Chapter 18
Quality of Instruction in Science Education

Knut Neumann, Alexander Kauertz, and Hans E. Fischer

International large-scale assessments revealed remarkable differences in students' science achievements between countries. In the 1995 iteration of the Third International Mathematics and Science Study (TIMSS), students' achievements were less than expected for countries as developed as the United States, Germany, and France (Beaton et al. 1997). These results were confirmed by the Programme for International Student Assessment (PISA) studies (e.g., Organisation for Economic and Cultural Development (OECD 2001). In consequence, a discussion arose in major western countries about the quality of education in general and the quality of instruction in particular.

Attempts to identify and describe quality of instruction and its components were undertaken already in the 1960s. These attempts were followed by extensive research programs on teacher effectiveness in the late 1960s and 1970s. Systemization of results from research on teacher effectiveness on the basis of quality of instruction models led to another boom in research in the late 1970s and 1980s – mainly comprising metaanalyses. Since these efforts were not satisfying with respect to explaining instructional outcomes in general, with the TIMSS study, a new attempt was

K. Neumann (✉)
Department of Physics Education,
Leibniz Institute for Science Education, 24116 Kiel, Germany
e-mail: neumann@ipn.uni-kiel.de

A. Kauertz
Department of Physics, University of Education at Weingarten,
88250 Weingarten, Germany
e-mail: kauertz@ph-weingarten.de

H.E. Fischer
Faculty of Physics, University of Duisburg-Essen,
45127 Essen, Germany
e-mail: hans.fischer@uni-due.de

B.J. Fraser et al. (eds.), *Second International Handbook of Science Education*,
Springer International Handbooks of Education 24, DOI 10.1007/978-1-4020-9041-7_18,

made to investigate instruction and to relate instructional characteristics to students' achievement. This was mainly because video analysis of lessons became technically possible. Video analyses allowed to record classrooms and analyze instruction in an extensive and thorough manner in multiple iterations.

This chapter presents a review of research on quality of instruction in science education including different general theoretical frameworks. Firstly, early attempts in modeling quality of instruction will be described. Based on these models, the extensive amount of studies on teacher effectiveness research will be summarized by the help of metaanalyses and research reviews. Furthermore, recent video-based studies and their results will be described. From the discussed works, finally, dimensions of quality of science instruction will be derived.

Models of School Learning

A first consideration of instructional quality can be found in John Carroll's (1963) model of school learning. In this model, students' degree of learning is described as the ratio of the time a student actually spends on learning and the time a student needs to spend on something in order to learn it. Carroll (1963) defined the time actually spent for learning as a function of opportunity and perseverance, and the time needed as a function of aptitude, ability to understand instruction, and quality of instruction. As to quality of instruction, he suggested a constituting set of characteristics – namely clarity of the learning goals, adequate presentation of the learning material as well as a planned series of learning steps (cf. Carroll 1989).

In the light of research on learning processes by Robert Gagné (1965), Benjamin Bloom (1976) takes a shift away from the relevance of time as such and towards the learning process itself. While he emphasizes the importance of students' prerequisites, in particular their cognitive abilities, for the learning process, he also identifies a set of characteristics influencing the learning process: According to him, cues and feedback have a moderate influence on achievement gains, while reinforcement and participation have a small influence only. However, the overall influence of quality of instruction as well as of students' affective characteristics on student achievement is considered to be only moderate while students' cognitive abilities are considered to have the highest influence (cf. Bloom 1976).

Two other works, by Robert Slavin (1987) and Bert Creemers (1994), set off to systematize existing results from research on instruction on the grounds of Carroll's (1963) model. Creemers (1994) described quality of instruction as the quality of curriculum and its implementation in instruction, grouping procedures as well as characteristics of teachers' behavior. Essential characteristics of teacher behavior are the structuring of content, clarity of presentation, questioning, immediate exercise after presentation, evaluating whether goals are achieved, and corrective instruction (van der Werf et al. 2000). Slavin (1987) reduced Carroll's (1963) model to four elements: quality of instruction, learning time, appropriate levels of instruction, and incentive. Whereas all four elements were considered equally important for effective

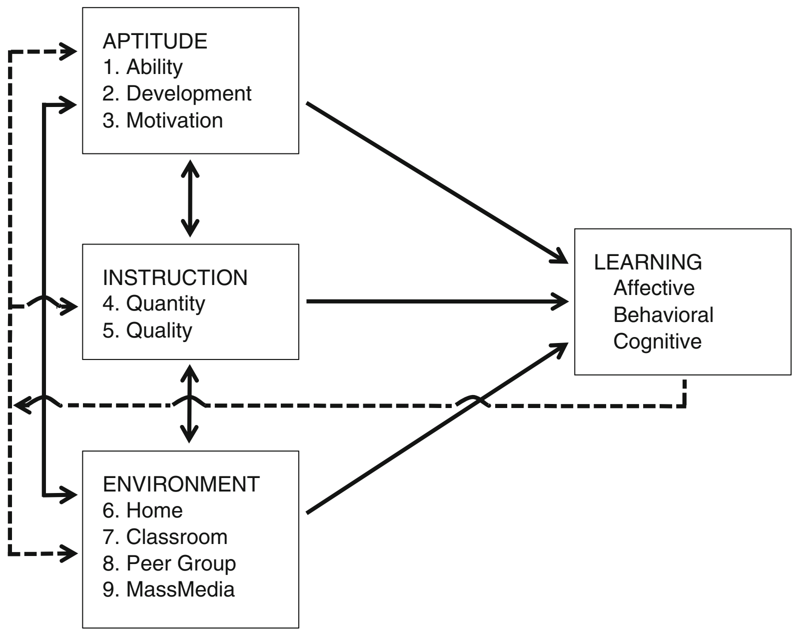

Fig. 18.1 Walberg's (1981) model of educational productivity. Adapted from Fraser et al. (1987, p. 157)

instruction, none of them can be compensated by one of the others. As to quality of instruction, Slavin (1987) compiles a list of characteristics similar to Creemers' (1994) list of teaching or teachers' characteristics, respectively (cf. Gruehn 2000).

Another model that has evolved from Carroll's (1963) model of school learning is the model of educational productivity proposed by Herbert Walberg (1981). Walberg (1981) presented a first systematization of research on modeling school learning and the products of school learning (Gruehn 2000). A major new feature in Walberg's (1981) model was the provision of the learning environment and its influence on students' learning time. Altogether, Walberg (1981) identifies at first seven and in later works nine factors that influence affective, behavioral and cognitive learning: ability or prior achievement, age and development, motivation or self-concept, quantity of instruction or time engaged in learning, quality of instruction, home environment, classroom environment, peer group environment, and the mass media (Fig. 18.1; cf. Fraser et al. 1987). Quality of instruction in this model is related to the degree of direct instruction (Rosenshine 1979).

Summarizing, it has to be maintained that within the above models instruction is described as a function of student individual characteristics, instructional characteristics, and characteristics of the learning environment providing information on the quality of the learning process and in consequence of instructional outcomes. Quality of instruction is considered a set of instructional characteristics, as for example, clarity and structure or teacher–student interactions. Outcomes can be affective, behavioral or cognitive, where the focus is mostly on the latter, that is, students' achievement. Walberg's (1981) model takes an exceptional position in scope of the discussed models. It is a synthesis of all proceeding models at least with respect to the first five factors it embraces, while it accounts for the learning environment through inclusion of the remaining four factors (Gruehn 2000). Finally, it describes quality of instruction on the basis of empirical research on teaching effectiveness. The models discussed so far are proposed for instruction and learning in general. Specific characteristics of individuals and environments are taken into account but domain specifics, that is, subject matter or subject specific learning processes, remain unconsidered.

Teacher Effectiveness Research

Early research on teacher effectiveness followed two different research approaches: The teaching process paradigm on the one hand and the criterion of effectiveness paradigm on the other (Gage 1972). Within the teaching process paradigm, what characterizes a good teacher was defined based on experts' experience or observations of classroom learning (Rosenshine and Furst 1971). The criterion of effectiveness approach on the other hand drew on outcome criteria, for example, student achievement, for identifying characteristics of effective teaching (Shavelson and Dempsey-Atwood 1976). A first major review of research on the latter is given by Barak Rosenshine and Norma Furst (1971). They derive a set of 11 different variables, amongst which Clarity, Variability, Enthusiasm, Task-oriented and/or Businesslike Behaviors, and Students' Opportunity to Learn Criterion Material are considered as particularly important. However, Rosenshine and Furst (1971) state a lack of substantial research on teachers' characteristics relating to higher student achievement and demand further research in this field to back up the relevance of the characteristics compiled by them.

In another attempt to summarize the general factors that influence classroom learning, Michael Dunkin and Bruce Biddle (1974) developed the so-called "process-product model" of classroom learning. The model embraces four classes of variables: teacher characteristics (e.g., personality), context variables (e.g., classroom environment), process variables (e.g., learning activities), and product variables (e.g., student achievement) (cf. Shuell 1996).

In the decade following Dunkin and Biddle's (1974) work, the research base has been considerably broadened. The 1970s and 1980s provided a substantial amount of correlational and experimental studies that documented causal relationships between teacher behaviors and student achievement. In reference to the model suggested by Dunkin and Biddle (1974), this research is termed *process-product research*. Studies provided evidence that classroom management influences student achievement (Good 1979). Other studies indicated that managing classrooms effectively begins on the first day of school with a systematic approach, advance preparation, and planning (Evertson 1985). With reference to the core idea of Carroll's (1963) model of school teaching and learning, much research focused on the investigation of time-on-task. Results documented the importance of time-on-task, pointing out that students must become actively engaged in learning during instruction time (Anderson 1981).

In a review of several metaanalyses, Ronald Anderson (1983) summarizes the results of research on teacher effectiveness specific to science education. His analysis confirms the superiority of an inquiry approach in, for example, curricula or teaching techniques, although effect sizes vary heavily between metaanalyses. Additionally, effects with respect to the teaching of process skills were found. Interestingly, effects were noticeably larger in studies testing students for specific techniques but small in those testing for scientific methods in general.

An all-embracing review of process-product research was written by Jere Brophy and Thomas Good (1986) identifying two dimensions of characteristics: characteristics related to quantity and pacing of instruction on the one hand and qualitative characteristics on the other. As to quantitative characteristics, they find the amount of opportunities to learn and the content covered, role definition/expectations/time allocation, classroom management/student engaged time, consistent success/academic learning time, and active teaching to have a positive impact on instructional outcomes. With respect to qualitative characteristics, giving information (including structuring, redundancy/sequencing, clarity, enthusiasm and pacing/waiting time), questioning the students (including difficulty level of questions, cognitive level of questions, clarity of questions, selecting the respondent, waiting for the student to respond), as well as reacting to students' responses (including, for example, reactions to correct and incorrect responses), handling seatwork, and homework are identified (cf. Brophy 1986). These results, although formulated in a different way, strongly support the characteristics of effective teaching found by Rosenshine and Furst (1971). The aspect of clarity can be found in both reviews; variability in Rosenshine and Furst's (1971) review relates to the cognitive level in discourse and, thus, is included in questioning students – as is enthusiasm. Task/business-like behaviors refer to characteristics subsumed under quantitative characteristics. In addition, Brophy and Good (1986) emphasize the importance of structuredness of content as suggested by David Ausubel (1968), Jerome Bruner (1966) and other cognitive structuralists.

Particularly interesting is the work of Barry Fraser et al. (1987) as it presents a synthesis of educational research. Based on Walberg's (1981) model of educational productivity, research reviews of the 1970s were analyzed, from which productive factors of learning were obtained. In addition, quantitative syntheses or metaanalyses of studies of these factors were accomplished. Fraser et al. (1987) found that three groups of aptitudinal, instructional, and environmental factors have influences on instructional outcomes, that is, cognitive, affective, and behavioral learning. The strongest effects were found for variables of students' aptitude, wherein intelligence was found to be the strongest factor. As to quality of instruction, Fraser et al. (1987) found a mean effect size for time and strong effects for reinforcement, instructional cues, engagement, and feedback. The works of Fraser et al. (1987) are remarkable in another way as well, as the authors derive a model to describe contextual and transactional influences on science outcomes, which after the work of Anderson (1983) is a particular attempt in describing a model of instructional quality specifically for science education. Fraser et al. (1987) found the strongest factor of quality of instruction to be the time between a teachers' question and students' answers, followed by focusing (e.g., organizers), students' hands-on activities, use of teacher questioning or – in line with Anderson (1983) – inquiry learning. The overall mean effect size of the factors established was one-third of a standard deviation (Fraser et al. 1987).

A further probe of the model of educational productivity is accomplished by Herbert Walberg et al. (1981) using data from the National Assessment of Educational Progress (NAEP) program. By regression analysis, the factors Socioeconomic Status, Motivation, Quality of Instruction (measured by a questionnaire on students' perception of the degree of direct, didactic instruction), Class (social psychological environment), and Home conditions were each found to be significant. While other factors such as race and gender where controlled, "Under a stringent probe, however, the Class social-*psychological* environment appears as the only unequivocal cause of science learning in the data" (Walberg et al. 1981, p. 233). These results are confirmed by Margaret Wang et al. (1990), who find classroom management and climate together with student-teacher interactions to form an important set of instructional characteristics related to effective instruction.

Altogether, from research on teacher effectiveness, five dimensions of variables may be identified: clarity, structuredness, cognitive activation, pacing, and classroom management. Clarity refers to the clarity of learning goals, the presented content and so on, and structuredness refers to a systematic approach in the design of instruction. Cognitive activation embraces all variables relevant to activate students cognitively, for example, the cognitive level of tasks as well as variables related to students' engagement. Pacing is related to the adequate sequencing of tasks, in which adequateness means adequate with respect to students' abilities rather than an adequate content structure. Finally, classroom management refers to an adequate learning climate that allows for an effective learning. An important characteristic, which is not part of the above dimensions, would be teacher enthusiasm. This characteristic is not considered part of the actual instruction but rather is part of a whole set of characteristics related to a teachers' traits. These characteristics certainly will have to be included in a model of quality of instruction as they influence design and implementation of instruction (Wayne and Youngs 2003).

Video Studies of Instruction

Quality of instruction research received a major revival with the so called TIMSS Video Study (Stigler et al. 1999). As video recording and analysis became technically possible, this offered a new approach to the analysis of instruction. Video analysis preserves classroom activity so it can be viewed several times allowing for a detailed examination of the complex actions taking place in classrooms. In scope of the TIMS Video Study, this method was used to analyze mathematics lessons from Germany, Japan, and the United States to identify instructional characteristics relevant for differences observed in students' achievements in the TIMS study (Beaton et al. 1997). Analysis covered the content of the lessons, the teachers' aims as well as teachers' and students' manuals, verbal activities, and the material used. The analysis revealed the existence of specific patterns of instruction in Germany, the United States, and Japan – so-called lesson scripts (Stigler and Hiebert 1997). While instruction in Japan is characterized by a rather constructivist approach, instruction in Germany was identified as narrowly guided and result-oriented. Lesson scripts were considered to be highly culture specific (Stigler and Hiebert 1997). Despite that, no explanation for performance differences between the participating countries could be found (Stigler et al. 1999).

Thus, an aim of a further video study in scope of the 1999 iteration of TIMSS was to investigate whether high achieving countries share a common method of teaching (Hiebert et al. 2003). This time, science instruction was also video recorded and analyzed. In mathematics, lessons were videotaped in Australia, the Czech Republic, Hong Kong, the Netherlands, Switzerland, and the United States. Additionally, Japanese lessons from the earlier study were reanalyzed. Results of the preceding video study could be confirmed in general. Again, lesson structures similar to the ones found in the scope of the TIMSS Video Study could be observed. Differences appear, however, when investigating the characteristics of tasks. While in most countries the majority of problems presented during instruction were of low complexity, in Japan about 40% of the problems used were of high complexity. Also, in Japan in over 40% of the tasks, a previous task's solution was used to solve the given task, whereas at least 65% of the tasks in other countries were repetitive, that is, a task was the same or mostly the same as the preceding one (Hiebert et al. 2003). Yet, as the majority of Japanese mathematics lessons dealt with geometry and was videotaped 4 years earlier, the interpretation is not very powerful.

Results of the science part of the study were published in 2006 by Kathleen Roth et al. (2006). Based on an extensive literature review of research on teacher effectiveness, criteria of instructional quality were compiled and categorized in three classes: science content, teacher actions, and student actions embedded in school culture. Analyzing science instruction in Australia, the Czech Republic, Japan, The Netherlands, and United States on the grounds of this framework, Roth et al. (2006) found that high achieving countries shared two common characteristics: high content standards and a content-focused instructional approach. However, these high content standards were embodied by different characteristics per country, as, for example, the density and challenge of content ideas or students being held responsible for their own independent learning.

In summary, while the TIMSS video studies provided an extensive description of mathematics and science instruction, they failed in relating instructional characteristics to student achievements. This lack of reliable findings on the influence of country-specific patterns of instruction on students' performance led to a series of research projects investigating instruction by means of video analysis.

In an effort to shed more light on the complex matter of science instruction, a video study was undertaken by the Institut für die Pädagogik der Naturwissenschaften (IPN) in Kiel, Germany. The scope of this video study of physics instruction was to investigate teaching and learning processes (Seidel et al. 2007). Based on the results of research on teacher and teaching effectiveness, taking the "complex mediating process from instructional activities to student learning" (Seidel et al. 2005, p. 552) into account, a theoretical framework was used as a basis of a multitrait multimethod approach to examine physics instruction. Classroom activity patterns were investigated, aspects of instructional quality were surveyed, and finally these findings were related to student reports on cognitive learning processes, quality of learning motivation, and perception of supportive learning conditions (Seidel et al. 2005). Results on physics instruction were in line with the findings from the TIMS video study on mathematics instruction: German physics instruction is characterized by a narrowly focused questioning–developing teaching style. This was confirmed by Thomas Reyer (2004) who found that physics instruction is mainly characterized by a teacher-centered instruction using demonstration experiments and seldomly by student-centered instruction using experimental group work. However, Tina Seidel et al. (2007) could not find an influence of either approach on student learning. A more in-depth analysis, though, provided empirical evidence for several assumptions on quality of instruction: Goal clarity and coherence have a positive influence on students' perceptions of supportive learning conditions. Interactions in class work were found to be related to motivational affective development (cf. Seidel et al. 2005). Further, students perceived themselves as being more self-determined and motivated in classrooms with high quality classroom discourse (Seidel et al. 2003), that is, with high cognitive activation. Analysis of the use of experiments pointed toward a lack of support and self-contained learning during experimental phases (Tesch and Duit 2004).

Similar results could be found in a Swiss-German cooperation project "Instructional Quality and Mathematical Understanding in Different Cultures" (Rakoczy et al. 2007). Based on an opportunity-to-learn model of instructional quality (Fig. 18.2), a three-lesson unit was videotaped in 20 German and 20 Swiss classes. Analysis was based on three dimensions of teaching quality: classroom management, cognitive activation, and student-centered orientation (Lipowsky et al. 2005) as well as structure of the content presented (Rakoczy et al. 2007). Results provided evidence that student achievement is higher in classes with high cognitive activation. Also, classroom discourse was found to have an influence on student achievement. Together, both characteristics explained 9% of students' achievement (Lipowsky et al. 2005). Additionally, a structured presentation of content was found to have a particular influence on student achievement (Rakoczy et al. 2007).

In another approach, the data were analyzed with respect to instructional patterns (Hugener et al. 2007). Altogether, three patterns with respect on how the solution to problems posed during instruction is handled could be identified: a presenting pattern, a development pattern, and a discovery pattern. In line with the results of the TIMS Video Study described above, the discovery pattern was related to the highest

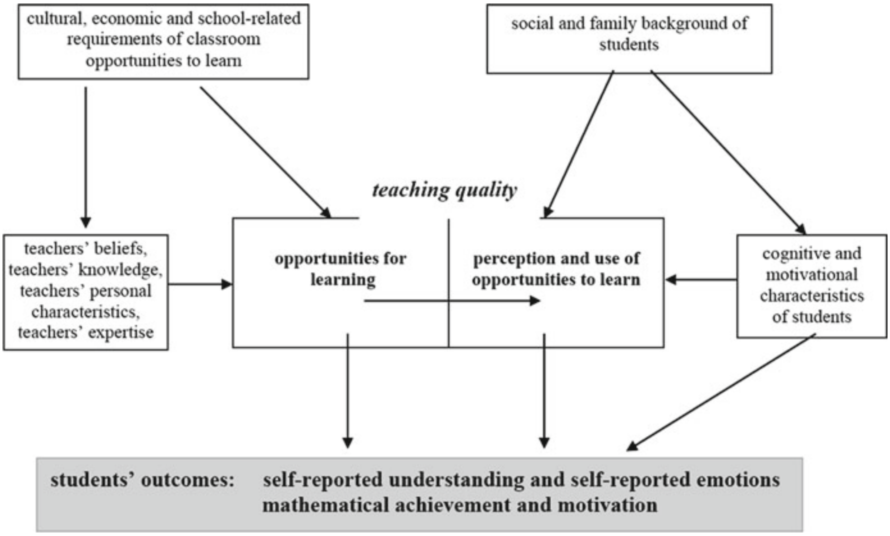

Fig. 18.2 Model of instructional quality. Taken from Lipowsky et al. (2005)

cognitive activation although again no influence on student achievement could be observed. This allows for the conclusion that while instruction might look the same on a surface level of instruction, instructional characteristics influencing students' achievement might be located on a deeper level.

Apart from the presented video studies investigating instruction as a whole, a lot of studies have taken into focus different aspects of instruction on a descriptive base or correlational base with respect to student outcome. Eduardo Mortimer and Phil Scott (2003), for example, focus on a description of classroom or student–teacher interaction, respectively, particularly on dialog structures in the classroom. Others investigate the teachers' role in supporting learning in different teaching-learning environments (e.g., Viiri and Saari 2004). However, it is too early, yet, to draw conclusions as more studies will be needed to confirm the findings and allow for metaanalyses to create a larger picture of how these characteristics relate to each other and how they contribute to quality of instruction in general.

In summary, earlier video studies of instruction were not able to establish a relation between characteristics of instruction and students' achievement, whereas later ones were more successful as they set a stronger focus on deep-level characteristics of instruction and were based on more elaborate models of instructional quality. Results of the later investigations show that clarity, classroom management, cognitive activation, and structuredness have an impact on outcome criteria. This confirms the dimensions that could be identified from teacher effectiveness research. And while these dimensions are not specific to science education, their relevance to science education can be concluded from the described studies.

Summary and Outlook

Early models of school learning describe quality of instruction as a set of instructional characteristics influencing the learning process and thus mediating the influence of students' prerequisites on students' outcomes. In later models, the extensive amount of research on teacher effectiveness is systematized leading to five dimensions of instructional quality: clarity, structuredness, cognitive activation, pacing, and classroom management. The rapidly developing video recording technology allowed for a large-scale use of video equipment to record and to analyze lessons. And while early video studies struggled to identify instructional characteristics, later ones were – on the basis of theoretically founded models of instructional quality – able to provide evidence on the importance of the above dimensions. However, more research is needed especially on science-specific aspects of instructional quality. That is, on science-specific operationalizations and the interplay of the above dimension as well as the relevance of science-specific instructional characteristics, that is, the use of experiments.

Moreover, further research should take characteristics of students, teachers, and the classroom environment and their influence on the above dimensions of instructional quality into account. This is especially important as there is evidence that a mere change of instructional patterns does not influence student outcome and that quality of instruction is to be sought on the deep level of instruction. This again means that teacher training programs seeking to improve quality of instruction have to focus on teachers' professional knowledge to efficiently change the way instruction is designed and implemented.

Finally, as it seems that aptitudes are powerful correlates of learning; they deserve inclusion in theories of educational productivity.

References

Anderson, L. W. (1981). Instruction and time-on-task: A review. *Journal of Curriculum Studies*, *13*, 289–303.

Anderson, R. D. (1983). A consolidation and appraisal of science meta-analyses. *Journal of Research in Science Teaching*, *20*, 497–509.

Ausubel, D. P. (1968). *Educational psychology: A cognitive view*. New York: Holt, Rinehart and Winston.

Beaton, A. E., Martin, M. O., Mullis, I. V., Gonzalez, E. J., Smith, T. A., & Kelly, D. S. (1997). *Science achievement in the middle school years: IEA's Third International Mathematics and Science Study (TIMSS)*. Chestnut Hill, MA: Center for the Study of Testing, Evaluation, and Educational Policy, Boston College.

Bloom, B. S. (1976). *Human characteristics and school learning*. New York: McGraw-Hill.

Brophy, J. E., & Good, T. L. (1986). Teacher behavior and student achievement. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (pp. 328–375). New York: Macmillan.

Brophy, J. (1986). Teacher influences on student achievement. *American Psychologist*, *41*, 1069–1077.

Bruner, J. S. (1966). *Toward a theory of instruction*. New York: W. W. Norton.

Carroll, J. B. (1963). A model of school learning. *Teachers College Record*, *64*, 723–733.
Carroll, J. B. (1989). The Carroll model: A 25-year retrospective and prospective veiw. *The Educational Researcher*, *18*, 26–31.
Creemers, B. P. (1994). *The effective classroom*. London: Cassell.
Dunkin, M. J., & Biddle, B. J. (1974). *The study of teaching*. New York: Holt, Rinehart and Winston.
Evertson, C. M. (1985). Training teachers in classroom management: An experimental study in secondary school classrooms. *Journal of Educational Research*, *79*, 51–58.
Fraser, B. J., Walberg, H. J., Welch, W. W., & Hattie, J. A. (1987). Synthesis of educational productivity research. *International Journal of Educational Research*, *11*, 145–252.
Gage, N. L. (1972). *Teacher effectiveness and teacher education: The search for a scientific basis.* Palo Alto, CA: Pacific.
Gagné, R. M. (1965). *The conditions of learning*. New York: Holt, Rinehart and Winston.
Good, T. (1979). Teacher effectiveness in elementary school. *Journal of Teacher Education*, *30*, 52–64.
Gruehn, S. (2000). *Unterricht und schulisches Lernen (Instruction and school learning).* Münster, Germany: Waxmann.
Hiebert, J., Gallimore, R., Garnier, H., Bogard Givvin, K., Hollingsworth, H., Jacobs, J., et al. (2003). *Teaching mathematics in seven countires: Results from the TIMSS 1999 video study* (NCES 2003–013 Revised). Washington, DC: U.S. Department of Education, National Center for Education Statistics.
Hugener, I., Pauli, C., & Reusser, K. (2007). Inszenierungsmuster, kognitive Aktivierung und Leistung im Mathematikunterricht (Instructional patterns, cognitive activation and achievement in mathematics instruction). In D. Lemmermöhle, M. Rothgangel, S. Bögeholz, M.Hasselhorn, & R. Watermann (Eds.), *Professionell Lehren – Erfolgreich Lernen* (Professional teaching – Successful learning) (pp. 109–121). Münster, Germany: Waxmann.
Lipowsky, F., Rakoczy, K., Vetter, B., Klieme, E., Reusser, K., & Pauli, C. (2005, April). *Quality of geometry instruction and its impact on the achievement of students with different characteristics*. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada.
Mortimer, E., & Scott, P. (2003). *Meaning making in the secondary science classroom.* Milton Keynes, UK: Open University Press.
Organisation for Economic and Cultural Development (OECD). (2001). *Knowledge and skills for lLfe – First results from PISA 2000*. Paris: OECD.
Rakoczy, K., Klieme, E., Drollinger-Vetter, B., Lipowsky, F., Pauli, C., & Reusser, K. (2007). Structure as a quality feature of instruction. In M. Prenzel (Ed.), *Studies on the educational quality of schools* (pp. 101–120). Münster, Germany: Waxmann.
Reyer, T. (2004). *Oberflächenmerkmale und Tiefenstrukturen im Unterricht* (Surface characteristics and deep level structures of instruction). Berlin: Logos.
Rosenshine, B. (1979). Content, time and direct instruction. In P. L. Peterson & H. Walberg (Eds.), *Research on teaching* (pp. 28–56). Berkeley, CA: McCutchan.
Rosenshine, B., & Furst, N. (1971). Research on teacher performance criteria. In B. O. Smith (Ed.), *Research in teacher education: A symposium* (pp. 37–72). Englewood Cliffs, NJ: Prentice Hall.
Roth, K. J., Druker, S. L., Garnier, H. E., Lemmens, M., Chen, C., Kawanaka, T., et al. (2006). *Teaching science in five countries: Results from the TIMS 1999 video study* (NCES 2006–011). Washington, DC: US Government Printing Office.
Seidel, T., Prenzel, M., Rimmele, R., Herweg, C., Kobarg, M., Schwindt, K., et al. (2007). Science teaching and learning in German physics classrooms. In M. Prenzel (Ed.), *Studies on the educational quality of schools* (pp. 79–99). Münster, Germany: Waxmann.
Seidel, T., Rimmele, R., & Prenzel, M. (2003). Gelegenheitsstrukturen beim Klassengespräch und ihre Bedeutung für die Lernmotivation – Videoanalysen in Kombination mit Schülerselbsteinschätzungen (The structure of opportunities during classroom discourse and

their influence on motivation to learn – Video analyses in combination with self-evaluations). *Unterrichtswissenschaft*, *31*(2), 142–165.
Seidel, T., Rimmele, R., & Prenzel, M. (2005). Clarity and coherence of learning goals as a scaffold for student learning. *Learning and Instruction*, *15*, 539–556.
Shavelson, R., & Dempsey-Atwood, N. (1976). Generalizability of measures of teaching behavior. *Review of Educational Research*, *46*, 553–611.
Shuell, T. J. (1996). Teaching and learning in a classroom context. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 726–764). New York: Macmillan.
Slavin, R. E. (1987). Quality, appropriateness, incentive and time: A model of instructional effectiveness. *International Journal of Educational Research*, *21*, 141–157.
Stigler, J. W., Gonzales, P., Kawanaka, T., Knoll, S., & Serrano, A. (1999). *The TIMSS Videotape Classroom Study: Methods and findings from an exploratory research project on eighth-grade mathematics instruction in Germany, Japan and the United States.* Washington, DC: National Center for Education Statistics.
Stigler, J., & Hiebert, J. (1997). Understanding and improving mathematics instruction: An overview of the TIMSS Video Study. *Phi Detla Kappa*, *79*(1), 14–21.
Tesch, M., & Duit, R. (2004). Experimentieren im Physikunterricht – Ergebnisse einer Videostudie (Experiments in physics instruction – Results from a video study). *Zeitschrift für Didaktik der Naturwissenschaften*, *10*, 51–69.
van der Werf, G., Creeemers, B., de Jong, R., & Klaver, E. (2000). Evaluation of school improvement through an educational effectiveness model: The case of Indonesia's PEQIP Project. *Comparative Education Review*, *44*, 329–355.
Viiri, J., & Saari., H. (2004). Teacher talk in science education. In A. Laine (Ed.), *Proceedings of the 21th annual symposium of the Finnish Association of Math and Science Education Research* (pp. 448–466). Helsinki, Finland: University of Helsinki, Department of Applied Sciences of Education.
Walberg, H. J. (1981). A psychological theory of educational productivity. In F. H. Farley & N. Gordon (Eds.), *Psychology and education: The state of the union* (pp. 81–108). Berkeley, CA: McCutchan.
Walberg, H. J., Haertel, G. D., Pascarella, E., Junker, L. K., & Boulanger, F. D. (1981). Probing a model of educational productivity in science with national assessment samples of early adolescents. *American Educational Research Journal*, *18*, 233–249.
Wang, M. C., Haertel, G. D., & Walberg, H. J. (1990). What influences learning? A content analysis of review literature. *Journal of Educational Research*, *84*, 30–43.
Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research*, *73*, 89–122.