ, was published) and 2013, there only is a single 5-year period (1969–1972) in which no new instruments were developed, compared to the two to three NOS instruments developed during all other 5-year periods and up to four instruments developed in the 5-year period from 1974 to 1978 (also see Figure 31.1).

Of the 32 instruments in Table 31.1, a small minority (4 or 12.5%) is of the open-ended generative type, where learners are asked to articulate their ideas in response to open-ended questions, scenarios, or prompts. In comparison, the overwhelming majority of instruments (28 or 87.5%) are of the forced-choice type, where respondents select an answer or indicate a preference from among a predetermined set of options. The 17 instruments using three-, four-, or five-point Likert-type items account for 61% of all forced-choice NOS assessments, with an additional 28% (8 instruments) comprising multiple-choice items and 11% (3 instruments) featuring agree/disagree type items. Most Likert-type instruments (70%, 12 instruments) use a five-point scale of the sort: strongly/completely agree, agree, undecided/neutral/no comment, disagree, and strongly/totally disagree, with the exception of the *Beliefs About Science and School Science Questionnaire* (BASSSQ; Aldridge et al., 1997; see Table 31.1). One instrument (6%) used a four-point and an additional four instruments (24%) used three-point Likert scales with a variety of descriptors, such as agree, neutral, disagree; and accurate, inaccurate, don't know/don't understand (e.g., the *WISP*).

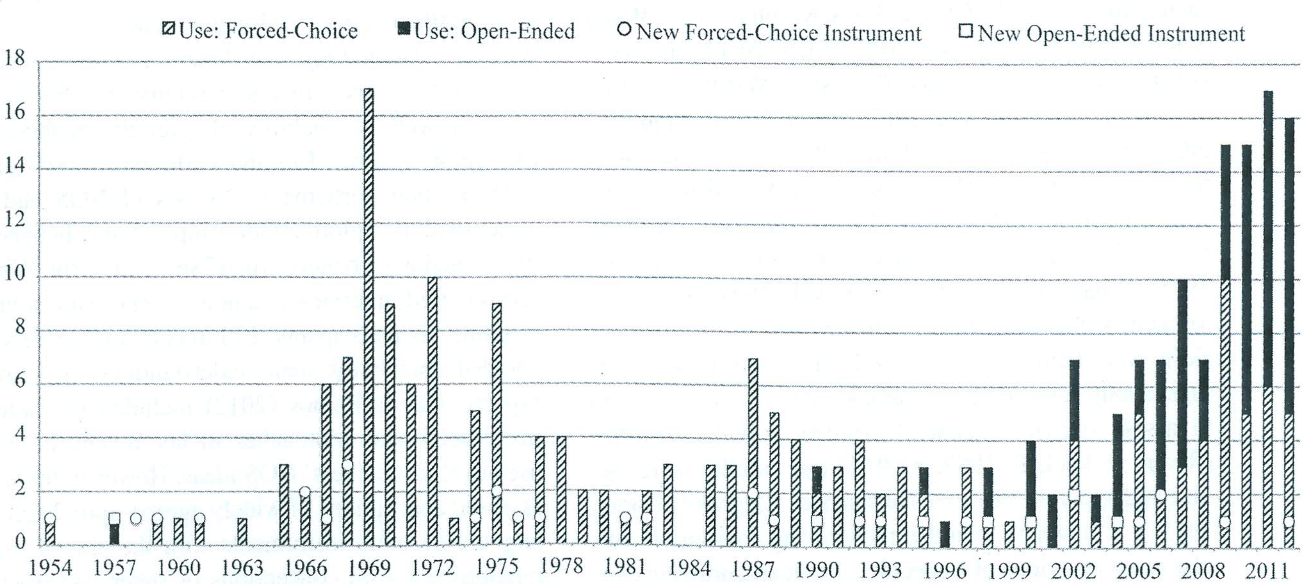

Figure 31.1 Frequency of use of forced-choice and open-ended NOS instruments in empirical studies (1954–2012).

The Evolving Landscape of NOS Assessments

An examination of the landscape of NOS assessment from a perspective focused on the number, frequency of development, and type of NOS instruments could be misleading. Indeed, data presented in the preceding section could support a number of inferences, such as that forced-choice assessments have dominated and continue to dominate the field; that the few open-ended instruments seem to have played a relatively less substantial role than forced-choice measures in gauging what learners understand about NOS; that a distinct pattern in the development of NOS assessments is lacking given that the field seems consistently preoccupied with the development of forced-choice measures—particularly of the Likert-type variety—with intermittent attempts to develop and use open-ended measures; or that the field has been and continues to be in disarray with regard to the construct of NOS as evident by the continuous development of instruments over the past six decades. Albeit consistent with the data provided in Table 31.1, such inferences, it will become evident, do not accurately capture the evolving nature of the NOS assessment field. Indeed, a very different and robust narrative emerges when the domain is examined from alternative lenses. The following sections examine the field's development from the perspective of the *use* of NOS assessments in empirical studies and explicate ways in which the alternative narrative that emerges from such examination is consistent with theoretical and conceptual advances underlying the very approaches to gauging learners' NOS understandings.

Use of NOS Assessments in Empirical Studies

While theoretical and conceptual debates about a specific definition for NOS continue or emerge at different points in time, it was argued earlier that, from a practical perspective, the only construct (or constructs) in currency in the field of science education is the NOS construct or are those constructs being assessed. Hence the significance of examining NOS assessments. Additionally, the extent to which assessments impact discourse, claims, and judgments in the field about progress, or lack thereof, toward achieving the much-desired NOS–related outcomes hinges on the extent to which these instruments are used in empirical studies. Nonetheless, the notion of "use" should not be understood as a mere quantification of an instrument's popularity. This notion is based on the assumption that researchers who opt to use a certain NOS instrument from among a number of instruments available at any one time have carefully examined these measures and (a) are in general agreement with the NOS model underlying their instrument of choice and (b) endorse the assessment approach embodied in this instrument, which they consider to be commensurate with the state of the art in terms of NOS assessment at the time of the study. It should not escape us that using a certain instrument might also be dictated, at least partially, by practical considerations. For instance, the use of an open-ended NOS assessment entails substantially more resources and effort in terms of collecting, analyzing, and interpreting data when compared to a forced-choice instrument. However, it is safe to assume that researchers would not use a NOS measure, which is incommensurate with their conceptions of the construct of NOS or which they consider would not generate a valid and reliable assessment of their participants' conceptions of this construct. In a sense, examining the use of various instruments would generate an empirical profile of the field's *collective wisdom* as it relates to NOS assessment.

A thorough review of the literature in the period from 1954 to 2013 resulted in a total of 241 empirical research studies published—with very few exceptions—in refereed journal articles, which used the instruments listed in Table 31.1. The exceptions included cases when the single use of an instrument by its developer(s) was not published as a peer-reviewed article, such as with the *Test of the Social Aspects of Science* (TSAS; Korth, 1969). The reader is reminded that, as explicated in the aforementioned inclusion criteria, use of a modified instrument in an empirical study was mostly counted as an instance of use of the original instrument (see previous discussion for the few exceptions). Nonetheless, while surely useful, the simple enumeration of instances of use would not have been sufficient to accurately depict the impact of different instruments. In particular, the use of an instrument—be it in its original or modified form—by researchers other than its original developer(s) is an important indicator of its impact. Indeed, some instruments have been used in a substantial number of studies. All these uses, however, were undertaken by the instrument developer(s). *The Scientific Epistemological Views* (SEVs) is a case in point: The *SEVs* was developed by Tsai and Liu (2005) and since then has been used in as many as eight empirical studies that were, nonetheless, undertaken by Tsai and his colleagues. Thus, while there is need to acknowledge and celebrate the work of prolific researchers, it was important to keep this limitation in mind. Secondly, longevity adds an important dimension to an instrument's impact, because it makes possible comparing investigations undertaken in different contexts over a period of time, which enable drawing meaningful and robust conclusions about the status of the field. At the same time, it was important not to overly depend on the longevity dimension given that many instruments have been in existence for substantial periods of time, which would disadvantage more recently developed instruments.

An "index of use" score was generated to account for these dimensions when examining the use of NOS instruments (see Table 31.2). This score was calculated for each instrument listed in Table 31.1 by assigning the instrument one point for each use in an empirical study undertaken by the instrument developer(s), three points for each use by researchers other than the developer(s) within the first 5 years following the instrument's publication, and five points for each use in an empirical study by other researchers beyond the

TABLE 31.2
Nature of Science Instruments Ranked by the Extent of Use in Empirical Research Studies

Instrument	Empirical studies	Used by			Index of use score[1]
		Author(s)	Others		
			≤ 5 yrs	> 5 yrs	
VNOS (Views of Nature of Science) Questionnaire		16	5	41	236
VNOS-A	Lederman & O'Malley (1990)	1	0	0	1
VNOS-B	Abd-El-Khalick & Akerson (2004); Abd-El-Khalick, Bell, & Lederman (1998); Akerson & Abd-El-Khalick (2003); Akerson, Abd-El-Khalick, & Lederman (2000); Akerson & Hanuscin (2007); Akerson & Volrich (2006); Bell, Lederman, & Abd-El-Khalick (2000)	5	0	2	15
VNOS-C	Abd-El-Khalick (2001, 2005); Abd-El-Khalick & Akerson (2009); Abd-El-Khalick & Lederman (2000b); Akçay & Koç (2009); Akerson, Hanson, & Cullen (2007); Borda, Burgess, Plog, DeKalb, & Luce (2009); Cakmakci (2012); Duncan & Arthurs (2012); Eastwood et al. (2012); Goff, Boesdorfer, & Hunter (2012); Hanuscin, Akerson, & Phillipson-Mower (2006); Howe (2007); Howe & Rudge (2005); Irez (2006, 2007); Kaya (2012); Kim & Nehm (2011); Kucuk (2008); Lederman, Schwartz, Abd-El-Khalick, & Bell (2001); Liu & Lederman (2007); Marchlewicz & Wink (2011); McDonald (2010); Nalçaci, Akarsu, Kariper (2011); Pegg & Gummer (2010); Posnanski (2010); Roehrig & Luft (2004); Russell & Aydeniz (2012); Russell & Weaver (2011); Salloum & Abd-El-Khalick (2010); Schwartz & Lederman (2002); Schwartz, Lederman, & Crawford (2004); Schwartz, Westerlund, Garcìa, & Taylor (2010); Thye & Kwen (2004); Urhahne, Kremer, & Mayer (2011); Walker & Zeidler (2007); Wong & Hodson (2009, 2010); Wong, Hodson, Kwan, & Yung (2008); Yalçinoğlu & Anagün (2012)	10	3	27	154
VNOS-D	Akerson, Buck, Donnelly, Nargund-Joshi, & Weiland (2011); Akerson, Cullen, & Hanson (2009); Akerson & Donnelly (2010); Akerson & Hanuscin (2007); Akerson & Volrich (2006); Hanuscin & Lee (2011); Leblebicioğlu, Metin, Yardimci, & Berkyürek (2011); Leblebicioğlu, Metin, Yardimci, & Cetin (2011); Metin & Leblebicioğlu (2011, 2012); Quigley, Pongsanon, & Akerson (2010, 2011)	0	2	10	56
VNOS-E	Demirbaş & Balci (2012); Walls (2012)	0	0	2	10
TOUS	Aikenhead (1974b); Anderson (1970); Broadhurst (1970); Choppin (1974); Cossman (1969); Crumb (1965); Durkee (1974); Engen, Smith, & Yager (1967–1968); Fisher (1979); Fisher & Fraser (1980); Fraser (1978); Fulton (1971a, 1971b); Ginns & Foster (1978); Glass & Yager (1970); Johnson (1972); Jones (1969); Jungwirth (1968, 1969); Jungwirth & Jungwirth (1972); Klopfer & Cooley (1963); Krockover (1971); Lavach (1969); Lowery (1967); Mackay (1971); Meyer (1969); Olstad (1969); Rampal (1992); Riley (1979); Rothman (1969); Rothman, Welch, & Walberg (1969); Tamir (1994); Tamir & Jungwirth (1975); Trent (1965); Voss (1965); Walberg (1969); Welch (1969, 1972); Welch & Rothman (1968); Welch & Walberg (1967–1968, 1970, 1972); Yager (1968); Yager & Wick (1966); Zingaro & Collette (1967–1968)	1	5	39	211
VOSTS	Abd-El-Khalick & BouJaoude (1997); Aikenhead (1987, 1997); Aikenhead, Fleming, & Ryan (1987); Ben-Chaim & Zoller (1991); Bennett & Hogarth (2009); Botton & Brown (1998); Bradford, Rubba, & Harkness (1995); Choi & Cho (2002); Dass (2005); Dogan & Abd-El-Khalick (2008); Fleming (1987, 1988); Haidar (1999, 2000); Kang, Scharmann, & Noh (2005); Kokkotas et al. (2009); Lin & Chen (2002); Marbach-Ad et al. (2009); Mbajiorgu & Ali (2003); Rubba & Harkness (1993); Walczak & Walczak (2009); Yalvac, Tekkaya, Cakiroglu, & Kahyaoglu (2007); Zoller & Ben-Chaim (1994); Zoller, Donn, Wild, & Beckett (1991a, 1991b); Zoller et al. (1990)	5	4	18	107
SPI	Aikenhead (1974b); Anderson (1970); Breedlove & Gessert (1970); Haukoos & Penick (1983, 1985, 1987); Lawrenz (1975); Lawrenz & Cohen (1985); Markle & Capie (1977); Rothman (1969); Rothman, Welch, & Walberg (1969); Spears & Zollman (1977); Tamir (1972); Tamir & Jungwirth (1975); Walberg (1969); Welch (1969, 1972, 1980); Welch & Lawrenz (1982); Welch & Pella (1967); Welch & Rothman (1968); Welch & Walberg (1967–1968, 1970, 1972)	10	4	10	72
NSKS	Chan (2005); Dienye (1987); Folmer, Barbosa, Soares, & Rocha, (2009); Güzel (2011); Lederman (1986a, 1986b); Lederman & Druger (1985); Lederman & Zeidler (1987); Lin & Chiu (2004); Lonsbury & Ellis (2002); Rubba & Andersen (1978); Sutherland & Dennick (2002); Tasar (2006); Walker & Zeidler (2007); Zeidler & Lederman (1989)	1	0	14	71
NOSS	Akindehin (1988); Andersen, Harty, & Samuel (1986); Duschl & Wright (1989); Kimball (1967–1968); Ogunniyi (1983); Scharmann (1988a, 1988b, 1989, 1994); Scharmann & Harris (1992)	1	0	9	46
WISP	Carey & Stauss (1968, 1970); Lawson, Nordland, & DeVito (1975); Leake & Hinerman (1973); Markle & Capie (1977); Simpson & Wasik (1978); Wood (1972)	0	3	4	29

(Continued)

TABLE 31.2 (Continued)

Instrument	Empirical studies	Used by			Index of use score[1]
		Author(s)	Others		
			≤ 5 yrs	> 5 yrs	
FAS	Cooley & Bassett (1961); Cossman (1969); Fulton (1971a, 1971b); Glass & Yager (1970); Jacobs & Bollenbacher (1960)	0	2	4	26
POSE	Abd-El-Khalick (2002); Khishfe (2008, 2012); Khishfe & Abd-El-Khalick (2002); Seker & Welsh (2006); Yacoubian & BouJaoude (2010)	2	1	3	20
SUSSI	Fergusson, Oliver, & Walter (2012); Golabek & Amrane-Cooper (2011); Liang et al. (2008); Liang et al. (2009); Miller, Montplaisir, Offerdahl, Cheng, & Ketterling (2010); Shim, Young, & Paolucci (2010)	2	3	1	16
TRI-S	Schwirian (1968, 1969); Schwirian & Thomson (1972); Simpson, Shrum, & Rentz (1972); Spears & Hathaway (1975); Symington & Fensham (1976)	3	1	2	16
SEVs	Chai, Deng, & Tsai (2012); Huang, Tsai, & Chang (2005); Lin & Tsai (2008); Liu, Lin, & Tsai (2011); Liu & Tsai (2008); Oskay, Yilmaz, Dinçol, & Erdem (2011); Tsai (2007); Tsai & Liang (2009); Tsai & Liu (2005)	8	0	1	13
MOSQ	Buaraphan (2009a, 2009b, 2010, 2011, 2012); Buaraphan & Sung-Ong (2009); Sarkar & Gomes (2010)	6	1	0	9
MaNS	Johnson & Peeples (1987); Scharmann (1990); Scharmann & Harris (1992)	1	2	0	7
TSAS	Chrouser (1975); Korth (1969)	1	0	1	6
CI	Kim & Irving (2010); Nott & Wellington (1995, 1996, 1998)	3	1	0	6
BASSSQ	Aldridge, Taylor, & Chen (1997); İrez, Çakir, & Şeker (2011)	1	0	1	6
VOSE	Chen (2006b); Demirbas, Bozdogan, & Özbek (2012)	1	0	1	6
VOS	Barufaldi, Bethel, & Lamb (1977); Hillis (1975)	1	1	0	4
VaPS	Apostolou & Koulaidis (2010); Koulaidis & Ogborn (1988, 1989, 1995)	4	0	0	4
LOS	Ogunniyi (1982, 1983)	1	0	0	1
SAQ	Wilson (1954)	1	0	0	1
ISSSC	Mead & Métraux (1957)	1	0	0	1
ASSC	Allen (1959)	1	0	0	1
ISAIA	Swan (1966)	1	0	0	1
KASSPP1	Aikenhead (1974a)	1	0	0	1
NOST	Billeh & Hasan (1975)	1	0	0	1
TUNS	Billeh & Malik (1977)	1	0	0	1
COST	Cotham & Smith (1981)	1	0	0	1
MNSKS	Meichtry (1992)	1	0	0	1
BNSSE	Pomeroy (1993)	1	0	0	1
NOSI	Hacıeminoğlu, Yılmaz-Tüzün, & Ertepınar (2012)	1	0	0	1

[1] "Index of use score" calculated by assigning an instrument 1 point for each use in an empirical study by the instrument developer(s), 3 points for each use by other researchers within the first 5 years following the development of the instrument, and 5 points for each use by other researchers beyond the 5-year period after the instrument development.

5-year period after which the instrument was developed. The 5-year cutoff derives from the aforementioned stipulation for not disadvantaging recently developed instruments given that many NOS instruments have been in existence for substantial periods of time. Consider the *VOSTS* (Aikenhead et al., 1987) as an example. The instrument was developed in 1987 and used in a total of 27 published studies, 5 of which were undertaken by Glen Aikenhead, Reg Fleming, and/or Alan Ryan, for which it received five points. Four more studies used the *VOSTS* in the first 5 years following its development, that is, between 1987 and 1992 (Ben-Chaim & Zoller, 1991; Zoller, Donn, Wild, & Beckett, 1991a, 1991b; Zoller et al., 1990), for which it was allotted an additional 12 points. Finally, researchers other than Aikenhead and his colleagues used the *VOSTS* in 18 additional studies starting in 1993 (see Table 31.2), for which the *VOSTS* received an additional 90 points to end up with an index of use score of 107 points. Table 31.2 lists NOS instruments ranked by their index of use score. It should be noted that uses of all forms of the *VNOS* (Lederman et al., 2002) were grouped together because the different forms share the same underlying model for NOS and approach to assessment (i.e., generic and context-based open-ended questions coupled with follow-up individual interviews) and differ only in the number, context, and phrasing of their items, which were designed to make them accessible to different populations, including science teachers and high school students (i.e., *VNOS–A* , *VNOS–B*, and *VNOS–C*), and elementary school students (i.e., *VNOS–D* / *VNOS–E*).

When examined from the use-of-instruments lens, the landscape of NOS assessment takes on different features. Most noteworthy is the observation that the seemingly large variance in assessments suggested by the proliferation of instruments over the past 60 years is *greatly* reduced. In terms of frequency of use in empirical studies, the top three instruments in Table 31.2—the *VNOS, TOUS,* and *VOSTS*—account for more than 50% of all instrument use in the past six decades. The preferential use of the *VNOS, TOUS,* and *VOSTS,* it will become evident, is closely associated with the field's evolution in terms of the different assumptions underlying approaches to assessing NOS understandings represented by each of the three instruments. Additionally, the top nine instruments in Table 31.2 account for more than 75% of all use of NOS measures: These instruments are the *VNOS* (Lederman et al., 2002), *TOUS* (Cooley & Klopfer, 1961), *VOSTS* (Aikenhead et al., 1987), *SPI* (Welch, 1966b), *NSKS* (Rubba, 1976), *NOSS* (Kimball, 1967–1968), *WISP* (Scientific Literacy Research Center, 1967), *Facts about Science Test* (FAS; Stice, 1958), and *Perspectives on Scientific Epistemology* (POSE; Abd-El-Khalick, 2002). The remaining 23 instruments account for 23% of all use in empirical studies, 17% of which were instances of use by the instruments' developers and 6% by other researchers.

Second, by employing the use of instruments as proxy for researcher sensibilities about accessing and gauging learner NOS understandings, the present approach defers judgment about available instruments to the collective of science education researchers. This approach's validity is substantiated by the fact that its outcomes resonate strongly with the conclusions of prior reviews of the domain. For instance, in their reviews of NOS assessments, Lederman (2007) and Lederman and colleagues (1998) examined the development and psychometric properties (both firsthand and by reference to other reviews) of many of the instruments being examined here. They listed all of the aforementioned nine instruments—except the *FAS*, which they considered not to address the NOS domain and which was included in the present review with qualification—as those considered to be valid and reliable measures of NOS understandings. Additionally, an examination of use in empirical studies shows that instruments that historically have been critiqued for addressing affective, attitudinal, and other outcomes deemed to be beyond the scope of the construct of NOS (see Lederman, 2007) do not seem to have played a significant role in the history of NOS assessment, as evidenced by their restricted use by researchers. The latter include instruments such as the SAQ (Wilson, 1954), ASSC (Allen, 1959), *ISAIA* (Swan, 1966), and *TRI-S* (Schwirian, 1968).

Third, a use-of-measures perspective presents a different profile in terms of the changing landscape of NOS assessment. Figure 31.1 shows instances of use of NOS measures over the past 60 years, when instruments are grouped by two broad types of items: forced-choice (i.e., agree/disagree, multiple choice, and Likert) and open-ended items. Figure 31.1 also shows the specific years when new instruments of either broad type were developed. It could be seen that between 1954 and 1998, the field was exclusively dominated by the use of forced-choice–type instruments. A substantial shift was evident in the decade from 1990 to 1999, where the use of open-ended instruments rose sharply from virtually zero in the previous four decades to account for 27% of all empirical studies that examined learners' NOS conceptions. This trend continued to assert itself in the period from 2000 to 2012, where the use of open-ended instruments more than doubled compared to the preceding decade and came to represent 57% of all instances of use. Interestingly, the latter shift becomes even more pronounced when geographical locations of researchers are taken into account. Indeed, in the North American context where research on NOS in science education has had the deepest roots and longest-running tradition, *no* new forced-choice NOS instruments were developed during the past two decades. Researchers outside the North American context have developed *all* forced-choice instruments listed in Table 31.1 in the period from 1993 to 2012. In that sense, the shift from forced-choice, "quantitative" to more open-ended, "qualitative" approaches to NOS assessment has been rather decisive over the course of the last two decades. The implications of this shift in terms of the validity of gauging learner NOS understandings are explicated below.

The grouping of instruments into the two broad types of forced-choice and open-ended measures is useful and sheds important light on the changing nature of the field. This grouping, nonetheless, is somewhat coarse: It does not account for the unique case of the *VOSTS* (Aikenhead et al., 1987), which stands in a category of its own among forced-choice instruments. While the *VOSTS* items are of the forced-choice type, its empirically driven development process renders it unique and affords it a measure of validity that could not be claimed in the case of theoretically driven forced-choice instruments. When the development of the *VOSTS* is taken into account, a pattern of historical progression from theoretically driven forced-choice to empirically driven forced-choice to open-ended instruments becomes the most prominent feature in the landscape related to the evolution of the field of NOS assessment. The next section focuses on this progression and its underlying conceptual bases.

Evolution of Approaches to NOS Assessment

Since the outset, the development of forced-choice NOS instruments has been theoretically driven: The model for NOS underlying an instrument, the domains of the scientific enterprise targeted by the instrument, and consequently the espoused respondent positions toward issues addressed in the instrument items all were derived from some theoretical delineation or portrayal of NOS. In the case of the very first instrument, Wilson (1954) emphasized the importance of developing students' "understanding of science and the methods by which scientific knowledge has been obtained" (p. 159). Toward assessing these understandings, Wilson prepared a set of 26 statements, which became the SAQ, by drawing on "two recent books dealing with the general subject of understanding science . . . *Science and Common Sense* by James B. Conant and *The Path of Science* by C. E. Kenneth Mees" (p. 161). Similarly, the elements underlying the *SPI* were derived from "books by Beveridge, Conant, Kemeny, Lachman, Nash, and [E. B.] Wilson" and then "presented to fourteen research scientists for validity judgment" (Welch & Pella, 1967, p. 64). The latter books also were general readers about the scientific enterprise, ranging from *The Art of Scientific Investigation* (Beveridge, 1950) to a *Philosopher Looks at Science* (Kemeny, 1959) to *The Foundations of Science* (Lachman, 1960). Many authors of the books consulted by researchers to design items for early NOS instruments, such as the SAQ, *TOUS*, and *SPI*, were historians or philosophers of science or drew on these disciplines in their writings. By the late 1960s, nonetheless, many instrument developers replaced the by-proxy or secondary access to the philosophy of science with firsthand consultation of this literature. For instance, toward developing the *NOSS,* Kimball (1967–1968) noted that its underlying "theoretical model of the nature of science was developed out of extensive study on the nature and philosophy of science" and continued, "While this model was only one of many possibilities, it was consistent in its agreement with views expressed by Conant and Bronowski; and additional support . . . was found among the writings of other philosophers of science" (p. 111). The significance of Kimball's statements, in addition to illustrating the point about the shift to firsthand consultation of the philosophy of science, is their being the first to explicitly point out that assessment instruments were undergirded by a "theoretical model" of NOS, which could be "one of many possibilities." This is a crucial point to which I return shortly. By the mid-1970s, use of the philosophical literature in instrument development was firmly established. Moreover, by the late 1970s, science education researchers seemed to have started being selective about which philosophical literature they would draw on to establish their model for NOS. For instance, Ogunniyi (1982) derived his *Language of Science* (LOS) instrument items from "statements about the language of science held by Carnap (1966), Hempel (1966), Frank (1962), Kemeny (1959), Nagel (1961), Nash (1963), and Popper (1962)" (p. 26). Ogunniyi, however, made the conscious decision to exclude the work of Thomas Kuhn from among the works consulted, arguing that "although Kuhn's influence has been felt since the 1960s his effort has been mainly on the historical and socio-psychological milieu of science rather than the language of science" (p. 26). The intention here is not to be critical of Ogunniyi or his philosophical choices. What is important is the fact that NOS models underlying various instruments drew on philosophical literatures that overlapped to various extents. Thus, by design, theoretically driven NOS instruments embodied their developer(s)' selective philosophical preferences in terms of, at least, which philosophical issues and stances were included and which were excluded.

An even more involved approach was reflected in the fact that some instrument developers seem to have taken sides on certain long-lived, contemporaneous philosophical debates in their models for NOS. For instance, Pomeroy (1993) adopted what she dubbed a "nontraditional" view of science, which seems to have endorsed a constructivist view of science and scientific knowledge and suggested, among other things, that science is just another way of knowing. In accordance with the NOS model embodied in Pomeroy's BNSSE, elementary teachers held more nontraditional (considered more informed) views of science when compared to scientists, who endorsed the most traditional views (considered more naïve). Again, this observation is not intended to criticize Pomeroy's model for NOS but to indicate that hers had taken sides on controversial philosophical debates, such as that related to the relative merits or "equivalence" of scientific versus indigenous ways of knowing (see Loving, 1997). Obviously, nothing is inherently flawed in a NOS assessment embodying a theoretical model for NOS, as long as such a model is explicitly articulated and made accessible to readers. In this regard, it should be noted that several researchers did not make their NOS model and, in some cases, their very instrument, accessible to other researchers; the *NOST* (Billeh & Hasan, 1975) is a case in point. Like with the design of any assessment, a theoretical or conceptual delineation of the dimensions and attributes of the target construct is a precursor to building valid assessments. However, when theoretically driven NOS instruments are packaged as assessment tools that draw on forced-choice items, the resulting assessment approach becomes very problematic.

Discontent with forced-choice NOS instruments dates back to the early 1970s. For instance, Mackay and White (1974) criticized the use of Likert-type and multiple-choice items in several instruments, including the *TOUS* and *FAS*. Even though the authors were mostly concerned with the extent to which these instruments enable valid assessment of students' *perceptions* of scientists, their criticisms reasonably apply to accessing understandings of other NOS dimensions. Mackay and White critiqued the *TOUS*'s use of multiple-choice items because "it may be difficult to communicate to the respondent the criteria to be used in selecting the best answer" (p. 132). Moreover, they questioned the validity of responses to Likert scales because a respondent's standing on such scales "is a measure of a complex combination of his standing, of his ability to perceive the intent of the scale, and of his set to respond in a particular way" (p. 132). By the mid-1980s, these initial concerns had developed into a highly articulate critique of the validity of forced-choice, theoretically driven NOS instruments in assessing learners' conceptions. The critique had two major dimensions.

First, Aikenhead and colleagues (Aikenhead, 1988; Aikenhead et al., 1989) argued that forced-choice instruments were based on the problematic assumption that respondents' perceptions and interpretations of an instrument's items are commensurate with those of its developers. They continued that ambiguities result from assuming that respondents understand a certain item in the manner that the instrument developers *intended* for it to be understood, or that respondents would agree or disagree with a statement, select one of the alternatives furnished in a multiple-choice item, or indicate preference on a Likert scale for reasons that coincide with those of the instrument developers. Thus, the validity of these instruments is seriously threatened because of difficulties with interpreting respondent choices. Lederman and O'Malley (1990) empirically supported this position. Second, as noted, forced-choice instruments often embodied specific theoretical models for NOS. Also, items provided choices that were theoretically driven and designed with certain philosophical stances in mind, and the espoused selections or answers often reflected the developers' philosophical positions and preferences. Being of the forced-choice category, instruments ended up imposing developers' views on respondents:

> Irrespective of the choices the respondents made, they often ended up being labeled as if they firmly held coherent, consistent philosophic stances such as inductivist, verificationist or hypothetico-deductivist . . . Thus, the views that ended up being ascribed to respondents were more likely an artifact of the instrument in use than a faithful representation of the respondents' conceptions of NOS.
>
> (Lederman et al., 2002, p. 502)

Indeed, research has demonstrated that students' and teachers' NOS views are rather fluid, fragmented, lacking overarching frameworks, and sometimes outright inconsistent (e.g., Abd-El-Khalick, 2004a; Abd-El-Khalick & BouJaoude, 2003; Lederman & O'Malley, 1990). It could be seen that theoretically driven forced-choice instruments generated data that were difficult to interpret and not necessarily representative of respondents' NOS conceptions.

An additional third criticism was directed at the usefulness and meaningfulness of data garnered from theoretically driven forced-choice assessments. Aikenhead (1974b) initiated a broad critique of the meaningfulness of any conclusions about student learning when it comes to "knowledge *about* the scientific enterprise" that could be derived from quantitative instruments. He argued that the resulting aggregate data, such as total scores or gain scores on instruments, such as the *TOUS* and *SPI,* masked the "specific ideas students tended to learn" (p. 26). Such data could not address questions such as "What specific ideas have students learned? What misunderstandings have they still retained?" (p. 24). Instead, by conducting what he termed "qualitative analyses," which entailed examining student responses and performance on specific *TOUS* and *SPI* items, Aikenhead was able to discern specific NOS aspects on which students made progress as a result of engagement with instruction using the *Harvard Project Physics* course (Holton, Watson, & Rutherford, 1967).

Lederman and O'Malley (1990), Abd-El-Khalick (1998), and Lederman et al. (2002) expanded this criticism, arguing that quantitative instruments have limited usefulness. Albeit more suitable for large-scale assessments because of ease of administration and scoring, the aggregate measures that these instruments produce are generally limited to labeling respondents' views as "adequate" or "inadequate"—mostly by assigning student views cumulative numerical values—rather than elucidating and clarifying such views. Lederman (1986b) also noted that developers never clarified what numerical values constituted an "adequate" view of NOS on their instruments. Thus, the use of quantitative instruments limits the feasibility of drawing meaningful conclusions regarding the nature of learners' NOS views and/or assessing the meaningfulness and importance of any gains in understanding NOS achieved by learners as a result of various instructional interventions.

In this context, Aikenhead and his colleagues embarked on an ambitious project in the early 1980s to develop an "empirically driven" NOS instrument. The *VOSTS*, which was developed over the course of 6 years and involved thousands of Grade 11 and 12 Canadian students, is an inventory of 114 multiple-choice items, each of which comprises a statement with several reasoned viewpoints or positions. The latter were not derived from examining the philosophical literature. Instead, a student-centered process was used to develop the content and form of *VOSTS* viewpoints or positions, which were drawn from high school students' responses to open-ended items and follow-up interview questions. By substituting theoretically driven positions with student response patterns, Aikenhead and colleagues (1989) constructed an empirically based instrument, which brings a high degree of validity to *VOSTS* responses, since "the meaning that students read into the VOSTS choices tends to be the same meaning that students would express if they were interviewed" (Ryan & Aikenhead, 1992, p. 576). It should be noted that Aikenhead and colleagues never generated a scoring scheme for the *VOSTS* or used cumulative scores when reporting on *VOSTS* data. Instead, frequency and percentage distributions of respondent choices to individual items were reported and interpreted for the purpose of drawing conclusions from *VOSTS* data (e.g., Aikenhead 1987, 1997).

Empirical evidence further supported the validity of the *VOSTS.* Aikenhead (1988) investigated the degree of ambiguity associated with the use of four different response formats to assess high school students' NOS understandings. These formats were Likert-type items, student-generated paragraphs, semistructured interviews, and *VOSTS* items. He reported that Likert-type responses "offer only a guess at student beliefs, and the chances of an evaluator guessing accurately are very remote" (p. 615). The ambiguity associated with using Likert-type items was close to 80%. It is noteworthy that 56% of all NOS instruments in Table 31.1 used one or another form of Likert-type scales. Written paragraph responses generated 35 to 50% ambiguity, followed by *VOSTS* empirically derived multiple-choice items at 15 to 20% ambiguity. Semistructured interviews "offered the most lucid and accurate data" with about 5% ambiguity (p. 625). Empirically driven items, thus, far surpassed theoretically driven items in gauging respondents' NOS understandings. Following its development and dissemination, the *VOSTS* became the NOS instrument of choice for a substantial period of time. As Table 31.2 indicates, it is the third most used instrument in empirical investigations into learners' NOS understandings and continues to be used into the present, especially in the case of large-scale studies where using resource-intensive qualitative approaches (i.e., open-ended questionnaires or interviews) simply is impractical (e.g., Dogan & Abd-El-Khalick, 2008).

The next epoch in the development of NOS assessment approaches started with some dissatisfaction related to the *VOSTS*. Lederman et al. (2002) argued that, when used outside the Canadian or Western contexts, the aforementioned criticisms of forced-choice instruments would as well apply to the *VOSTS*. From the perspective of non–Canadian or non–Western respondents, *VOSTS* viewpoints, in an important sense, would impose on them a certain set of responses like with the case of other forced-choice instruments. Additionally, the *VOSTS* multiple-choice format limits the space of answers available to respondents. Indeed, Abd-El-Khalick and BouJaoude (1997) found that, when *VOSTS* items were modified to allow respondents to articulate whatever ideas they deemed to be representative of their views, participant Lebanese science teachers indicated that their positions on some issues were either not represented among or were combinations of the provided *VOSTS* positions. Some participants expressed viewpoints that were totally different from ones presented in *VOSTS* items. More importantly, the development of the next generation of NOS assessments was motivated by the demonstrated value of interviews in generating valid profiles of learners' understandings of NOS (e.g., Driver et al., 1996; Lederman & O'Malley, 1990). As Aikenhead (1988) had found, compared to other response formats, semistructured individual interviews generated the least ambiguity (about 5%) in interpreting responses to NOS–related prompts. Interviews not only help to elucidate respondents' views on issues related to NOS but also provide them with opportunities to articulate the reasoning underlying as well as the interrelationships among their views.

Lederman et al. (2002) chronicled the development of the *VNOS* instrument (Forms A, B, and C), which started as a semistructured interview protocol (Lederman & O'Malley, 1990) and developed into an open-ended questionnaire whose administration is always coupled with follow-up semistructured interviews with a subset of respondents. The detailed, follow-up interview protocol (see Abd-El-Khalick, 2004a) is intended to probe and clarify responses in order to reduce ambiguities, which might result from interpreting written responses, and ensure a high degree of congruence between researchers' interpretation and respondents' intended meaning in relation to their NOS understandings. The *VNOS* is among a very few NOS instruments whose *construct* validity has been empirically established (Bell, 1999). Its use enables researchers to avoid the threats to validity associated with the use of theoretically driven forced-choice instruments and generate expansive profiles of respondents' conceptions of several important NOS dimensions and the reasoning underlying these conceptions, thus producing more meaningful data as compared to empirically driven forced-choice instruments. While the high resource burden associated with administering the *VNOS* and analyzing the resultant data renders it less than ideal for use in large-scale studies, *VNOS* data enable pinpointing changes in respondents' NOS views and linking such changes to specific elements in instructional interventions used in empirical studies. The *VNOS* has become the most used NOS instrument (see Table 31.2).

The evolution of approaches to NOS assessment from theoretically based forced-choice instruments (e.g., *TOUS, SPI, NSKS*) to empirically based forced-choice instruments (i.e., *VOSTS*) and then to open-ended instruments (e.g., *VNOS, CI, POSE*) was based in robust conceptual arguments related to the validity and meaningfulness of these assessments, as well as supporting empirical evidence, especially the work of Aikenhead (1988). This evolution also is reflected in NOS instruments use: The pattern is most clearly evident in the use of the *TOUS*, *VOSTS*, and *VNOS*, which together account for 51% of the use of all instruments in empirical research studies (see Table 31.2). Figure 31.2 shows the frequency of this use

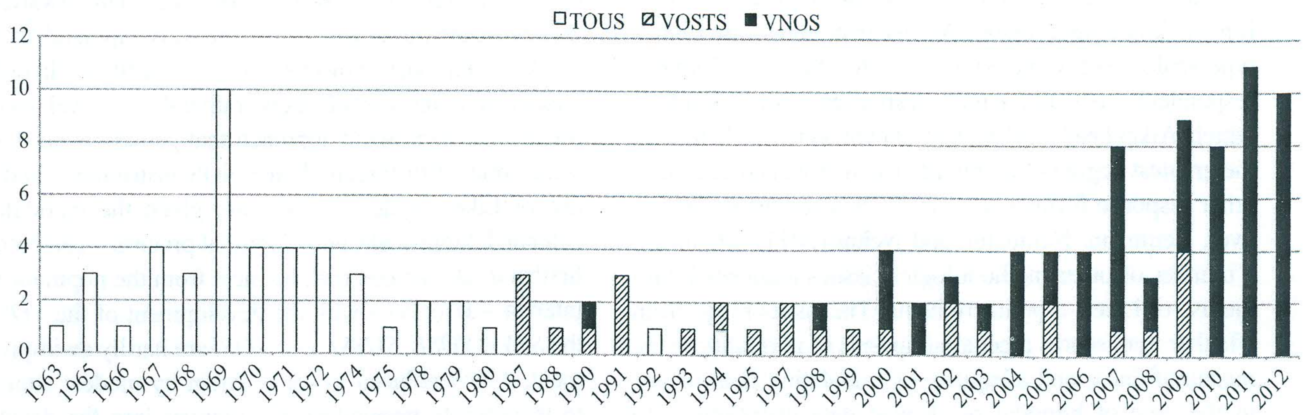

Figure 31.2 Use of the TOUS, VOSTS, and VNOS in empirical studies (1963–2012). Not shown are years in which none of the three instruments was used.

over the period from 1963—the first time the *TOUS* was used in a published study, to 2012.

Use of the *TOUS* peaked in 1969 with 10 published studies, but by the 1980s it had fallen out of favor. Since its formal introduction in 1987, the *VOSTS* dominated the field till around 2000 despite the introduction of the *VNOS–A* in 1990 (Lederman & O'Malley, 1990) and *VNOS–B* in 1998 (Abd-El-Khalick et al., 1998), as well as five other forced-choice instruments—namely, the Methods and Nature of Science (MaNS; Johnson & Peeples, 1987), Views about Philosophy of Science (VaPS; Koulaidis & Ogborn, 1988), *MNSKS* (Meichtry, 1992), BNSSE (Pomerory, 1993), and *BASSSQ* (Aldridge et al., 1997). Indeed, between 1987 and 2000, the *VOSTS* accounted for 35% of all use of NOS instruments in the field. It continues to be the forced-choice instrument of choice, especially for use in large-scale studies (e.g., Bennett & Hogarth, 2009; Dogan & Abd-El-Khalick, 2008; Kang et al., 2005). By around 2000, at the time the *VNOS–C* was introduced (Abd-El-Khalick, 1998; Abd-El-Khalick & Lederman, 2000a), use of the instrument picked up. The *VNOS* accounted for 57% of all NOS instruments use between 2000 and 2012 (see Table 31.2 and Figure 31.2). However, as Table 31.1 shows, researchers continued to develop new forced-choice instruments during the past decade. This observation invites close examination in the context of the narrative presented here about the development of NOS assessment approaches.

More Forced-Choice Instruments: Progression or Regression?

Since 2000, five new forced-choice instruments have been developed: *SEVs* (Tsai & Liu, 2005), *Views on Science and Education Questionnaire* (VOSE; Chen, 2006), *Student Understanding of Science and Scientific Inquiry* (SUSSI; Liang et al., 2006), *Myths of Science Questionnaire* (MOSQ; Buaraphan, 2009a), and *Nature of Science Instrument* (NOSI; Hacıeminoğlu et al., 2012). All five instruments, it was noted earlier, were developed outside the North American context. Indeed, no new forced-choice instruments were developed in the latter context since 1998. Close examination of these five instruments shows that they hardly provide substantially more viable alternatives than their theoretically or empirically driven forced-choice predecessors. All five instruments use Likert-type scales (see Table 31.1; note that the *SUSSI* invites respondents to explain their responses with examples), which Aikenhead (1988) had shown generated, by far, the greatest degree of ambiguity (up to 80%) compared to other response formats used in NOS assessments. Moreover, Neumann, Neumann, and Nehm (2011) articulated a number of other methodological issues associated with the use of Likert-type instruments. The issues range from whether item scores meet assumptions of normality to the meaningfulness and adequacy of calculating item means in the case of bimodal or skewed data distributions to whether "participants answered all items in a consistent fashion, that is, whether all category options were comparably scaled by the respondent, independent of the item" (p. 1375). Consider the *MOSQ* (Buaraphan, 2009a, 2009b): This instrument's development was limited to selecting a set of 14 broad theoretically driven statements, establishing their face and content validity by a panel of researchers, and "piloting" the instrument with a small number of preservice and inservice science teachers. The *MOSQ* psychometric properties do not seem to have been explored or established, and it would not qualify as a valid and reliable instrument.

Statements on the *SEVs* also are theoretically driven, while the remaining three instruments include what could best be described as a combination of theoretically driven and by-proxy empirically driven statements. Development of statements for the *VOSE*, *SUSSI*, and *NOSI* drew on *VOSTS* items and viewpoints; student response patterns that emerged from other empirical studies, which had used open-ended tools, such as the *VNOS* and semistructured interviews; and/or results from small-scale pilot studies conducted by the instrument developers, some of which involved individual interviews. For instance, in lieu of an empirically driven process that drew on the positions of participants from the population of interest, such as that undertaken by Aikenhead and colleagues, *VOSE* items "were revised from VOSTS or generated according to statements that emerged from the pilot study and the recent literature, such as Khishfe and Abd-El-Khalick (2002), Lederman et al. (2002)" (Chen, 2006, p. 807). However, *VOSTS* viewpoints were derived in the Canadian context, and studies by Khishfe and Abd-El-Khalick (2002) and those reported by Lederman et al. (2002) were undertaken in substantially different contexts (e.g., Lebanon and the United States) from the Taiwanese context, for which the *VOSE* was intended. This concern also applies to the *SUSSI* and *NOSI*.

Development of the *SEVs*, *VOSE*, *SUSSI*, and *NOSI* involved the use of expert review panels, small-scale pilot studies coupled with follow-up individual interviews, and the establishment of the instruments' psychometric properties, mostly by the use of factor analysis techniques. Developers clearly were aware of and explicitly attempted to address the well-documented shortcomings of forced-choice theoretically driven instruments. This awareness and associated attempts surely are to be applauded. The resulting approach, however, could as well be characterized as including a set of "fixes" rather than a novel or comprehensive assessment approach that genuinely addresses issues related to forced-choice NOS instruments and the use of Likert scales. To be sure, given the meticulous, extended, large-scale, and empirical process—centered on firsthand interactions with learners from the population of interest—associated with the development of the *VOSTS*; the *SEVs*, *VOSE*, *SUSSI*, and *NOSI* are hardly an improvement. Albeit reduced to some extent by modest attempts to incorporate respondent perspectives into the development process, these instruments still are subject to validity threats entailed by assuming that respondents would select a preference on a Likert scale on any of these instruments' items for reasons that correspond with or are similar to those intended or implied by instrument developers. Additionally, the criticism related to the meaningfulness of the data generated by these forced-choice instruments still holds (Lederman et al., 2002). In this regard, these new instruments are equivalent with their older counterparts, such as the *NSKS* and *NOSS*.

These new instruments also have come under criticism from a psychometric perspective. Like most quantitative NOS instruments to date, the *SEVs*, *VOSE*, *SUSSI*, and *NOSI* were developed within the context of classical test theory. Neumann and colleagues (2011) applied item response theory (Rasch Modeling) to assess the validity, reliability, and dimensionality of the *SUSSI* (Liang et al., 2006). Neumann and colleagues (2011) reported that, first, "a construct analysis revealed that the instrument did not match published operationalizations of NOS concepts" (p. 1373). Their analyses showed that a two-dimensional Rasch model—one that accounts for NOS and the other for scientific inquiry (SI)—had a significantly better fit than the one-dimensional model underlying the *SUSSI*'s development. The two distinct dimensions, Neumann's group continued, corroborate theoretical distinctions in the literature between NOS and SI (see Abd-El-Khalick et al., 1998) and detract from the *SUSSI*'s theoretical model, which suggests the two domains are more closely aligned and could serve as a singular construct. These findings cast serious doubt on the validity of the *SUSSI*. Neumann and colleagues' (2011) work suggests that newly developed quantitative, forced-choice NOS instruments would benefit from the use of state-of-the-art item response theory to investigate their psychometric properties, including measures of validity and reliability. Based on these arguments, it could be asserted that the recent development of new forced-choice NOS instruments surely does *not* represent advancement in the field. Such development, nonetheless, speaks to the need for "quantification" of NOS understandings for purposes of investigating relationships with other important attributes and science learning outcomes. This need deserves substantial attention.

On "Context" and NOS Assessment: Issues With Inference From Actions to Beliefs

Approaches to NOS assessment, it could be seen, are of two general types. The first are generative: These ask learners to *produce responses* toward explicating and reflecting on their understandings and thinking about NOS. Generative assessments use individual interviews (e.g., Abd-El-Khalick & BouJaoude, 2003; Bianchini & Solomon, 2003; Driver et al., 1996; Lederman & O'Malley, 1990), open-ended questionnaires (e.g., Lederman et al., 2002: *VNOS*), prompts (e.g., Mead & Métraux, 1957: ISSSC), and/or scenarios (e.g., Nott & Wellington, 1995: CI). The second type includes forced-choice varieties, which ask learners to *select responses* from alternatives on multiple-choice items (e.g., Aikenhead et al., 1987: *VOSTS;* Cooley & Klopfer, 1961: *TOUS*) or *indicate a preference* on a Likert-type scale (e.g., Cotham & Smith, 1981: *COST;* Kimball, 1967–1968: *NOSS*). The preceding section spoke to the differential validity and meaningfulness of data collected as well as benefits and burdens associated with using these types of assessments. Both types, nonetheless, seem to be "susceptible" to the same factor, namely, that of "context." This factor, however, is used in two very different manners in the literature related to NOS assessment.

First, "context of assessment" is used to refer to whether and the extent to which an item, prompt, or scenario on a NOS measure is devoid of or anchored in some specific referent and the extent to which such referent is familiar, accessible, or understandable to respondents. Referents might include, among other things, a specific scientific theory or controversy, a socioscientific issue, or a context that could be typified as belonging to school science (versus authentic scientific practice), such as description of an experiment or an inquiry activity performed by students or the respondent. For example, Item #41 on the *NSKS* (Rubba, 1976) asks respondents to indicate their degree of agreement or disagreement with the statement, "Scientific theories are discovered, not created by man" (p. 5). In comparison, the *COST* provides a three-paragraph description of geological theories and then asks respondents to indicate the extent of their agreement with the statement, "Dietz and Hess didn't invent the theory of plate tectonics. They objectively derived it from the facts" (Item #2, p. 145). Clearly, both items are designed to access respondents' conceptions related to the same NOS aspect, specifically, whether scientific constructs are discovered from facts or invented to account for such facts. While the *NSKS* item stood decontextualized, the latter was anchored in a discussion of plate tectonics theory. The *COST* also purports to assess learners' familiarity with each of the theories used to contextualize its items by asking the respondent to "describe the state of your knowledge of the . . . subjects [e.g., theories or topics]" used in the questionnaire on a five-point scale from "mastery" to "no knowledge" (p. 153). Further, the *COST* asks each respondent to "estimate to what extent your responses . . . were influenced by personal convictions independent of your understanding of the subjects," again, on a five-point scale from "complete" influence to "none" (p. 153). Cotham and Smith (1981) argued that the latter information provided for additional confidence in interpreting learner responses to their context-based items. However, as with the case of the *COST,* no robust evidence currently exists about the *specific* ways in which and the extent to which context, science content knowledge, and/or background understandings or "convictions" impact the ways in which instruments provide faithful access to respondent understandings of NOS aspects and domains.

There is some evidence to suggest that context does impact responses to NOS–related items. Driver and colleagues (1996) reported that the images of science in which their participant students anchored their thinking about NOS–related interview questions impacted their expressed views. In particular, Driver and colleagues noted that students' responses differed when they thought about the familiar and proximal context of school science, as compared to the more abstract and distal context of "real" science. Nonetheless, there is need for more systematic evidence to establish and delineate these findings and gauge the extent of any associated impacts on the validity and reliability of NOS measures. Such evidence will need to draw on responses collected from the same population—as compared to drawing on a set of studies across contexts and populations as Driver and her colleagues had done—and will require the use of fully counterbalanced research designs both in terms of sequencing the prompts as well as the very content of these prompts.

Evidence related to the sensitivity of NOS views to science content knowledge has been, at best, equivocal. A host of early studies failed to establish clear or strong correlations between students' and teachers' NOS conceptions and their science content knowledge and/or science achievement (Billeh & Hasan, 1975; Carey & Stauss, 1968, 1970; Olstad, 1969; Scharmann, 1988a, 1988b; Welch & Walberg, 1972; Wood, 1972; Yager & Wick, 1966). The extent to which these findings were associated with the aforementioned problems related to the validity of the forced-choice NOS instruments used in these studies remains unclear. There is some evidence, nonetheless, to suggest that responses to NOS–related items are sensitive to the *quality* and *depth* of learners' science content understandings (as compared to the previously and often used gross measures of science content knowledge, such as grade point average or number of science credit hours). Teachers' responses to *VNOS* items is a case in point. For instance, Items #6 and #7 on the *VNOS–C* (Lederman et al., 2002, p. 509) are designed to shed light on respondents' views of the same NOS aspects, namely, the inferential, creative, and tentative NOS. Item #6 makes use of atomic theory as a context:

> Science textbooks often represent the atom as a central nucleus composed of protons (positively charged particles) and neutrons (neutral particles) with electrons (negatively charged particles) orbiting that nucleus. How certain are scientists about the structure of the atom? What specific evidence *do you think* scientists used to determine what an atom looks like?

In comparison, Item #7 uses the biological construct of species to provide such a context, while essentially asking the same question:

> Science textbooks often define a species as a group of organisms that share similar characteristics and can interbreed with one another to produce fertile offspring. How certain are scientists about their characterization of what a species is? What specific evidence *do you think* scientists used to determine what a species is?

Abd-El-Khalick (2004a) found that prospective secondary science teachers and college science students explicated differing views in relation to these two items depending on whether they had majored in physical versus biological sciences. Some respondents expressed views that could be characterized as more informed in the context of one item and ones that clearly were more naïve in the context of the other. This effect, however, mostly was observed in the case of participants who had completed advanced science courses in the physical or biological sciences and less evident in the case of those majoring, for example, in the geological sciences. Abd-El-Khalick concluded that (also see Leach, Millar, Ryder, & Séré, 2000; Ryder et al., 1999)

> For purposes of exploring college students' NOS views . . . disciplinary nuances and differences . . . need to be brought into the mix . . . researchers who aim to assess college students' NOS views have the task of further delineating and refining the specific realm and NOS aspects they aim to explore, as well as providing students with more specific contexts in their NOS assessments.
>
> (Abd-El-Khalick, 2004a, p. 417)

While there is need for more systematic research in this area, the evidence to date suggests that context—in the specific sense delineated above—matters. Given that reform efforts (e.g., AAAS, 1990; NGSS Lead States, 2013) call for developing students' understandings about the nature of the scientific enterprise writ large, it should prove very fruitful to pursue research questions about the specific ways in which learner NOS understandings are intertwined with the context in which they are learned, elicited, and assessed; the extent and feasibility of transfer from context-bound to generalized NOS understandings; and the ways in which context can facilitate or anchor as well as inform the interpretation of learner responses to various NOS measures.

Albeit much less prevalent, "context" also is used in a second, radically different sense when it comes to NOS assessment. In this second sense, some researchers argue that students' expressed ideas hardly are a faithful representation of their NOS understandings as indicated by the fact that, while engaged with the doing of science, students *are observed* to behave in ways that seem incommensurate with how they talk or write about science (e.g., Sandoval & Morrison, 2003). The implication of these findings, the argument continues, is that observing student behavior in situ or in "context," such as when they are engaged with doing inquiry activities, provides for a more valid assessment of their underlying NOS understandings (Kelly, Chen, & Crawford, 1998). Irrespective of the form that such an approach to assessment would take, it will necessarily involve inference to beliefs from actions and, thus, is apt to be problematic. Indeed, even practicing scientists do not necessarily *do* science in accordance with an articulated epistemological framework: As Kuhn (1996/1962) convincingly argued, such a framework rarely is explicated in scientific education or apprenticeships. While scientists' actions might be consistent with certain epistemological frameworks underlying the disciplinary traditions into which they were initiated, these actions might not tell much about a particular scientist's underlying epistemological beliefs. It only is during times of severe disciplinary crises, Kuhn showed, that scientists revert to explicating and examining the assumptions, rules, and frameworks underlying their practice (for a detailed discussion, see Abd-El-Khalick, 2012b).

Assessments involving inferences to beliefs from actions are based on the problematic assumption that learners' actions as they engage with *doing* science are necessarily reflective of and somehow consistent with an underlying framework on their thinking *about* science. This assumption is inconsistent with evidence indicating that students' NOS understandings often are contextual, fluid, and/or fragmented and lack overarching or consistent frameworks (Abd-El-Khalick, 2004a; Leach et al., 2000). Additionally, robust empirical evidence indicates that teachers' instructional practices—how they enact science and science instruction in their classrooms—often are incommensurate with their NOS understandings (e.g., Abd-El-Khalick et al., 1998; Lederman, 1999, 2007). More importantly, I argued (see Abd-El-Khalick, 2004c) that inferences to beliefs from actions run the risk of imposing an observer's own framework about NOS on observed behaviors, not because learners necessarily ascribe to such a framework but because observers approach the task with one or more coherent frameworks in mind (this is the theory-laden nature of observation in action!). This situation is similar to the case of forced-choice assessment instruments, which often suggested that students held some consistent NOS framework or another, which later turned out to be an artifact of the fact that instrument developers had designed these assessments with specific philosophical stances in mind (Aikenhead et al., 1987; Lederman et al., 2002).

Beyond Silence: Benchmarking "Controversial" NOS for Assessment

As noted earlier, the seeming dominant narrative in the field is one focused on a plethora of assessments and continued disagreement about the construct of NOS. Earlier, I examined the landscape of NOS assessment from a use-of-instruments lens. The examination was intended as a proxy for the collective judgment of the field about NOS measures and indicated a greatly reduced variance by virtue of differential use vis-à-vis the multitude of available instruments. This section will focus on the construct of NOS as *embodied* in these measures through a content analysis of instruments listed in Table 31.1, with the aim of gauging the extent to which the construct, as practically assessed, is contested.

A most important attribute of all NOS measures is that none embodies a covering or overarching model for NOS. This fact hardly is surprising given the very nature of scholarship in history, philosophy, sociology, and psychology of science (HPSSS; Abd-El-Khalick, 2012a). As Laudan and colleagues (1986, p. 42) put it, we currently "have no well-confirmed general picture of how science works, no theory of science worthy of general assent" (see also Rosenberg, 2005). Since the 1950s, science educators have aimed to help precollege students learn *some ideas* about the scientific enterprise and/or the development and validation of scientific knowledge, which were deemed central to a scientifically literate citizenry (Pella, O'Hearn, & Gale, 1966; Wilson, 1954). Researchers who set out to assess understandings of these ideas were keenly aware of the fact that any NOS model they had adopted toward developing their particular instrument "was only one of many possibilities" and would require some measure of consistency—though by no means complete agreement—among historians and philosophers of science (Kimball, 1967–1968, p. 111). This attribute was not explicitly acknowledged in early instruments, which outlined their domains in seemingly covering terms, such as "methods and aims of science," "assumptions of science," "science as an institution," or "activities, products, and ethics of science" (see Table 31.1). However, by the late 1960s, researchers started to explicate what is best described as an "aspects approach to NOS" (Abd-El-Khalick, 2012b), which translated into a set of descriptors signaling the NOS domains or ideas underlying the instrument at hand. The model for the *NOSS* (Kimball, 1967–1968), for instance, embodied the notions that science is driven by curiosity, empirical, reproducible, dynamic, parsimonious, tentative, imposes human constructs on nature, and is not characterized by a single "scientific method." An examination of Table 31.1 indicates that, with very few exceptions (i.e., the *NOST*, *TUNS*, and *VOSTS*), Kimball's (1967–1968) approach became commonplace, and models for NOS underlying the development of various instruments were outlined in terms of sets of NOS domains and aspects. None of these models, it should be noted, outlined an overarching framework to coherently connect and synthesize the respective domains or aspects. Below, I examine the implications of the latter, seemingly intentional omission. Toward understanding the extent to which the NOS aspects and domains targeted by various instruments were consistent across the past 60 years, a content analysis of the *individual items* for all instruments listed in Table 31.1 was undertaken. The analysis, it should be noted, did not include items that clearly were not related to NOS, such as those that addressed views of science education (*VOSE*, BNSSE), school science (*BASSSQ*), scientific inquiry (SUSSI), or acceptance of evolutionary theory (MaNS).

Analyses identified three clusters of NOS aspects. The first included an almost universal set of aspects, which were consistently addressed in the overwhelming majority of all instruments over the course of the past six decades (70–100% of instruments). These included understanding that (a) Scientific knowledge is *tentative* (changing, not certain or absolute, not to be equated with the Truth, all categories of scientific knowledge are amenable to change); (b) science is *empirical*, drawing on observations of the natural world toward developing scientific claims or adjudicating between conflicting claims; (c) scientific inquiry is *theory laden* or *theory driven:* data and observations are not inherently objective or neutral but are vulnerable to scientists' perspectives, mindsets, and theories. Data and observations do not speak for themselves: They require and are susceptible to interpretations, which are guided by scientists' perspectives and theories; (d) the development of scientific knowledge necessarily requires human *creativity*, imagination, and/or intuition, which could not be accounted for by strictly rational models of scientific discovery; (e) science is a *social* activity undertaken in the context of communities and institutions with well-defined structures and practices. The social nature of science is crucial to the production and validation of scientific knowledge, and the incorporation of such knowledge into the scientific cannon; (f) no one universal "scientific method" exists or characterizes the actual practice of scientists, and such a method does not guarantee the development of infallible claims to scientific knowledge; and (g) science is embedded in, affects, and is affected by the lager social and cultural milieu in which it is practiced. This *social and cultural embeddedness* of science also includes ideas related to the reciprocal influences of science and society (and, in some cases, technology).

The second cluster of NOS aspects was less prevalent and evident in 30 to 70% of all instruments. This cluster included ideas related to the nature of theories (e.g., role and functions of theories, theory testing, and adjudication between rival theories), nature of laws (e.g., function and epistemological status of laws), and the relationship between different types of scientific knowledge, including theories, laws, hypotheses, and/or facts. Two other aspects were focused on the distinction between scientific claims and their underlying evidence, that is, the inferential nature of scientific knowledge; and the aims of science, which often were cast in terms of contrasting utilitarian and practical perspectives on these aims with the value inherent to developing robust knowledge about, and explanations of, natural phenomena. Like the first cluster, this second set of NOS aspects could be identified in instruments throughout the past six decades, albeit not as continually addressed or revisited as with aspects in the first cluster.

The third cluster included aspects of two varieties that were much less prominent across NOS instruments. First were aspects characteristically addressed in early instruments up to the late 1960s but rarely addressed in later instruments. These include attitudes toward science, scientists, science-related careers, and products of science (especially in terms of beneficial or detrimental effects of products, such as medical technologies or nuclear weapons); perceptions of scientists (mainly focused on emphasizing that scientists were like other professionals and that being a scientist does not require inherent characteristics, such as being objective or genius); and processes of science. The second variety received occasional emphasis across the past six decades and included defining science (sometimes in contradistinction with technology or theological knowledge); limitations of science (e.g., answering questions about the natural world or informing rather than dictating policy-related decisions); models in science; assumptions of science (e.g., consistency of patterns across natural events; nature is comprehensible through systematic, empirical study); the question of demarcation (i.e., science versus pseudoscience), ethics and morality in science; experimentation and experimental design; and a few other descriptors of science and scientific knowledge (e.g., parsimonious, unified, unique, and developmental).

The domains and aspects that cut across all or the greater majority of instruments over the past 60 years are strikingly similar to sets of NOS "consensus" lists, which currently provide the basis for research and development efforts in the field (see Lederman, Chapter 31 this volume). NOS consensus lists converge irrespective of whether researchers go about the process by analyzing HPSSS literatures (e.g., AAAS, 1990; Abd-El-Khalick et al., 1998; Hodson, 1991; NRC, 1996), empirically identifying points of agreement about NOS in global science education reform documents (McComas & Olson, 1998), or establishing agreement among "experts" (Osborne, Collins, Ratcliffe, Millar, & Duschl, 2003). Indeed, at certain levels of generality, the very understandings that were deemed desirable across analyzed instruments also were strikingly similar. In this sense, the features of the NOS construct embodied in various instruments surely are *not* nearly as divergent as they are sometimes made to be by some researchers (e.g., Alters, 1997; Rudolph, 2000; Wong & Hodson, 2010). However, this is not to say that instruments featured no disagreements about desired NOS understandings. Indeed, disagreements were evident. For instance, compared to the *VNOS* (Lederman et al., 2002), the BNSSE (Pomeroy, 1993) endorsed a relativistic perspective on the nature of scientific knowledge. Similarly, the *COST* (Cotham & Smith, 1981) would allow a more prominent role for subjective factors in choosing between rival theories compared, for instance, to the *TOUS* (Cooley & Klopfer, 1961). Such disagreements mirror continuing debates within HPSSS, which bring us back to the very point underlying the "aspects approach" to conceptualizing NOS and the aforementioned seemingly intentional—indeed, necessary—omission in terms of articulating overarching, coherent frameworks to connect various aspects outlined in the models underlying NOS instruments.

Embracing "Controversy": A Developmental Framework for Benchmarking NOS Understandings for Assessment

Starting in the late 1960s, mindful of debates within HPSSS, researchers adopted the approach of outlining a set of valued NOS aspects, coupled with an (at least, implied) argument for the pragmatic irrelevance of high-level controversies about NOS to precollege students (Abd-El-Khalick, 2012b). Most researchers and instrument developers leveraged consensus, remained *silent* on controversial issues, and did not commit to a specific philosophical stance. This approach is best exemplified in articulations of NOS in science education reform documents. For example, the AAAS (1990) and NRC (1996) documents do not take a stand on continuing debates between empiricists (e.g., van Fraassen, 1998) and realists (e.g., Musgrave, 1998) as to the ontological status of entities postulated by scientific theories, among many other issues. The ontological status of scientific entities, it should be noted, is targeted in several NOS instruments. Reform documents do not even adopt a single stance on NOS, such as constructive empiricism (van Fraassen, 1980), scientific realism (Boyd, 1983), revolutions and normal science (Kuhn, 1996/1962), sophisticated falsificationism (Lakatos, 1970), or any other well-articulated theory of science. Science educators have highlighted the resulting inherent tensions in reform documents, such as affirming that scientific knowledge is both tentative and durable (AAAS, 1990, pp. 2–3; see Good & Shymansky, 2001). What is important to emphasize is that images of NOS conveyed in reform documents are qualitatively *more* informed and accurate than those often emphasized in science textbooks (Abd-El-Khalick et al., 2008), taught by most science teachers, and endorsed by most precollege students (Lederman, 2007).

From an applied perspective, the discussed framework has enabled science educators to navigate difficulties associated with incorporating informed NOS conceptions in science curricula and instruction without being paralyzed by continuing HPSSS debates. However, major issues will continue to plague the field related to NOS assessment in terms of benchmarking the construct. More often than not, some researchers seem to misconstrue the underpinnings and logic of the self-imposed silence on controversial NOS issues and actually *take sides* on some debates, advocating certain philosophical stances. This includes, for example, advocacy for relativist perspectives on the status of scientific knowledge or empiricist/instrumentalist perspectives on the nature of scientific entities (e.g., Pomeroy, 1993). Thus, while the NOS domains and aspects often addressed in various instruments are very similar, the stances espoused by these instruments—for, at least, a subset of these NOS ideas—likely will continue to be susceptible to the instrument developers' and researchers' philosophical background, preferences, and stances.

Thus, when it comes to benchmarking NOS understandings for assessment, there is need to move the field forward by extending the current consensus framework. Researchers cannot continue to sidestep controversies by remaining silent on contested aspects of the attributes of scientific knowledge or how scientific knowledge is produced and validated. The alternative frame should remain faithful to the controversial nature of some NOS dimensions while simultaneously enabling the development and enactment of science curricula, instruction, and assessment to help students internalize images of NOS that are qualitatively more informed than the naïve portrayals of science to which they continue to hold steadfastly. Abd-El-Khalick (2012b) outlined such a framework, which continues to focus on aspects of NOS currently emphasized in reform documents and widely supported by science educators (see Chapter 31 this volume). Nonetheless, these NOS aspects would be addressed at increasing levels of depth along a *continuum* from general, simple, and unproblematic in elementary grades to a treatment that is specific, complex, and problematized/controversial in the context of teacher education (see Table 31.3 for illustrative examples). The continuum also should take learners' developmental levels into consideration. The level of generality at which NOS aspects are addressed in elementary grades would render them noncontroversial but genuinely more informed than currently prevalent naïve conceptions. Secondary school students would discuss NOS aspects with reasonable levels of sophistication that go beyond mere generalization, such as that scientific knowledge depends on empirical evidence or is tentative. All along this continuum, the interrelatedness of NOS aspects would be progressively examined with greater depth to provide learners with ample opportunities to construct, reconstruct, and consolidate their own internally consistent frameworks about the workings of science. In teacher education and professional development contexts, prospective and practicing science teachers would tackle nuanced complexities about NOS, including major debates and controversies within HPSSS. As a result, teachers would be well positioned to not only support student NOS learning but also to calibrate their teaching to the level of depth at which they address NOS and to the specific interests and abilities of their students (for a detailed discussion see Abd-El-Khalick, 2012b).

This developmental framework transforms the nature of conceptualizing and deploying NOS assessments and associated measures. Assessment developers would not only need to delineate the set of aspects or domains that underlie their assessments but also the level of depth of targeted ideas, as well as the manner in which and the sophistication with which these NOS ideas are interrelated. The framework firstly entails calibrating the level of generality at which a set of NOS ideas are noncontroversial for the *specific* learner audience in mind and the measure of agreement—or lack thereof—among HPSSS scholars, which is associated with this level of generality. For instance, there hardly is any disagreement about the idea that "observation is always selective. It needs a chosen object, a definite task, an interest, a point of view, a problem" (Popper, 1963, p. 45). As Darwin had put it in a letter to Henry Fawcett in 1861, "all observation must be for or against a point of view if it is to be of any service!" (Barlow, 1993/1958, p. 161). In this general sense, the notion that observation is *not* inherently neutral, objective, or presuppositionless is virtually noncontroversial (O'Hear, 1989), an aspect that has consistently been

TABLE 31.3
A Developmental Framework for Benchmarking NOS Understandings for Assessment: Illustrative Examples

Educational level	Examples of target NOS aspects				Level of depth
	Tentative NOS	Theory-Laden NOS	Empirical NOS	Social NOS	
College; teacher education and professional development	There are debates as to whether scientific knowledge grows by accretion (commensurable across theoretical frames) or through paradigmatic shifts (incommensurable across theoretical frames).	There are debates as to the nature and significance of the theory-ladenness of observation. Is theory-ladenness significant beyond situations in which the evidence lies at the very edge of the human perceptual apparatus and/or observational instruments?	Scientific theories are underdetermined by evidence. Debates continue about the extent to which rationality versus value judgment mediates the use of evidence in the process of theory choice in science.	Debates continue about the viability of social constructionist conceptions in accounting for science's success in the absence of "realist" conceptions. Only "miracle" can explain such success if science was not getting closer to the "true" nature of phenomena!	Specific, complex, problematized (controversial)
Secondary school	Scientific knowledge is expanded, revised, or rejected because of two fundamental reasons: (1) New evidence is brought to bear (empirical NOS), and/or (2) existing evidence is reinterpreted in light of theoretical advances (theory-laden NOS). Scientific knowledge changes in at least two fundamental ways: (1) It is expanded through accretion, and/or (2) discarded and altogether replaced with new knowledge.	Theories might determine what scientists "see" when conducting investigations through selective attention and/or influencing the interpretation of "raw" inputs from the environment. Theories are crucial to conducting scientific investigations because theories enable scientists both to choose what evidence to collect (and what evidence to disregard) and how to interpret the collected evidence.	The relationship between scientific knowledge and evidence is indirect. Theories can only be tested by comparing their consequences with empirical observations. Hypotheses do not "jump out" from evidence: Inference beyond the evidence is usually involved (creative NOS). The relationship between knowledge claims and evidence is mediated by theory (theory-laden NOS).	The social character of science contributes to its objectivity: Intersubjective critical discourse through established value-driven channels (e.g., double-blind review procedures) minimizes the subjectivities of participant scientists. Science is conducted within the context of social institutions and has established and identifiable norms.	
Elementary school	Scientific knowledge is subject to change over time.	Scientific investigations always involve theories and empirical observations.	Science demands evidence: Scientific knowledge is derived from and/or supported by observations of the natural world.	Different groups of scientists contribute to the development of scientific knowledge.	General, simple, unproblematic

emphasized in assessments and a crucial idea about NOS for secondary school students to understand. However, a developmental framework entails, secondly, that instrument developers become aware of debates or controversies associated with the ideas at hand and tailor their assessment accordingly. To continue with this example, NOS assessments should avoid taking sides in the form of commitments related to the theory-laden NOS, such as that "The course of scientific discovery resembles the process of reaching a difficult judicial decision" (Pomeroy, 1993, p. 275: BNSSE, Item 17) when working with learners at certain levels of sophistication. Science teachers, for instance, are better positioned to teach about NOS if they are, in this case, aware of the continuing debates about the nature of the relationship between theoretical commitments and empirical evidence and about the extent to which theory overrides evidence when it comes to choosing between rival theories (see Abd-El-Khalick 2012a; Curd & Cover, 1998). While similar in several respects, making judicial decisions and adjudicating between scientific theories—at least for a subset of NOS scholars—still differ in crucial ways, which are dependent on observations of the behavior of natural phenomena.

Table 31.3 provides additional illustrative examples about core NOS aspects whose assessment should be approached at different levels of generality and sophistication. It could be seen that instruments and measures designed to assess secondary science teachers' and college students' NOS understandings would need to be calibrated accordingly and create "spaces" for eliciting respondents' understanding about debates associated with and the interrelatedness of such core NOS aspects. These "spaces" could take the form of appropriate alternatives in forced-choice instruments or carefully crafted contexts, scenarios, and prompts in open-ended measures. In this sense, some level of understanding of relevant NOS debates and controversies for certain groups of audiences, especially science teachers and college students, becomes *the* espoused view underlying NOS measures. This suggested approach to NOS assessment, it should be emphasized, requires robust understandings of the *developmental appropriateness* of the various NOS ideas along a continuum, such as that outlined in Table 31.3. Currently, our understandings of such developmental appropriateness are anecdotal and informed by localized studies rather than based in a systematic body of literature. Investigating the developmental appropriates of NOS ideas along the K–16 educational ladder should prove to be a very fruitful line of research for those interested in the domain of NOS in science education in general and in the field of NOS assessment in particular.

Concluding Remarks and Moving Forward

As evident in the present review, the landscape of assessment related to NOS is much more streamlined than it is often perceived or made to be. The collective wisdom of practice reflected in the extent and patterns of use of available NOS assessments in empirical studies indicates that robust comparisons across studies, populations, and interventions are possible by virtue of the selective use in research of a rather small of number of instruments. Analyses also revealed a clear pattern for the evolution of approaches to NOS assessment from theoretically driven to empirically driven forced-choice instruments and finally to open-ended assessment approaches. This evolution carried with it substantial improvements in terms of the validity of assessments and the meaningfulness of the data generated about learners' NOS conceptions. Use, coupled with content analyses of NOS instruments developed over the past six decades, revealed wide agreement around a set of core NOS aspects and domains, which are deemed central to a robust precollege science education. Nonetheless, the very nature of HPSSS scholarship creates tensions for NOS assessment, especially around contested areas related to these core NOS aspects. Toward addressing this important tension, I argued for embracing "controversy" about NOS by adopting a developmental framework for benchmarking NOS for purposes of assessment. Additionally, I argued that context *matters* for NOS assessment and should be accounted for in related measures. However, it also was argued that approaches purporting to assess learners' views of NOS by inference from learner actions to their beliefs are very problematic.

Finally, I argued that forced-choice instruments that continue to be developed into the present despite robust theoretical arguments and empirical evidence that cast serious doubt on their validity do *not* represent an advancement in the field. The continued development of such instruments, however, reflects the need for resource-lean assessments and quantification of learners' NOS understandings for the respective purposes of use in large-scale studies and statistical examination of the interrelationship between learner NOS understandings and a host of other variables of interest. Instead of *going back* to the development and use of forced-choice instruments, the next chapter in the evolution of NOS assessment is the development of valid and reliable as well as efficient ways to quantify student responses to open-ended instruments. Abd-El-Khalick and his students have been working on the development of a rubric to "score" *VNOS–C* responses (Abd-El-Khalick, 2004b; Abd-El-Khalick, Belarmino, & Summers, 2012). The approach capitalizes on the virtues of open-ended assessments in eliciting valid representations of respondents' NOS views and then uses detailed rubrics and decision-making trees to generate scores that reflect student understandings on a scale, which would allow valid comparisons both within and across research studies. Presently, however, the approach remains resource intensive and will continue to present challenges to large-scale studies. A possible measure to ameliorate the time and effort needed to implement this approach would be to capitalize on the emerging technologies of natural language processing to significantly reduce the load associated with the qualitative analysis and consequent scoring of learner responses to open-ended NOS instruments.

Acknowledgments

Thanks to my students John Myers and Ryan Summers for their respective help with searching the voluminous literature on NOS and examining the content of NOS instruments. Thanks also to Norman G. Lederman and Valarie Akerson, who reviewed this chapter.

References

Abd-El-Khalick, F. (1998). *The influence of history of science courses on students' conceptions of the nature of science.* Unpublished doctoral dissertation, Oregon State University, Oregon.

Abd-El-Khalick, F. (2001). Embedding nature of science instruction in preservice elementary science courses: Abandoning scientism, but . . . *Journal of Science Teacher Education, 12*(3), 215–233.

Abd-El-Khalick, F. (2002, April). *The development of conceptions of the nature of scientific knowledge and knowing in the middle and high school years: A cross-sectional study.* Paper presented at the annual meeting of the National Association for Research in Science Teaching, New Orleans, LA.

Abd-El-Khalick, F. (2004a). Over and over and over again: College students' views of nature of science. In L. B. Flick & N. G. Lederman (Eds.), *Scientific inquiry and nature of science: Implications for teaching, learning, and teacher education* (pp. 389–425). Dordrecht, the Netherlands: Kluwer.

Abd-El-Khalick, F. (2004b, April). *The relationship between students' views of nature of science and their conceptual understanding of stoichiometry: An empirical assessment.* Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Abd-El-Khalick, F. (2004c, April). *Assessing and influencing epistemological views: The moderate fallacies of inference to beliefs from action and influence of action on beliefs.* Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Abd-El-Khalick, F. (2005). Developing deeper understandings of nature of science: The impact of a philosophy of science course on preservice science teachers' views and instructional planning. *International Journal of Science Education, 27*(1), 15–42.

Abd-El-Khalick, F. (2012a). Examining the sources for our understandings about science: Enduring conflations and critical issues in research on nature of science in science education. *International Journal of Science Education, 34*(3), 353–374.

Abd-El-Khalick, F. (2012b). Nature of science in science education: Toward a coherent framework for synergistic research and development. In B. J. Fraser, K. Tobin, & C. McRobbie (Eds.), *Second international handbook of science education* (Vol. 2, pp. 1041–1060). Dordrecht, the Netherlands: Springer.

Abd-El-Khalick, F. (2013). Teaching *with* and *about* nature of science, and science teacher knowledge domains. *Science & Education, 22*(9), 2087–2107.

Abd-El-Khalick, F., & Akerson, V.L. (2004). Learning as conceptual change: Factors that mediate the development of preservice elementary teachers' views of nature of science. *Science Education, 88*(5), 785–810.

Abd-El-Khalick, F., & Akerson, V.L. (2009). The influence of metacognitive training on preservice elementary teachers' conceptions of nature of science. *International Journal of Science Education, 31*(16), 2161–2184.

Abd-El-Khalick, F., Belarmino, J., & Summers, R. (2012, March). *Development and validation of a rubric to score the views of nature of science questionnaire.* Paper presented at the annual meeting of the National Association for Research in Science Teaching, Indianapolis, IN.

Abd-El-Khalick, F., Bell, R.L., & Lederman, N.G. (1998). The nature of science and instructional practice: Making the unnatural natural. *Science Education, 82*(4), 417–436.

Abd-El-Khalick, F., & BouJaoude, S. (1997). An exploratory study of the knowledge base for science teaching. *Journal of Research in Science Teaching, 34*(7), 673–699.

Abd-El-Khalick, F., & BouJaoude, S. (2003). Lebanese middle school students' views of nature of science. *Mediterranean Journal of Educational Studies, 8*(1), 61–79.

Abd-El-Khalick, F., & Lederman, N.G. (2000a). Improving science teachers' conceptions of the nature of science: A critical review of the literature. *International Journal of Science Education, 22*(7), 665–701.

Abd-El-Khalick, F., & Lederman, N.G. (2000b). The influence of history of science courses on students' views of nature of science. *Journal of Research in Science Teaching, 37*(10), 1057–1095.

Abd-El-Khalick, F., Waters, M., & Le, A. (2008). Representations of nature of science in high school chemistry textbooks over the past four decades. *Journal of Research in Science Teaching, 45*(7), 835–855.

Aikenhead, G.S. (1973). The measurement of high school students' knowledge about science and scientists. *Science Education, 57*(4), 539–549.

Aikenhead, G.S. (1974a). Course evaluation. I: A new methodology for test construction. *Journal of Research in Science Teaching, 11*(1), 17–22.

Aikenhead, G.S. (1974b). Course evaluation. II: Interpretation of student performance on evaluative tests. *Journal of Research in Science Teaching, 11*(1), 23–30.

Aikenhead, G.S. (1987). High-school graduates' beliefs about science-technology-society. III. Characteristics and limitations of scientific knowledge. *Science Education, 71*(4), 459–487.

Aikenhead, G.S. (1988). An analysis of four ways of assessing student beliefs about STS topics. *Journal of Research in Science Teaching, 25*(8), 607–629.

Aikenhead, G.S. (1997). Student views on the influence of culture on science. *International Journal of Science Education, 19*(4), 419–428.

Aikenhead, G.S., Fleming, R.W., & Ryan, A.G. (1987). High-school graduates' beliefs about science-technology-society. I. Methods and issues in monitoring student views. *Science Education, 71*(2), 145–161.

Aikenhead, G.S., & Ryan, A.G. (1992). The development of a new instrument: "Views on Science-Technology-Society" (VOSTS). *Science Education, 76*(5), 477–491.

Aikenhead, G.S., Ryan, A., & Fleming, R. (1989). *Views on science-technology-society* (from CDN.mc.5). Saskatoon, Canada: Department of Curriculum Studies, University of Saskatchewan.

Akçay, B., & Koç, I. (2009). Inservice science teachers' views about the nature of science. *Hasan Ali Yücel Faculty of Education Journal, 11,* 1–11.

Akerson, V.L., & Abd-El-Khalick, F. (2003). Teaching elements of nature of science: A yearlong case study of a fourth grade teacher. *Journal of Research in Science Teaching, 40*(10), 1025–1049.

Akerson, V.L., Abd-El-Khalick, F., & Lederman, N.G. (2000). Influence of a reflective explicit activity-based approach on elementary teachers' conceptions of nature of science. *Journal of Research in Science Teaching, 37*(4), 295–317.

Akerson, V.L., Buck, G.A., Donnelly, L.A., Nargund-Joshi, V., & Weiland, I.S. (2011). The importance of teaching and learning nature of science in the early childhood years. *Journal of Science Teacher Education, 20,* 537–549.

Akerson, V.L., Cullen, T.A., & Hanson, D.L. (2009). Fostering a community of practice through a professional development program to improve elementary teachers' views of nature of science and teaching practice. *Journal of Research in Science Teaching, 46*(10), 1090–1113.

Akerson, V., & Donnelly, L.A. (2010). Teaching nature of science to K-2 students: What understandings can they attain? *International Journal of Science Education, 32*(1), 97–124.

Akerson, V.L., Hanson, D.L., & Cullen, T.A. (2007). The influence of guided inquiry and explicit instruction on K–6 teachers' views of nature of science. *Journal of Science Teacher Education, 18,* 751–772.

Akerson, V.L., & Hanuscin, D.L. (2007). Teaching the nature of science through inquiry: Results of a 3-year professional development programs. *Journal of Research in Science Teaching, 44*(5), 653–680.

Akerson, V.L., & Volrich, M.L. (2006). Teaching nature of science explicitly in a first-grade internship setting. *Journal of Research in Science Teaching, 43*(4), 377–394.

Akindehin, F. (1988). Effect of an instructional package on preservice science teachers' understanding of the nature of science and acquisition of science-related attitudes. *Science Education, 72*(1), 73–82.

Aldridge, J., Taylor, P., & Chen, C.C. (1997, March). *Development, validation and use of the beliefs about science and school science questionnaire (BASSSQ).* Paper presented at the annual meeting of the National Association for Research in Science Teaching, Oak Brook, IL.

Allen, H., Jr. (1959). *Attitudes of certain high school seniors toward science and scientific careers.* New York: Teachers College Press.

Alters, B.J. (1997). Whose nature of science? *Journal of Research in Science Teaching, 34*(1), 39–55.

American Association for the Advancement of Science (AAAS). (1990). *Science for all Americans.* New York: Oxford University Press.

Andersen, H.O., Harty, H., & Samuel, K.V. (1986). Nature of science, 1969 and 1984: Perspectives of preservice secondary science teachers. *School Science and Mathematics, 86*(1), 43–50.

Anderson, E.J., DeMelo, H.T., Szabo, M., & Toth, G. (1975). Behavioral objectives, science processes, and learning for inquiry-oriented instructional materials. *Science Education, 59*(2), 263–271.

Anderson, G.J. (1970). Effects of classroom social climate on individual learning. *American Educational Research Journal, 7*(2), 135–152.

Anderson, O.R., & Callaway, J. (1986). Studies of information processing rates in science learning and related cognitive variables. II: A first approximation to estimating a relationship between science reasoning skills and information acquisition. *Journal of Research in Science Teaching, 23*(1), 67–72.

Apostolou, A., & Koulaidis, V. (2010). Epistemology and science education: A study of epistemological views of teachers. *Research in Science & Technological Education, 28*(2), 149–166.

Barlow, N. (Ed.). (1993/1958). *The autobiography of Charles Darwin: 1809–1882.* New York: W.W. Norton.

Barufaldi, J.P., Bethel, L.J., & Lamb, W.G. (1977). The effect of a science methods course on the philosophical view of science among elementary education majors. *Journal of Research in Science Teaching, 14*(4), 289–294.

Bell, R.L. (1999). *Understandings of the nature of science and decision making on science and technology based issues.* Unpublished doctoral dissertation, Oregon State University, Oregon.

Bell, R.L., Lederman, N.G., & Abd-El-Khalick, F. (2000). Developing and acting upon one's conceptions of the nature of science: A follow-up study. *Journal of Research in Science Teaching, 37*(6), 563–581.

Ben-Chaim, D., & Zoller, U. (1991). The STS outlook profiles of Israeli high-school students and their teachers. *International Journal of Science Education, 13*(4), 447–458.

Bennett, J., & Hogarth, S. (2009). Would you want to talk to a scientist at a party? High school students' attitudes to school science and to science. *International Journal of Science Education, 31*(14), 1975–1998.

Beveridge, W.I.B. (1950). *The art of scientific investigation.* New York: Vintage Books.

Bianchini, J.A., & Solomon, E.R. (2003). Constructing views of science tied to issues of equity and diversity: A study of beginning science teachers. *Journal of Research in Science Teaching, 40,* 53–76.

Billeh, V.Y., & Hasan, O.E. (1975). Factors influencing teachers' gain in understanding the nature of science. *Journal of Research in Science Teaching, 12*(3), 209–219.

Billeh, V.Y., & Malik, M.H. (1977). Development and application of a test on understanding the nature of science. *Science Education, 61*(4), 559–571.

Biological Sciences Curriculum Study (BSCS). (1962). *Processes of science test.* New York: The Psychological Corporation.

Borda, E.J., Burgess, D.J., Plog, C.J., DeKalb, N.C., & Luce, M.M. (2009). Concept maps as tools for assessing students' epistemologies of science. *Electronic Journal of Science Education, 13*(2). Retrieved from http://ejse.southwestern.edu/article/view/7804

Botton, C., & Brown, C. (1998). The reliability of some VOSTS items when used with preservice secondary science teachers in England. *Journal of Research in Science Teaching, 35*(1), 53–71.

Boyd, R.N. (1983). On the current status of scientific realism. *Erkenntnis, 19,* 45–90.

Bradford, C.S., Rubba, P.A., & Harkness, W.L. (1995). Views about science-technology-society interactions held by college students in general education physics and STS courses. *Science Education, 79*(4), 355–373.

Breedlove, C.B., & Gessert, W.L. (1970). Use of an "electronic blackboard" in a physics teaching project. *School Science and Mathematics, 70*(2), 154–168.

Broadhurst, N.A. (1970). A study of selected learning outcomes of graduating high school students in south Australian schools. *Science Education, 54*(1), 17–21.

Buaraphan, K. (2009a). Preservice and inservice science teachers' responses and reasoning about the nature of science. *Educational Research and Review, 4*(11), 561–581.

Buaraphan, K. (2009b). Thai in-service science teachers' conceptions of the nature of science. *Journal of Science and Mathematics Education in Southeast Asia, 32*(2), 188–217.

Buaraphan, K. (2010). Pre-service and in-service science teachers' conceptions of the nature of science. *Science Educator, 19*(2), 35–47.

Buaraphan, K. (2011). Pre-service physics teachers' conceptions of nature of science. *US-China Education Review, 8*(2), 137–148.

Buaraphan, K. (2012). Embedding nature of science in teaching about astronomy and space. *Journal of Science Education and Technology, 21,* 353–369.

Buaraphan, K., & Sung-Ong, S. (2009). Thai pre-service science teachers' conceptions of the nature of science. *Asia-Pacific Forum on Science Learning and Teaching, 10*(1). Retrieved from www.ied.edu.hk/apfslt/

Buffler, A., Lubben, F., & Ibrahim, B. (2009). The relationship between students' views of the nature of science and their views of the nature of scientific measurement. *International Journal of Science Education, 31*(9), 1137–1156.

Cakmakci, G. (2012). Promoting pre-service teachers' ideas about nature of science through educational research apprenticeship. *Australian Journal of Teacher Education, 37*(2). Retrieved from http://ro.ecu.edu.au/ajte/vol37/iss2/8

Carey, R.L., & Stauss, N.G. (1968). An analysis of the understanding of the nature of science by prospective secondary science teachers. *Science Education, 52*(4), 358–363.

Carey, R.L., & Stauss, N.G. (1970). An analysis of experienced science teachers' understanding of the nature of science. *School Science and Mathematics, 70*(5), 366–376.

Chai, C.C., Deng, F., & Tsai, C.C. (2012). A comparison of scientific epistemological views between Mainland China and Taiwan high school students. *Asia Pacific Education Review, 13*(1), 17–26.

Chan, K.S. (2005). Exploring the dynamic interplay of college students' conceptions of the nature of science. *Asia-Pacific Forum on Science Learning and Teaching, 6*(2), 1–16.

Chen, S. (2006). Development of an instrument to assess views on nature of science and attitudes toward teaching science. *Science Education, 90,* 803–819.

Choi, K., & Cho, H.H. (2002). Effects of teaching ethical issues on Korean school students' attitudes towards science. *Journal of Biological Education, 37*(1), 26–30.

Choppin, B.H. (1974). The introduction of new science curricula in England and Wales. *Comparative Education Review, 18*(2), 196–206.

Chrouser, W.H. (1975). Outdoor vs. indoor laboratory techniques in teaching biology to prospective elementary teachers. *Journal of Research in Science Teaching, 12*(1), 41–48.

Cobern, W.W., & Loving, C.C. (2002). Investigation of preservice elementary teachers' thinking about science. *Journal of Research in Science Teaching, 39*(10), 1016–1031.

Cooley, W.W., & Bassett, R.D. (1961). Evaluation and follow-up study of a summer science and mathematics program for talented secondary school students. *Science Education, 45*(3), 209–216.

Cooley, W.W., & Klopfer, L.E. (1961). *TOUS: Test on understanding science.* Princeton, NJ: Education Testing Service.

Cossman, G.W. (1969). The effects of a course in science and culture for secondary school students. *Journal of Research in Science Teaching, 6*(3), 274–283.

Cotham, J.S., & Smith, E.L. (1981). Development and validation of the conceptions of scientific theories test. *Journal of Research in Science Teaching, 18*(5), 387–396.

Crumb, G.H. (1965). Understanding of science in high school physics. *Journal of Research in Science Teaching, 3,* 246–250.

Curd, M., & Cover, J.A. (1998). *Philosophy of science: The central issues.* New York: W.W. Norton & Co.

Dass, P.M. (2005). Understanding the nature of scientific enterprise (NOSE) through a discourse with its history: The influence of an undergraduate "history of science" course. *International Journal of Science and Mathematics Education, 3,* 87–115.

Demirbaş, M., & Balci, F. (2012). The impact of the explicit reflective approach in teaching the nature of science upon Turkish students' perceptions of science. *International Journal of Asian Social Science, 2*(8), 1255–1260.

Demirbas, M., Bozdogan, A.E., & Özbek, G. (2012). An analysis from different variables of views of pre-service science teachers in Turkey on the nature of science. *Research Journal of Recent Sciences, 1*(8), 29–35.

Dienye, N.E. (1987). The effect of inservice science education. *British Journal of In-Service Education, 14*(1), 48–55.

Dogan, N., & Abd-El-Khalick, F. (2008). Turkish grade 10 students' and science teachers' conceptions of nature of science: A national study. *Journal of Research in Science Teaching, 45*(10), 1083–1112.

Doran, R.L., Guerin, R.O., & Cavalieri, J. (1974). An analysis of several instruments measuring "nature of science" objectives. *Science Education, 58*(3), 312–329.

Driver, R., Leach, J., Millar, R., & Scott, P. (1996). *Young people's images of science.* Buckingham, UK: Open University Press.

Duncan, D.K., & Arthurs, L. (2012). Improving student attitudes about learning science and student scientific reasoning skills. *Astronomy Education Review, 11*(1). Retrieved from http://aer.aas.org/resource/1/aerscz/v11/i1/p010102_s1

Durkee, P. (1974). An analysis of the appropriateness and utilization of TOUS with special reference to high-ability students studying physics. *Science Education, 58*(3), 343–356.

Duschl, R.A., & Wright, E. (1989). A case study of high school teachers' decision making models for planning and teaching science. *Journal of Research in Science Teaching, 26*(5), 467–501.

Eastwood, J.L., Sadler, T.D., Zeidler, D.L., Lewis, A., Amiri, L., & Applebaum, S. (2012). Contextualizing nature of science instruction in socioscientific issues. *International Journal of Science Education, 34*(15), 2289–2315.

Engen, H.B., Smith, D.D., & Yager, R.E. (1967–1968). Student outcomes with rotation of teachers in secondary biology. *Journal of Research in Science Teaching, 5,* 230–234.

Faryniarz, J.V. (1992). Effectiveness of microcomputer simulations in stimulating environmental problem solving by community college students. *Journal of Research in Science Teaching, 29*(5), 453–470.

Fergusson, J., Oliver, C., & Walter, M.R. (2012). Astrobiology outreach and the nature of science: The role of creativity. *Astrobiology, 12*(12), 1143–1153.

Fisher, D.L. (1979). The impact of the inclusion of ASEP materials on some cognitive outcomes in different types of Tasmanian schools. *Research in Science Education, 9,* 111–118.

Fisher, D.L., & Fraser, B.J. (1980). Evaluating the impact of a national curriculum on content-free cognitive outcomes. *European Journal of Science Education, 2*(1), 45–59.

Fleming, R.W. (1987). High-school gradates' beliefs about science-technology-society. II. The interaction among science, technology, and society. *Science Education, 71*(2), 163–186.

Fleming, R.W. (1988). Undergraduate science students' views on the relationship between science, technology and society. *International Journal of Science Education, 10*(4), 449–463.

Flores, F., Lopez, A., Gallegos, L., & Barojas, J. (2000). Transforming science and learning concepts of physics teachers. *International Journal of Science Education, 22*(2), 197–208.

Folmer, V., Barbosa, N.B., Soares, F.A., & Rocha, J.B.T. (2009). Experimental activities based on ill-structured problems improve Brazilian school students' understanding of the nature of scientific knowledge. *Revista Electrónica de Enseñanza de las Ciencias, 8*(1), 232–254.

Fraser, B.J. (1978). Development of a test of science-related attitudes. *Science Education, 62,* 509–515.

Fulton, H.F. (1971a). An analysis of student outcomes utilizing two approaches to teaching BSCS biology. *Journal of Research in Science Teaching, 8*(1), 21–28.

Fulton, H.F. (1971b). Individualized vs. group teaching of BSCS biology. *The American Biology Teacher, 33*(5), 277–279, 291.

Gardner, P.L. (1975). Attitudes to science: A review. *Studies in Science Education, 2*(1), 1–41.

Gess-Newsome, J. (2002). The use and impact of explicit instruction about the nature of science and science inquiry in an elementary science methods course. *Science & Education, 11,* 55–67.

Ginns, I.S., & Foster, W.J. (1978). A comparison of teaching strategies and their effect on attitudes to and understanding of science. *South Pacific Journal of Teacher Education, 6*(2), 154–158.

Glass, L.W., & Yager, R.E. (1970). Individualized instruction as a spur to understanding the scientific enterprise. *The American Biology Teacher, 32*(6), 359–361.

Goff, P., Boesdorfer, S.B., & Hunter, W. (2012). Using a multicultural approach to teach chemistry and the nature of science to undergraduate non-majors. *Cultural Studies of Science Education, 7,* 631–651.

Golabek, C., & Amrane-Cooper, L. (2011). Trainee teachers' perceptions of the nature of science and implications for pre-service teacher training in England. *Research in Secondary Teacher Education, 1*(2), 9–13.

Good, R., & Shymansky, J. (2001). Nature-of-science literacy in *Benchmarks* and *Standards:* Post-modern/relativist or modern/realist? *Science & Education, 10,* 173–185.

Guerra-Ramos, M.T. (2012). Teachers' ideas about the nature of science: A critical analysis of research approaches and their contribution to pedagogical practice. *Science & Education, 21*(5), 631–655.

Guerra-Ramos, M.T., Ryder, J., & Leach, J.J. (2010). Ideas about the nature of science in pedagogically relevant contexts: Insights from a situated perspective of primary teachers' knowledge. *Science Education, 94,* 282–307.

Güzel, H. (2011). Opinions of university students about the nature of science. *World Applied Sciences Journal, 12*(7), 1005–1013.

Hacıeminoğlu, E., Yılmaz-Tüzün, O., & Ertepınar, H. (2012). Development and validation of nature of science instrument for elementary school students. *Education 3–13: International Journal of Primary, Elementary and Early Years Education.* doi:10.1080/03004279.2012.671840

Haidar, A.H. (1999). Emirates pre-service and inservice teachers' views about the nature of science. *International Journal of Science Education, 21*(8), 807–822.

Haidar, A.H. (2000). Professors' views on the influence of Arab society on science and technology. *Journal of Science Education and Technology, 9*(3), 257–273.

Hanuscin, D.L., Akerson, V.L., & Phillipson-Mower, T. (2006). Integrating nature of science instruction into a physical science content course for preservice elementary teachers: NOS views of teaching assistants. *Science Education, 90,* 912–935.

Hanuscin, D.L., & Lee, M.H. (2011). Elementary teachers' pedagogical content knowledge for teaching the nature of science. *Science Education, 95,* 145–167.

Haukoos, G.D., & Penick, J.E. (1983). The influence of classroom climate on science process and content achievement of community college students. *Journal of Research in Science Teaching, 20*(7), 629–637.

Haukoos, G.D., & Penick, J.E. (1985). The effects of classroom climate on college science students: A replication study. *Journal of Research in Science Teaching, 22*(2), 163–168.

Haukoos, G.D., & Penick, J.E. (1987). Interaction effect of personality characteristics, classroom climate, and science achievement. *Science Education, 71*(5), 735–743.

Hillis, S.R. (1975). The development of an instrument to determine student views of the tentativeness of science. In E.J. Montague (Ed.), *Research and curriculum development in science education: Science teacher behavior and student affective and cognitive learning* (Vol. 3, pp. 34–40). Austin: University of Texas Press. (ERIC Document Reproduction Service No. ED 124 404)

Hodson, D. (1991). Philosophy of science and science education. In M.R. Matthews (Ed.), *History, philosophy, and science teaching: Selected readings* (pp. 19–32). New York: Teachers College Press.

Holton, G., Watson, F.G., & Rutherford, F.J. (1967). Harvard project physics: A progress report. *The Physics Teacher, 5*(5).

Howe, E.M. (2007). Addressing nature-of-science core tenets with the history of science: An example with sickle-cell anemia & malaria. *American Biology Teacher, 69*(8), 467–472.

Howe, E.M., & Rudge, D.W. (2005). Recapitulating the history of sickle-cell anemia research: Improving students' NOS views explicitly and reflectively. *Science & Education, 14,* 423–441.

Huang, C.M., Tsai, C.C., & Chang, C.Y. (2005). An investigation of Taiwanese early adolescents' views about the nature of science. *Adolescence, 40*(159), 645–654.

Hungerford, H., & Walding, H. (1974). *The modification of elementary methods students' concepts concerning science and scientists.* Paper presented at the annual meeting of the National Science Teachers Association.

İrez, S. (2006). Are we prepared? An assessment of preservice science teacher educators' beliefs about nature of science. *Science Education, 90,* 1113–1143.

İrez, S. (2007). Reflection-oriented qualitative approach in beliefs research. *Eurasia Journal of Mathematics, Science & Technology Education, 3*(1), 17–27.

İrez, S., Çakir, M., & Şeker, H. (2011). Exploring nature of science understandings of Turkish pre-service science teachers. *Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education, 5*(2), 6–17.

Irzik, G., & Nola, R. (2011). A family resemblance approach to the nature of science for science education. *Science & Education, 20*(7), 591–607.

Jacobs, J.N., & Bollenbacher, J.K. (1960). Teaching ninth-grade biology by television. *Audiovisual Communication Review, 8,* 176–191.

Johnson, G. (1972). An integrated two year chemistry-physics course compared with consecutively taught separate courses. *Science Education, 56*(2), 143–154.

Johnson, R.L., & Peeples, E.E. (1987). The role of scientific understanding in college: Student acceptance of evolution. *American Biology Teacher, 49*(2), 96–98.

Jones, K.M. (1969). The attainment of understandings about the scientific enterprise, scientists, and the aims and methods of science by students in a college physical science course. *Journal of Research in Science Teaching, 6*(1), 47–49.

Jungwirth, E. (1968). Teaching for "understanding of science." *Journal of Biological Education, 2*(1), 39–51.

Jungwirth, E. (1969). Active understanding of the processes of science: An evaluation of certain aspects of the first two years of B.S.C.S. biology teaching in Israel. *Journal of Biological Education, 3*(1), 45–55.

Jungwirth, E., & Jungwirth, E. (1972). TOUS revisited: A longitudinal study of the development of understanding of science. *Journal of Biological Education, 6*(3), 187–195.

Kang, S., Scharmann, L.C., & Noh, T. (2005). Examining students' views on the nature of science: Results from Korean 6th, 8th, and 10th graders. *Science Education, 89,* 314–334.

Kaya, S. (2012). An examination of elementary and early childhood pre-service teachers' nature of science views. *Procedia-Social and Behavioral Science, 46,* 581–585.

Kelly, G.J., Chen, C., & Crawford, T. (1998). Methodological considerations for studying science-in-the-making in educational settings. *Research in Science Education, 28*(1), 23–49.

Kemeny, J.G. (1959). *A philosopher looks at science.* Princeton, NJ: D. Van Nostrand.

Khishfe, R. (2008). The development of seventh graders' views of nature of science. *Journal of Research in Science Teaching, 45*(4), 470–496.

Khishfe, R. (2012). Relationship between nature of science understandings and argumentation skills: A role for counterargument and contextual factors. *Journal of Research in Science Teaching, 49*(4), 489–514.

Khishfe, R., & Abd-El-Khalick, F. (2002). Influence of explicit reflective versus implicit inquiry-oriented instruction on sixth graders' views of nature of science. *Journal of Research in Science Teaching, 39*(7), 551–578.

Kim, S.Y., & Irving, K.E. (2010). History of science as an instructional context: Student learning in genetics and nature of science. *Science & Education, 19,* 187–215.

Kim, S.Y., & Nehm, R.H. (2011). Sun Young & Ross H.: A cross-cultural comparison of Korean and American science teachers' views of evolution and the nature of science. *International Journal of Science Education, 33*(2), 197–227.

Kimball, M.E. (1967–1968). Understanding the nature of science: A comparison of scientists and science teachers. *Journal of Research in Science Teaching, 5,* 110–120.

Klopfer, L.E., & Cooley, W.W. (1963). The *history of science cases for high schools* in the development of student understanding of science and scientists: A report on the HOSC instruction project. *Journal of Research in Science Teaching, 1*(1), 33–47.

Kokkotas, P., Piliouras, P., Malamitsa, K., & Stamoulis, E. (2009). Teaching physics to in-service primary schools teachers in the context of the history of science: The case of falling bodies. *Science & Education, 18,* 609–629.

Korth, W.W. (1968). *The use of history of science to promote student's understanding of the social aspects of science.* Unpublished doctoral dissertation, Stanford University, Stanford, California.

Korth, W.W. (1969, February). *Test every senior project: Understanding the social aspects of science.* Paper presented at the annual meeting of the National Association for Research in Science Teaching, Pasadena, CA. (ERIC Document Reproduction Service No. ED 028 087)

Koulaidis, V., & Ogborn, J. (1988). Use of systemic networks in the development of a questionnaire. *International Journal of Science Education, 10*(5), 497–509.

Koulaidis, V., & Ogborn, J. (1989). Philosophy of science: An empirical-study of teachers views. *International Journal of Science Education, 11*(2), 173–184.

Koulaidis, V., & Ogborn, J. (1995). Science teachers philosophical assumptions: How well do we understand them? *International Journal of Science Education, 17*(3), 273–283.

Krockover, G.H. (1971). Individualizing secondary school chemistry instruction. *School Science and Mathematics, 71*(6), 518–524.

Kucuk, M. (2008). Improving preservice elementary teachers' views of the nature of science using explicit-reflective teaching in a science, technology and society course. *Australian Journal of Teacher Education, 33*(2). Retrieved from http://ro.ecu.edu.au/ajte/vol33/iss2/1/

Kuhn, T.S. (1996). *The structure of scientific revolutions* (3rd ed.). Chicago: University of Chicago Press. (First published 1962)

Lachman, S.J. (1960). *The foundations of science.* New York: Vantage.

Laforgia, J. (1988). The affective domain related to science education and its evaluation. *Science Education, 72*(4), 407–421.

Lakatos, I. (1970). Falsification and the methodology of scientific research programmes. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 91–195). New York: Cambridge University Press.

Lakin, S., & Wellington, J. (1994). Who will teach the "nature of science"? Teachers' views of science and their implications for science education. *International Journal of Science Education, 16*(2), 175–190.

Laudan, L., Donovan, A., Laudan, R., Barker, P., Brown, H., Leplin, J., Thagard, P., & Wykstra, S. (1986). Scientific change: Philosophical models and historical research. *Synthese, 69,* 141–223.

Lavach, J.F. (1969). Organization and evaluation of an in-service program in the history of science. *Journal of Research in Science Teaching, 6,* 166–170.

Lawrenz, F. (1975). The relationship between science teacher characteristics and student achievement and attitude. *Journal of Research in Science Teaching, 12*(4), 433–437.

Lawrenz, F., & Cohen, H. (1985). The effect of methods classes and practice teaching on student attitudes toward science and knowledge of science processes. *Science Education, 69*(1), 105–113.

Lawson, A.E., Nordland, F.H., & DeVito, A. (1975). Relationship of formal reasoning to achievement, aptitudes, and attitudes in preservice teachers. *Journal of Research in Science Teaching, 4,* 423–431.

Leach, J., Millar, R., Ryder, J., & Séré, M.-G. (2000). Epistemological understanding in science learning: The consistency of representations across contexts. *Learning and Instruction, 10,* 497–527.

Leake, J.B., & Hinerman, C.O. (1973). Scientific literacy and school characteristics. *School Science and Mathematics, 73*(9), 772–782.

Leblebicioğlu, G., Metin, D., Yardimci, E., & Berkyürek, I. (2011). Teaching the nature of science in the nature: A summer science camp. *Elementary Education Online, 10*(3), 1037–1055.

Leblebicioğlu, G., Metin, D., Yardimci, E., & Cetin, P.S. (2011). The effect of informal and formal interaction between scientists and children at a science camp on their images of scientists. *Science Education International, 22*(3), 158–174.

Lederman, J.S., & Khishfe, R. (2002). *Views of nature of science, Form D.* Unpublished paper. Illinois Institute of Technology, Chicago, IL.

Lederman, J.S., & Ko, E.K. (2002). *Views of nature of science, Form E.* Unpublished paper. Illinois Institute of Technology, Chicago, IL.

Lederman, N.G. (1986a). Relating teacher behavior and classroom climate to changes in students' conceptions of the nature of science. *Science Education, 70*(1), 3–19.

Lederman, N.G. (1986b). Students' and teachers' understanding of the nature of science: A reassessment. *School Science and Mathematics, 86*(2), 91–99.

Lederman, N.G. (1992). Students' and teachers' conceptions of the nature of science: A review of the research. *Journal of Research in Science Teaching, 29*(4), 331–359.

Lederman, N.G. (1999). Teachers' understanding of the nature of science and classroom practice: Factors that facilitate or impede the relationship. *Journal of Research in Science Teaching, 36,* 916–929.

Lederman, N.G. (2007). Nature of science: Past, present, and future. In S.K. Abell & N.G. Lederman (Eds.), *Handbook of research on science education* (pp. 831–879). Mahwah, NJ: Lawrence Erlbaum.

Lederman, N.G., Abd-El-Khalick, F., Bell, R.L., & Schwartz, R. (2002). Views of nature of science questionnaire (VNOS): Toward valid and meaningful assessment of learners' conceptions of nature of science. *Journal of Research in Science Teaching, 39*(6), 497–521.

Lederman, N.G., & Druger, M. (1985). Classroom factors related to changes in students' conceptions of the nature of science. *Journal of Research in Science Teaching, 22*(7), 649–662.

Lederman, N.G., & O'Malley, M. (1990). Students' perceptions of tentativeness in science: Development, use, and sources of change. *Science Education, 74,* 225–239.

Lederman, N.G., Schwartz, R., Abd-El-Khalick, F., & Bell, R.L. (2001). Preservice teachers' understanding and teaching of nature of science: An intervention study. *Canadian Journal of Science, Mathematics and Technology Education, 1*(2), 135–160.

Lederman, N.G., Wade, P., & Bell, R. (1998). Assessing the nature of science: What is the nature of our assessments? *Science and Education, 7,* 595–615.

Lederman, N.G., & Zeidler, D.L. (1987). Science teachers' conceptions of the nature of science: Do they really influence teaching behavior? *Science Education, 71*(5), 721–734.

Liang, L.L., Chen, S., Chen, X., Kaya, O.N., Adams, A.D., Macklin, M., & Ebenezer, J. (2006, April). *Student understanding of science and scientific inquiry (SUSSI): Revision and further validation of an assessment instrument.* Paper presented at the annual meeting of the National Association for Research in Science Teaching, San Francisco, CA.

Liang, L.L., Chen, S., Chen, X., Kaya, O.N., Adams, A.D., Macklin, M., & Ebenezer, J. (2008). Assessing preservice elementary teachers' views on the nature of scientific knowledge: A dual-response instrument. *Asia-Pacific Forum on Science Learning and Teaching, 9*(1). Retrieved from www.ied.edu.hk/apfslt/v9_issue1/liang/liang3.htm

Liang, L.L., Chen, S., Chen, X., Kaya, O.N., Adams, A.D., Macklin, M., & Ebenezer, J. (2009). Preservice teachers' views about nature of scientific knowledge development: An international collaborative study. *International Journal of Science and Mathematics Education, 7,* 987–1012.

Lin, C.C., & Tsai, C.C. (2008). Exploring the structural relationships between high school students' scientific epistemological views and their utilization of information commitments toward online science information. *International Journal of Science Education, 30*(15), 2001–2022.

Lin, H.S., & Chen, C.C. (2002). Promoting preservice chemistry teachers' understanding about the nature of science through history. *Journal of Research in Science Teaching, 39*(9), 773–792.

Lin, H.S., & Chiu, H.L. (2004). Student understanding of the nature of science and their problem-solving strategies. *International Journal of Science Education, 26*(1), 101–112.

Liu, S.Y., & Lederman, N.G. (2007). Exploring prospective teachers' worldviews and conceptions of nature of science. *International Journal of Science Education, 29*(10), 1281–1307.

Liu, S.Y., Lin, C.S., & Tsai, C.C. (2011). College students' scientific epistemological views and thinking patterns in socioscientific decision making. *Science Education, 95*(3), 497–517.

Liu, S.Y., & Tsai, C.C. (2008). Differences in the scientific epistemological views of undergraduate students. *International Journal of Science Education, 30*(8), 1055–1073.

Lombrozo, T., Thanukos, A., & Weisberg, M. (2008). The importance of understanding the nature of science for accepting evolution. *Evolution: Education and Outreach, 1,* 290–298.

Lonsbury, J.G., & Ellis, J.D. (2002). Science history as a means to teach nature of science concepts: Using the development of understanding related to mechanisms of inheritance. *Electronic Journal of Science Education, 7*(2). Retrieved from http://ejse.southwestern.edu/article/view/7703/5470

Loving, C.C. (1997). From the summit of truth to its slippery slopes: Science education's journey through positivist-postmodern territory. *American Educational Research Journal, 34*(3), 421–452.

Lowery, L.F. (1967). An experimental investigation into the attitudes of fifth grade students toward science. *School Science and Mathematics, 67,* 569–579.

Mackay, L.D. (1971). Development of understanding about the nature of science. *Journal of Research in Science Teaching, 8*(1), 57–66.

Mackay, L.D., & White, R.T. (1974). Development of an alternative to Likert scaling: Tests of perceptions of scientists and self (TOPOSS). *Research in Science Education, 4*(1), 131–139.

Marbach-Ad, G., McGinnis, J.R., Dai, A.H., Pease, R., Schalk, K.A., & Benson, S. (2009). Promoting science for all by way of student interest in a transformative undergraduate microbiology laboratory for nonmajors. *Journal of Microbiology & Biology Education, 10,* 58–67.

Marchlewicz, S.C., & Wink, D.J. (2011). Using the activity model of inquiry to enhance general chemistry students' understanding of nature of science. *Journal of Chemical Education, 88,* 1041–1047.

Markle, G., & Capie, W. (1977). Assessing a competency-based physics course: A model for evaluating science courses serving elementary teachers. *Journal of Research in Science Teaching, 14*(2), 151–156.

Mayer, V.J., & Richmond, J.M. (1982). An overview of assessment instruments in science. *Science Education, 66*(1), 49–66.

Mbajiorgu, N.M., & Ali, A. (2003). Relationship between STS approach, scientific literacy, and achievement in biology. *Science Education, 87,* 31–39.

McComas, W.F., & Olson J.K. (1998). The nature of science in international science education standards documents. In W.F. McComas (Ed.), *The nature of science in science education: Rationales and strategies* (pp. 41–52). Dordrecht, the Netherlands: Kluwer.

McDonald, C.V. (2010). The influence of explicit nature of science and argumentation instruction on preservice primary teachers' views of nature of science. *Journal of Research in Science Teaching, 47*(9), 1137–1164.

Mead, M., & Métraux, R. (1957). Image of the scientist among high-school students: A pilot study. *Science, 126,* 384–390.

Meichtry, Y.J. (1992). Influencing student understanding of the nature of science: Data from a case of curriculum development. *Journal of Research in Science Teaching, 29*(4), 389–407.

Mellado, V. (1997). Preservice teachers' classroom practice and their conceptions of the nature of science. *Science & Education, 6,* 331–354.

Metin, D., & Leblebicioğlu, G. (2011). How did a science camp affect children's conceptions of science? *Asia-Pacific Forum on Science Learning and Teaching, 12*(1). Retrieved from www.ied.edu.hk/apfslt/v12_issue1/kilic/index.htm

Metin, D., & Leblebicioğlu, G. (2012). Effect of a science camp on the children's views of tentative nature of science. *Journal of Studies in Education, 2*(1). Retrieved from www.macrothink.org/journal/index.php/jse/article/view/1348

Meyer, J.H. (1969). The influence of the invitations to enquiry. *The American Biology Teacher, 31*(7), 451–453.

Miller, M.C.D., Montplaisir, L.M., Offerdahl, E.G., Cheng, F.C., & Ketterling, G.L. (2010). Comparison of views of the nature of science between natural science and nonscience majors. *CBE Life Sciences Education, 9,* 45–54.

Moore, R., & Sutman, F. (1970). The development of field test and validation of an inventory of scientific attitudes. *Journal of Research in Science Teaching, 7,* 85–94.

Munby, H. (1982). The impropriety of "panel of judges" validation in science attitude scales. *Journal of Research in Science Teaching, 19,* 617–619.

Murcia, K., & Schibeci, R. (1999). Primary student teachers' conceptions of the nature of science. *International Journal of Science Education, 21*(11), 1123–1140.

Musgrave, A. (1998). Realism versus constructive empiricism. In M. Curd & J.A. Cover (Eds.), *Philosophy of science: The central issues* (pp. 1088–1113). New York: Norton.

Nalçaci, Ġ. Ö., Akarsu, B., & Kariper, A. I (2011). Effects of the nature of science course on science prospective teachers' knowledge and opinions. *Ahmet Kelebodlu Journal of Education, 32,* 337–352.

National Research Council (NRC). (1996). *National science education standards.* Washington, DC: National Academies Press.

Neumann, I., Neumann, K., & Nehm, R. (2011). Evaluating instrument quality in science education: Rasch-based analyses of a nature of science test. *International Journal of Science Education, 33*(10), 1373–1405.

NGSS Lead States. (2013). *Next generation science standards: For states, by states.* Washington, DC: National Academies Press.

Nott, M., & Wellington, J. (1995). Critical incidents in the science classroom and the nature of science. *School Science Review, 76,* 41–46.

Nott, M., & Wellington, J. (1995b). Probing teachers' views of the nature of science: How should we do it and where should we be looking? In F. Finley et al. (Eds.), *Proceedings of the Third International History, Philosophy, and Science Teaching Conference* (Vol. 2, pp. 864–872). Minneapolis: University of Minnesota.

Nott, M., & Wellington, J. (1996). When the black box springs open: Practical work in school science and the nature of science. *International Journal of Science Education, 18*(7), 807–818.

Nott, M., & Wellington, J. (1998). Eliciting, interpreting and developing teachers' understandings of the nature of science. *Science & Education, 7,* 579–594.

Ogunniyi, M.B. (1982). An analysis of prospective science teachers' understanding of the nature of science. *Journal of Research in Science Teaching, 19*(1), 25–32.

Ogunniyi, M.B. (1983). Relative effects of a history/philosophy of science course on student teachers' performance on two models of science. *Research in Science & Technological Education, 1*(2), 193–199.

O'Hear, A. (1989). *An introduction to the philosophy of science.* New York: Oxford University Press.

Olstad, R.G. (1969). The effect of science teaching methods on the understanding of science. *Science Education, 53*(1), 9–11.

Orgren, J., & Doran, R.L. (1975). The effects of adopting the revised New York state regents earth science syllabus on selected teacher and student variables. *Journal of Research in Science Teaching, 12*(1), 15–24.

Osborne, J., Collins, S., Ratcliffe, M., Millar, R., & Duschl, R. (2003). What "ideas-about science" should be taught in school science? A Delphi study of the expert community. *Journal of Research in Science Teaching, 40*(7), 692–720.

Osborne, J., Simon, S., & Collins, S. (2003). Attitudes toward science: A review of the literature and its implications. *International Journal of Science Education, 25*(9), 1049–1079.

Oskay, Ö. Ö., Yilmaz, A., Dinçol, S., & Erdem, A. (2011). Determination of relationship between prospective chemistry teachers' scientific epistemological views, information commitments and online searching achievement. *Procedia Social and Behavioral Sciences, 15,* 3484–3489.

Pearl, R.E. (1974). The present status of science attitude measurement: History, theory, and availability of measurement. *School Science and Mathematics, 74*(5), 375–381.

Pegg, J., & Gummer, E. (2010). The influence of a multidisciplinary scientific research experience on teachers views of nature of science. *The Montana Mathematics Enthusiast, 7*(2 & 3), 447–460.

Pella, M.O., O'Hearn, G.T., & Gale, C.W. (1966). Referents to scientific literacy. *Journal of Research in Science Teaching, 4,* 199–208.

Pomeroy, D. (1993). Implications of teachers' beliefs about the nature of science: Comparison of the beliefs of scientists, secondary science teachers, and elementary teachers. *Science Education, 77,* 261–278.

Popper, K.R. (1963). *Conjectures and refutations: The growth of scientific knowledge.* London, UK: Routledge & Kegan Paul.

Posnanski, T.J. (2010). Developing understanding of the nature of science within a professional development program for inservice elementary teachers: Project nature of elementary science teaching. *Journal of Science Teacher Education, 21,* 589–621.

Quigley, C., Pongsanon, K., & Akerson, V.L. (2010). If we teach them, they can learn: Young students' views of nature of science aspects to early elementary students during an informal science education program. *Journal of Science Teacher Education, 21,* 887–907.

Quigley, C., Pongsanon, K., & Akerson, V.L. (2011). If we teach them, they can learn: Young students' views of nature of science during an informal science education program. *Journal of Science Teacher Education, 22,* 129–149.

Rampal, A. (1992). Images of science and scientists: A study of school teachers' views. I. Characteristics of scientists. *Science Education, 76*(4), 415–436.

Riban, D.B. (1976). Examination of a model for field studies in science. *Science Education, 60*(1), 1–11.

Riban, D.B., & Koval, D.B. (1971). An investigation of the effect of field studies in science on the learning of the methodology of science. *Science Education, 55*(3), 291–294.

Richardson, J.S., & Showalter, V. (1967). *Effects of a unified science curriculum on high school graduates.* Columbus: Ohio State University. (ERIC Document Reproduction Service No. 024 593)

Riley, J.P., II (1979). The influence of hands-on science process training on preservice teachers' acquisition of process skills and attitude toward science and science teaching. *Journal of Research in Science Teaching, 16*(5), 373–384.

Rivers, R.H., & Vockell, E. (1987). Computer simulations to stimulate scientific problem solving. *Journal of Research in Science Teaching, 24*(5), 403–415.

Roehrig, G.H., & Luft, J.A. (2004). Constraints experienced by beginning secondary science teachers in implementing scientific inquiry lessons. *International Journal of Science Education, 26*(1), 3–24.

Rosenberg, A. (2005). *Philosophy of science: A contemporary introduction* (2nd ed.). New York: Routledge.

Rothman, A.I. (1969). Teacher characteristics and student learning. *Journal of Research in Science Teaching, 6,* 340–348.

Rothman, A.I., Welch, W.W., & Walberg, H.J. (1969). Physics teacher characteristics and student learning. *Journal of Research in Science Teaching, 6,* 59–63.

Rubba, P. (1976). *Nature of scientific knowledge scale.* Bloomington: School of Education, Indiana University.

Rubba, P. (1977). *The development, field testing, and validation of an instrument to assess secondary school students' understanding of the nature of scientific knowledge.* Unpublished doctoral dissertation, Indiana University, Bloomington, IN.

Rubba, P., & Harkness, W.L. (1993). Examination of preservice and in-service secondary science teachers' beliefs about science-technology-society interactions. *Science Education, 77*(4), 407–431.

Rubba, P.A., & Andersen, H. (1978). Development of an instrument to assess secondary school students' understanding of the nature of scientific knowledge. *Science Education, 62*(4), 449–458.

Rubba, P.A., Horner, J.K., & Smith, J.M. (1981). A study of two misconceptions about the nature of science among junior high school students. *School Science and Mathematics, 81*(3), 221–226.

Rudge, D.W., Cassidy, D.P., Fulford, J.M., & Howe, E.M. (2013). Changes observed in views of nature of science during a historically based unit. *Science & Education.* doi:10.1007/s11191-012-9572-3

Rudolph, J.L. (2000). Reconsidering the "nature of science" as a curriculum component. *Journal of Curriculum Studies, 32*(3), 403–419.

Russell, C.B., & Weaver, G.C. (2011). A comparative study of traditional, inquiry-based, and research-based laboratory curricula: Impacts on understanding of the nature of science. *Chemistry Education Research and Practice, 12,* 57–67.

Russell, T., & Aydeniz, M. (2012). Traversing the divide between high school science students and sophisticated nature of science understandings: A multi-pronged approach. *Journal of Science Education and Technology.* doi:10.1007/s10956-012-9412-x

Ryan, A.G., & Aikenhead, G.S. (1992). Students' preconceptions about the epistemology of science. *Science Education, 76,* 559–580.

Ryder, J., Leach, J., & Driver, R. (1999). Undergraduate science students' images of science. *Journal of Research in Science Teaching, 36*(2), 201–219.

Salloum, S., & Abd-El-Khalick, F. (2010). A study of practical-moral knowledge in science teaching: Case studies in physical science classrooms. *Journal of Research in Science Teaching, 47*(8), 929–951.

Sandoval, W.A., & Morrison, K. (2003). High school students' ideas about theories and theory change after a biological inquiry unit. *Journal of Research in Science Teaching, 40*(4), 369–392.

Sarkar, M.A., & Gomes, J.J. (2010). Science teachers' conceptions of nature of science: The case of Bangladesh. *Asia-Pacific Forum on Science Learning and Teaching, 11*(1). Retrieved from www.ied.edu.hk/apfslt/

Scharmann, L.C. (1988a). Locus of control: A discriminator of the ability to foster an understanding of the nature of science among preservice elementary teachers. *Science Education, 72*(4), 453–465.

Scharmann, L.C. (1988b). The influence of sequenced instructional strategy and locus of control on preservice elementary teachers' understandings of the nature of science. *Journal of Research in Science Teaching, 25*(7), 589–604.

Scharmann, L.C. (1989). Developmental influences of science process skill instruction. *Journal of Research in Science Teaching, 26*(8), 715–726.

Scharmann, L.C. (1990). Enhancing the understanding of the premises of evolutionary theory: The influence of diversified instructional strategy. *School Science and Mathematics, 90*(2), 91–100.

Scharmann, L.C. (1994). Teaching evolution: The influence of peer teachers' instructional modeling. *Journal of Science Teacher Education, 5*(2), 66–76.

Scharmann, L.C., & Harris, W.M., Jr. (1992). Teaching evolution: Understanding and applying the nature of science. *Journal of Research in Science Teaching, 29*(4), 375–388.

Schwartz, R.S., & Lederman, N.G. (2002). "It's the nature of the beast": The influence of knowledge and intentions on learning and teaching nature of science. *Journal of Research in Science Teaching, 39*(3), 205–236.

Schwartz, R.S., Lederman, N.G., & Crawford, B.A. (2004). Developing views of nature of science in an authentic context: An explicit approach to bridging the gap between nature of science and scientific inquiry. *Science Education, 88,* 610–645.

Schwartz, R.S., Westerlund, J.F., Garcìa, D.M., & Taylor, T.A. (2010). The impact of full immersion scientific research experiences on teachers' views of the nature of science. *Electronic Journal of Science Education, 14*(1). Retrieved from http://ejse.southwestern.edu/article/view/7325

Schwirian, P.A. (1968). On measuring attitudes toward science. *Science Education, 52*(2), 172–179.

Schwirian, P.A. (1969). Characteristics of elementary teachers related to attitudes toward science. *Journal of Research in Science Teaching, 6*(3), 203–213.

Schwirian, P.A., & Thomson, B. (1972). Changing attitudes toward science: Undergraduates in 1967 and 1971. *Journal of Research in Science Teaching, 9*(3), 253–259.

Scientific Literacy Research Center. (1967). *Wisconsin inventory of science processes.* Madison: University of Wisconsin.

Seker, H., & Welsh, L.C. (2006). The use of history of mechanics in teaching motion and force units. *Science & Education, 15,* 55–89.

Shim, M.K., Young, B.J., & Paolucci, J. (2010). Elementary teachers' views on the nature of scientific knowledge: A comparison of inservice and preservice teachers approach. *Electronic Journal of Science Education, 14*(1). Retrieved from http://ejse.southwestern.edu/article/view/7335/5619

Simpson, R.D., Shrum, J.W., & Rentz, R. (1972). The science support scale: Its appropriateness with high school students. *Journal of Research in Science Teaching, 9*(2), 123–126.

Simpson, R.D., & Wasik, J.L. (1978). Correlation of selected affective behaviors with cognitive performance in a biology course for elementary teachers. *Journal of Research in Science Teaching, 15*(1), 65–71.

Smith, M.U., Lederman, N.G., Bell, R.L., McComas, W.F., & Clough, M.P. (1997). How great is the disagreement about the nature of science: A response to Alters. *Journal of Research in Science Teaching, 34*(10), 1101–1103.

Spears, J., & Zollman, D. (1977). The influence of structured versus unstructured laboratory on students' understanding the process of science. *Journal of Research in Science Teaching, 14*(1), 33–38.

Spears, J.D., & Hathaway, C.E. (1975). Student attitudes toward science and society. *American Journal of Physics, 43*(4), 343–348.

Starr, R.J. (1972). A study of invitations to enquiry and their effect on student knowledge of science processes. *School Science and Mathematics, 72*(8), 714–717.

Stice, G. (1958). *Facts about science test.* Princeton, NJ: Educational Testing Service.

Sutherland, D., & Dennick, R. (2002). Exploring culture, language and the perception of the nature of science. *International Journal of Science Education, 24*(1), 1–25.

Swan, M.D. (1966). Science achievement as it relates to science curricula and programs at the sixth grade level in Montana public schools. *Journal of Research in Science Teaching, 4,* 112–123.

Symington, D.J., & Fensham, P.J. (1976). Elementary school teachers' closed-mindedness, attitudes toward science, and congruence with a new curriculum. *Journal of Research in Science Teaching, 13*(5), 441–447.

Tamir, P. (1972). Understanding the process of science by students exposed to different science curricula in Israel. *Journal of Research in Science Teaching, 9*(3), 239–245.

Tamir, P. (1994). Israeli students' conceptions of science and views about the scientific enterprise. *Research in Science & Technological Education, 12*(2), 99–116.

Tamir, P., & Jungwirth, E. (1975). Students' growth as a result of studying BSCS biology for several years. *Journal of Research in Science Teaching, 12*(3), 263–279.

Tasar, M.F. (2006). Probing preservice teachers' understandings of scientific knowledge by using a vignette in conjunction with a paper and pencil test. *Eurasia Journal of Mathematics, Science and Technology Education, 2*(1), 53–70.

Thye, T.L., & Kwen, B.H. (2004). Assessing the nature of science views of Singaporean pre-service teachers. *Australian Journal of Teacher Education, 29*(2). Retrieved from http://ro.ecu.edu.au/ajte/vol29/iss2/1/

Trembath, R.J. (1972). The structure of science. *The Australian Science Teachers Journal, 18*(2), 59–63.

Trent, J. (1965). The attainment of the concept "understanding science" using contrasting physics courses. *Journal of Research in Science Teaching, 3,* 224–229.

Tsai, C.C. (2007). Teachers' scientific epistemological views: The coherence with instruction and students' views. *Science Education, 91*(2), 222–243.

Tsai, C.C., & Liang, J.C. (2009). The development of science activities via on-line peer assessment: The role of scientific epistemological views. *Instructional Science, 37*(3), 293–310.

Tsai, C.C., & Liu, S.Y. (2005). Developing a multi-dimensional instrument for assessing students' epistemological views toward science. *International Journal of Science Education, 27*(13), 1621–1638.

Turgut, H. (2011). The context of demarcation in nature of science teaching: The case of astrology. *Science & Education, 20,* 491–515.

Urhahne, D., Kremer, K., & Mayer, J. (2011). Conceptions of the nature of science—are they general or context specific? *International Journal of Science and Mathematics Education, 9,* 707–730.

van Fraassen, B.C. (1980). *The scientific image.* Oxford, UK: Oxford University Press.

van Fraassen, B.C. (1998). Arguments concerning scientific realism. In M. Curd & J.A. Cover (Eds.), *Philosophy of science: The central issues* (pp. 1064–1087). New York: Norton.

Vesterinen, V.-M., & Aksela, M. (2012). Design of chemistry teacher education course on nature of science. *Science & Education*, 22: 2193–2225. doi:10.1007/s11191-012-9506-0

Voss, B.E. (1965). "Great Experiments in Biology"—A summer program for academically talented high school students. *The American Biology Teacher, 27*(4), 257–259.

Wahbeh, N., & Abd-El-Khalick, F. (2013). Revisiting the translation of nature of science understandings into instructional practice: Teachers' nature of science pedagogical context knowledge. *International Journal of Science Education, 36*(3): 425–466. doi:10.1080/09500693.2013.786852

Walberg, H.J. (1969). Physics, femininity, and creativity. *Developmental Psychology, 1*(2), 47–54.

Walczak, M.M., & Walczak, D.E. (2009). Do student attitudes toward science change during a general education chemistry course? *Chemical Education Research, 86*(8), 985–991.

Walker, K.A., & Zeidler, D.L. (2007). Promoting discourse about socio-scientific issues through scaffolded inquiry. *International Journal of Science Education, 29*(11), 1387–1410.

Walls, L. (2012). Third grade African American students' views of the nature of science. *Journal of Research in Science Teaching, 49*(1), 1–37.

Welch, W.W. (1966a). *The development of an instrument for inventorying knowledge of the processes of science.* Unpublished doctoral dissertation, University of Wisconsin, Madison.

Welch, W.W. (1966b). *Welch science process inventory (form D).* Madison, WI: Author.

Welch, W.W. (1969). Correlates of courses satisfaction in high school physics. *Journal of Research in Science Teaching, 6,* 54–58.

Welch, W.W. (1972). Evaluation of the PSNS course. I: Design and implementation. *Journal of Research in Science Teaching, 9*(2), 139–145.

Welch, W.W. (1980). A possible explanation for declining test scores or learning less science but enjoying it more. *School Science and Mathematics, 80*(1), 22–28.

Welch, W.W., & Lawrenz, F. (1982). Characteristics of male and female science teachers. *Journal of Research in Science Teaching, 19*(7), 587–594.

Welch, W.W., & Pella, M.O. (1967). The development of an instrument for inventorying knowledge of the processes of science. *Journal of Research on Science Teaching, 5,* 64–68.

Welch, W.W., & Rothman, A.I. (1968). The success of recruited students in a new physics course. *Science Education, 52*(3), 270–273.

Welch, W.W., & Walberg, H.J. (1967–1968). An evaluation of summer institute programs for physics teachers. *Journal of Research in Science Teaching, 5,* 105–109.

Welch, W.W., & Walberg, H.J. (1970). Pretest and sensitization effects in curriculum evaluation. *American Educational Research Journal, 7*(4), 605–614.

Welch, W.W., & Walberg, H.J. (1972). A national experiment in curriculum evaluation. *American Educational Research Journal, 9*(3), 373–383.

Wilson, L.L. (1954). A study of opinions related to the nature of science and its purpose in society. *Science Education, 38*(2), 159–164.

Windschitl, M. (2004). Folk theories of "inquiry": How preservice teachers reproduce the discourse and practices of an theoretical scientific method. *Journal of Research in Science Teaching, 41*(5), 481–512.

Wong, S.L., & Hodson, D. (2009). From the horse's mouth: What scientists say about scientific investigation and scientific knowledge. *Science Education, 93,* 109–130.

Wong, S.L., & Hodson, D. (2010). More from the horse's mouth: What scientists say about science as a social practice. *International Journal of Science Education, 32*(11), 1431–1463.

Wong, S.L., Hodson, D., Kwan, J., & Yung, B.H.W. (2008). Turning crisis into opportunity: Enhancing student-teachers' understanding of nature of science and scientific inquiry through a case study of the scientific research in severe acute respiratory syndrome. *International Journal of Science Education, 30*(11), 1417–1439.

Wood, R.L. (1972). University education students' understanding of the nature and processes of science. *School Science and Mathematics, 72*(1), 73–79.

Yacoubian, H.A., & BouJaoude, S. (2010). The effect of reflective discussions following inquiry-based laboratory activities on students' views of nature of science. *Journal of Research in Science Teaching, 47*(10), 1229–1252.

Yager, R.E. (1968). Critical thinking and reference materials in the physical science classroom. *School Science and Mathematics, 68*(8), 743–746.

Yager, R.E., & Wick, J.W. (1966). Three emphases in teaching biology—a statistical comparison of results. *Journal of Research in Science Teaching, 4,* 16–20.

Yalçinoğlu, P., & Anagün, S.S. (2012). Teaching nature of science by explicit approach to the preservice elementary science teachers. *Elementary Education Online, 11*(1), 118–136.

Yalvac, B., Tekkaya, C., Cakiroglu, J., & Kahyaoglu, E. (2007). Turkish pre-service science teachers' views on science–technology–society issues. *International Journal of Science Education, 29*(3), 331–348.

Zeidler, D.L., & Lederman, N.G. (1989). The effect of teachers' language on students' conceptions of the nature of science. *Journal of Research in Science Teaching, 26*(9), 771–783.

Zimmermann, E., & Gilbert, J.K. (1998). Contradictory views of the nature of science held by a Brazilian secondary school physics teacher: Educational value of interviews. *Educational Research and Evaluation: An International Journal on Theory and Practice, 4*(3), 213–234.

Zingaro, J.S., & Collette, A.T. (1967–1968). A statistical comparison between inductive and traditional laboratories in college physical science. *Journal of Research in Science Teaching, 5,* 269–275.

Zoller, U., & Ben-Chaim, D. (1994). Views of prospective teachers versus practising teachers about science, technology and society issues. *Research in Science & Technological Education, 12*(1), 77–89.

Zoller, U., Donn, S., Wild, R., & Beckett, P. (1991a). Students' versus their teachers' beliefs and positions on science/technology/society-oriented issues. *International Journal of Science Education, 13*(1), 25–36.

Zoller, U., Donn, S., Wild, R., & Beckett, P. (1991b). Teachers' beliefs about views on selected science-technology-society topics: A probe into STS literacy versus indoctrination. *Science Education, 75*(5), 541–561.

Zoller, U., Ebenezer, J., Morley, K., Paras, S., Sandberg, V., West, C., Wolthers, T., & Tan, S.H. (1990). Goal attainment in science-technology-society (S/T/S) education and reality: The case of British Columbia. *Science Education, 74*(1), 19–36.