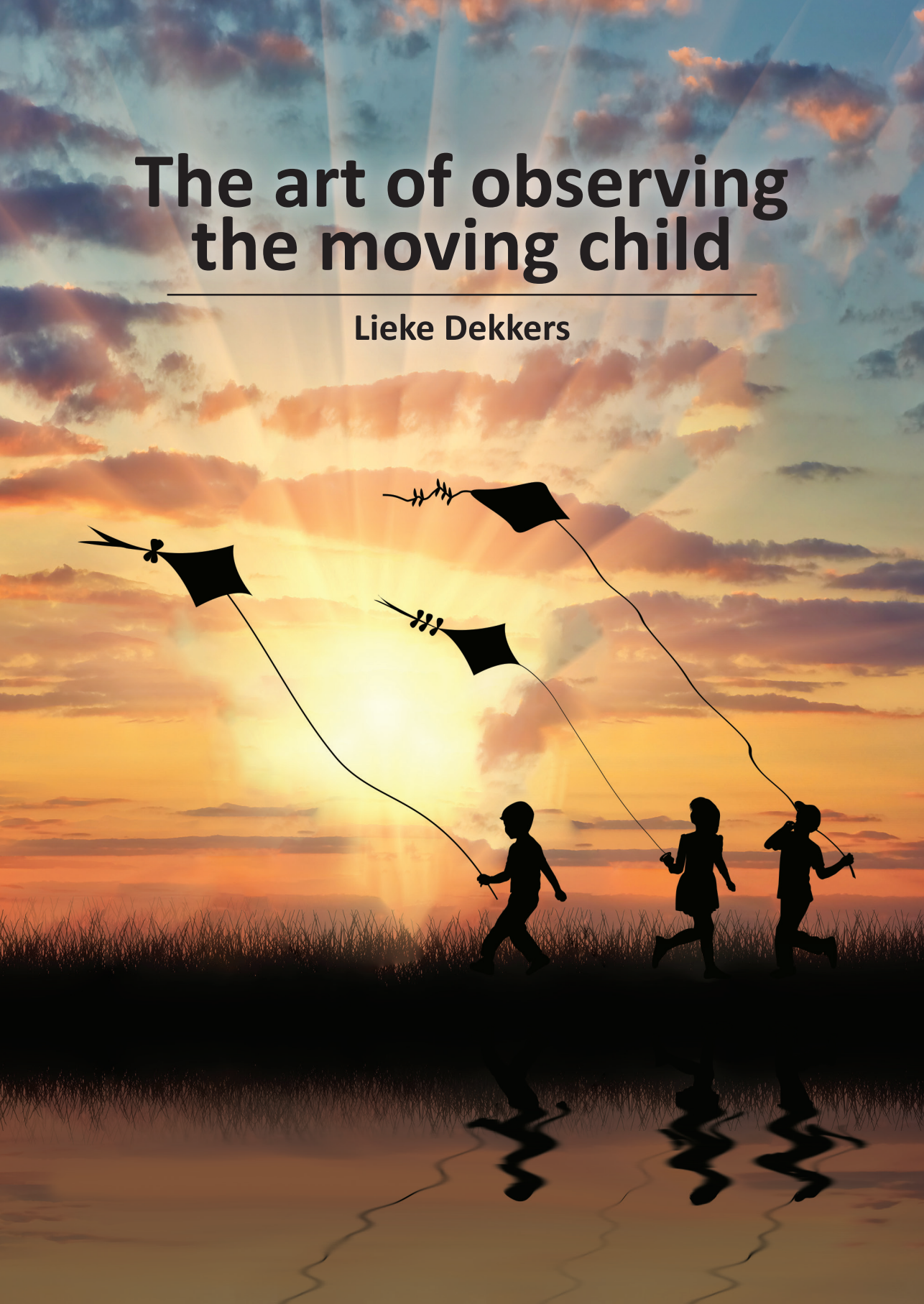
The art of observing the moving child

Lieke Dekkers



The art of observing the moving child

Lieke Dekkers

Lieke Dekkers
The art of observing the moving child
176 pages

The work presented in this thesis was carried out within the department of rehabilitation of the Radboud Institute for Health Sciences, Nijmegen.
Financial support for the printing of this thesis has kindly been provided by Radboudumc, the Scientific College Physiotherapy (WCF) of the Royal Dutch Society for Physiotherapy (KNGF), and the Dutch Association for Pediatric Physiotherapy (NVFK).
This research was supported by the Netherlands Organisation for Scientific Research (NWO) under project number: 023.004.037

ISBN: 978-94-6380-708-1

Layout and Cover: Dennis Hendriks, ProefschriftMaken.nl
Printed: ProefschriftMaken.nl

The art of observing the moving child

Proefschrift

ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,
volgens besluit van het college van decanen
in het openbaar te verdedigen op woensdag 8 april 2020
om 14.30 uur precies

door
Lieke Marie Anne Dekkers
geboren op 11 februari 1972
te Zeeland

Promotoren:	Prof. dr. M.W.G Nijhuis-van der Sanden
	Prof. dr. P.J. van der Wees
Copromotoren:	Dr. A.J.W.M. Janssen
	Dr. B.J.M. de Swart
Manuscriptcommissie:	Prof. dr. B. Steenbergen
	Prof. dr. M.A.A.P. Willemsen
	Prof. dr. M.J. Jongmans (Universiteit Utrecht)

Content

Chapter 1

General Introduction

The art of observing the moving child

Movement is an essential part of our lives. Children are born with the ability to move, but while growing up, they acquire countless motor skills, and those skills become increasingly controlled.[1] Guthrie[2] defined motor skills as 'the ability to bring about some end result with maximum certainty and minimum outlay of energy, or of time and energy' (p. 136). The acquisition of motor skills progresses from basic motor skills – such as walking, reaching and grasping, chewing and talking[3] – to more specific skills based on environmental needs and demands – such as writing a letter, swimming or playing a hockey match. Basic motor skills are mostly learned by trying and discovering, whereas more specific skills are often learned by imitating others or from verbal explanations of how to perform the particular skill. During the acquisition and reacquisition of motor skills, both quantitative and qualitative changes can be observed.[4] Quantitative changes can be seen as the acquisition of new and more complex motor skills,[5] whereas movement quality gives an impression of how movements are executed.[6] The observation of movements is a core aspect of diagnosis by physiotherapists as a basis for interventions to improve functional abilities in children and adults.[7,8] Currently available discriminative motor tests in paediatric physiotherapy specifically assess quantitative aspects by comparing individuals with their peers; these tests are norm referenced and validated. To assess movement quality, however, no generic instruments are available to assess children over time and for all age categories. Therefore, this thesis will focus on measurement of movement quality.

This introductory chapter starts with a general definition of the theoretical concepts for: learning and improving motor skills – that is, motor development, motor learning and motor control – before describing movement quality; observation of movements and their meaning for paediatric physiotherapy; and education on observational skills. Subsequently, this chapter describes the development of a newly developed measurement instrument to assess movement quality in children, the Observable Movement Quality (OMQ)-scale, followed by a description of the validation process for measurement instruments. The introduction ends with the aims and outline of this thesis.

Motor development, motor learning and motor control

In relation to learning and improving motor skills, 3 terms are frequently used in the literature: motor development, motor learning and motor control. Motor development refers to the process of increasing the capability to use the body to move, from basic movement skills to more specific and demanding motor skills. Typical motor development occurs in a predictable sequence of learning skills and is based on context and experience. Motor development is a nonlinear process, which varies for each child in age when a skill is mastered, and is affected by many factors. These factors consist of features of the child – such as body weight, muscle power, presence of diseases or disorders – and components of the environment – such as availability of toys, housing conditions and family composition.[3,9] Furthermore,

genetic factors are important for their role in genetically determined neurodevelopmental processes, which determine possibilities for variation – that is, the presence of a repertoire of options for achieving a goal – and adaptability – that is, the capacity to select the most efficient strategy from the repertoire in a specific situation.[10-12] In motor development, the quality of activity is a fundamental characteristic of the development of neural networks.[13-15] Motor learning focusses on the understanding of the acquisition and/or modification of movements.[16] When a child practises motor skills, the result is almost always an improved performance level, which could be measured in several ways – such as being able to ride a bike, reducing the time to run 500 metres or building a higher tower using blocks. However, improved performance does not define learning by itself; it is merely an indication that learning may have occurred.[17] By observing the systematic changes in motor performance that occur with practice, the gain in the underlying capability for skilled performances is represented; this improved capability leads to improved performance and illustrates the level of motor learning.[17] The type of instruction or feedback provided during practice influences the speed and degree of skill learning.[6] Motor learning has been described as a set of processes associated with practice or experiences leading to a relatively permanent change in the capabilities for skilled movement.[17]

Motor control refers to the planning and execution of movements and is defined as the ability to regulate or direct the mechanisms essential to movement.[18,19] Motor control focusses on understanding the control of a movement already acquired.[19] In motor control research, for example, studies focus on how the central nervous system organises the many individual muscles and joints into coordinated functional movements or how sensory information from the environment and the body is used to select and control movements.[6,18] Although the understanding of why a child moves the way he or she does (motor control) is an important topic in paediatric physiotherapy, the focus of this thesis is on the changes in performances during the acquisition or modification of motor skills (motor development and motor learning) observable in the quality of movements.

Movement quality

Wallbott[20] defined movement quality as 'the way in which human movement is executed with respect to the dimensions of time and space' (p. 345). Movement quality demonstrates subtle characteristics, such as velocity, fluency, accuracy and automation of movements, and gives an impression of how movements are controlled and coordinated.[5,6] In this way, movement quality represents the interaction among personal characteristics and experiences, the task difficulty and the environmental conditions; and it gives insight into the possibilities and potential of the person's movement system for reacting or adapting to changing conditions.[6,7,16,21,22]

For the assessment of movement quality, physiotherapists must rely on observational skills, which involves gathering, organising and giving meaning to visual, auditory and sensory information obtained while observing the moving person.[7,8] Movement quality has universal

characteristics that change under the influence of maturation, development and learning. Although descriptions of movement quality are frequently used, they are not standardised; the descriptions differ among physiotherapists, depending on the theoretical construct used in clinical reasoning, which precludes comparability and longitudinal evaluation.[7,23-25] Several studies have been performed in the field of paediatric physiotherapy to develop appropriate measurement instruments; however, available and commonly used qualitative measurement instruments focus mostly on specific diagnostic groups, such as children with cerebral palsy (Quality of Function Measure),[26] and are specifically designed to assess the functioning of extremities (Quality of Upper Extremity Skills Test)[27,28] or children in a specific age frame (General Movements, Infant Motor Profile).[29,30] To allow for comparison of movement quality between paediatric physiotherapists and longitudinal evaluation alongside discriminative motor tests, movement quality should be assessed independently of a specific age, motor task and predetermined theoretical construct.[25] Although currently available discriminative motor tests specifically assess quantitative aspects by comparison with peers, these tests are norm referenced and validated; no generic instrument is available to assess movement quality over time for all age categories in paediatric physiotherapy.

Observation

Observation is a fundamental skill for physiotherapists.[7] Curricula for bachelor students in physiotherapy address knowledge and skills for observation, measurement and interpretation of the quantity and quality of movements. To observe is, according to the Oxford English Dictionary, 'to watch carefully and attentively, to notice, to perceive and to register it as being significant'.[31] Humans, however, have the tendency to colour observation with an interpretation of what is seen, even though initial observation should be without any judgement.[32] For example, while observing a child jumping into 5 squares with 2 feet together and without stopping, one could describe in an interpretative manner that strength regulation and automated movements are lacking, or one could describe more objectively that the child is not able to jump with both feet together and fails to stop when landing.

General principles for observation were described by Boudreau et al.,[32] who identified several core principles. These principles are that observation has sensory, perceptive and cognitive components; that observation is distinct from inference and is made concrete through description; and that observations are goal oriented, occur over time, carry ethical obligations and occur on different levels. These levels of observation refer to the whole person observed, a body part, the personal or environmental context, behaviours and interactions, and the characteristics of the observer on, for example, emotional and aesthetic planes.[32,33] It seems evident that the observer should be considered an influencing factor when teaching and evaluating observational skills[34,35] because there is a tendency for perception – the interpretation of what is seen – to interfere with observation.[32,36,37]

Education on observational skills

Effective education on practical skills can alter clinical behaviour, positively influence patient outcomes and reduce the risk of patient harm.[38] Observation is one of the most prized and valued clinical skills, but it is a difficult skill to learn, especially for students who must rely on their descriptive abilities to describe their observational findings.[39] Furthermore, disputes are ongoing over the details of what constitutes (good) clinical observation, its conceptual basis and how it is learned or developed.[32] Observational skill teaching can have the tendency to emphasise the identification of memorised clinical signs rather than actually teaching students how to observe and describe.[40] However, the need to consider perception when teaching and evaluating observational skills seems evident.[34-37] Such an explicit education in observational skills could contribute to turning the students looking into a deeper seeing.[41]

The process of learning observational skills shares many of the features of skills learning in general.[32] As early as 1910, Flexner[42] referenced the insights of Cabot and Locke:[43]

Learning medicine is not fundamentally different than learning anything else. If one had one hundred hours in which to learn how to ride a horse or to speak in public, one might profitably spend perhaps an hour (in divided doses) in being told how to do it, four hours in watching a teacher do it, and the remaining ninety-five hours in practice, at first with close supervision, later under general oversight.

Although the importance of repetitive practice was emphasised at that time, detailed feedback on performance is now known to be critically important as well.[32,44] Both repetition and feedback enable the student to build knowledge constructions. This focusses the attention on knowledge's storage in, and retrieval from, memory using contextual cues to facilitate the transfer of learning from the learning context to the application context.[45] Furthermore, it has been emphasised that learning occurs when students are facilitated to learn in their zone of proximal development, as Vygotskiĭ[46] described. Moreover, learners may also need support in their learning process and personal development, which can be provided by a supportive social context.[47,48] Another important aspect for learning is motivation; an intrinsic motivation initiates learning activities and helps the student self-initiate learning behaviour.[49-52]

Observable Movement Quality (OMQ)-scale

The Observable Movement Quality-scale (OMQ-scale) was developed by Janssen et al.[25] and designed for children from 3 months to 16 years of age to assess movement quality, over time, for all age categories and diagnostic groups as a generic evaluative measurement instrument. The OMQ-scale is a 15-item scale, including 15 individual items on movement quality, which needs to be filled in alongside an age-adequate discriminative motor test or disease-specific test. The OMQ-scale is developed based on a 3-phase study involving paediatric physiotherapists. The study starts with semistructured interviews, followed by a structured meeting – using a nominal group technique[53,54] – to explore existing perspectives on the complex phenomenon of quality of movement and to identify all relevant conceptual

aspects to be included in the OMQ-scale.[25] During the following Delphi rounds, the final selection of items is made and their definitions set, resulting in a 15-item scale.[25]
The selection of OMQ-scale items is based on a conceptual construct. Each individual item focusses on a different element of observable movement quality; together, the items form the whole construct, resulting in a formative measurement model.[55] During the development of the scale, the challenge was to identify all items contributing to the construct to ensure a comprehensive measurement of the construct movement quality. All items included in the OMQ-scale should be relevant and comprehensible, showing an adequate reflection of the construct movement quality,[56] which is confirmed by the establishment of content validity.[25]
The individual items in a formative measurement model do not necessarily correlate with each other and, thus, are not interchangeable;[57,58] therefore, analyses of the internal structure – important for deciding how items might be combined into scales or subscales – for formative models can be ignored in the developmental process.[55]
In the OMQ-scale, the following aspects are included: appropriate fine motor movements, appropriate gross motor movements, fluency of movements, reduced muscle tone, increased muscle tone, tremors, slow and/or delayed movements, accelerated and/or abrupt movements, asymmetry in movements, accuracy, strength regulation, variation in movements, involuntary movements, automated movements and stereotyped movements.[25]
Because each individual item focusses on an element of observable movement quality, the expected level of performance – based on the child's age and developmental stage – the task performed and the environmental circumstances must be taken into account. Scoring demands an introspective judgement of movement quality based on systematic observations and internal reflection, which incorporates the therapist's knowledge, reasoning, and personal and working experiences.[7,8] The 15 items of the OMQ-scale are scored on a 5-point Likert scale. Consequently, the sum scores of the scale can range from 15 to 75. Lower scores indicate lower movement quality. Although the development process has established the OMQ-scale's content validity,[25] studies on remaining measurement properties are needed to validate its use in clinical practice.

Validation of measurement instruments

During the development of a measurement instrument, measurement properties must be established.[55,59] These measurement properties can help clinicians select the most appropriate measurement instrument.[60] The Consensus-based Standards for the Development of Measurement Instruments (COSMIN)[55,61] defined and tested methodological guidelines for measurement properties, which were followed during our studies.
Following the COSMIN guidelines,[56,60] the remaining measurement properties to be established – reliability, measurement error, hypotheses testing for construct validity and responsiveness – are considered equally important. Reliability refers to the extent to which the scores for patients who have not changed are the same for repeated measurements under several conditions, such as over time (test-retest), by different raters on the same

occasion (interrater reliability) or by the same raters on different occasions (intrarater reliability).[56] Furthermore, reliability includes measurement error, which is explained as the systematic and random error of a score that is not attributed to true changes in the construct to be measured.[56] Establishing the reliability of a measurement instrument is a valuable step in the process of determining its usefulness in clinical practice.

Hypotheses testing for construct validity is described as showing the degree to which the scores of a measurement instrument are consistent with hypotheses, with regard to, for example, internal relationships, relationships to scores of other instruments or differences among relevant groups, based on the assumption that the measurement instrument validly measures the construct to be measured.[56] A basic principle of construct validity is that hypotheses are formulated regarding the relationship of scores on the measurement instrument under study with scores of other instruments measuring similar or dissimilar constructs or the differences in the instrument scores among subgroups of patients.[55,56]

Responsiveness specifies the ability of a measurement instrument to detect change over time and therefore refers to the validity of a change in score.[56,62] Whenever a measurement instrument is to be used to assess the effectiveness of interventions – the score must change in proportion to the patient's state change and must remain stable when the patient is unchanged – then responsiveness is an important quality.[62] Not only will the assessment of responsiveness ensure that a change in performance over time will be large enough to be statistically significant for research purposes – that is, the smallest change that can be detected by the instrument beyond measurement error – but also precise enough to reflect meaningful change in clinical situations.[62] This meaningful change is the smallest change that is important to patients and is named minimal important change.[63,64] Knowledge of the responsiveness of a measurement instrument enables physiotherapists to determine the possibilities of a measurement instrument to evaluate patients over time.[65,66]

Aim and outline of this thesis

Aim

In physiotherapy, the assessment of movement quality is relevant to recognising motor problems, evaluating interventions and predicting recovery. The OMQ-scale was developed to observe and score movement quality relative to what is expected for a child's age, and content validity has been established;[25] however, the remaining measurement properties of the scale are yet to be assessed.

The aim of this thesis was to determine the reliability, validity and responsiveness of the OMQ-scale. A second aim was to investigate what students in physiotherapy need in their educational programs to develop observational skills as well as which didactical principles facilitate this learning.

The following research questions are being studied in this thesis:

- What is the construct validity of the OMQ-scale?

- What are the interrater reliability, intrarater reliability and measurement error of the OMQ-scale?
- What is the responsiveness of the OMQ-scale?
- What are the design principles for an educational program to develop observational skills for students of physiotherapy?

Outline

To answer the research questions, we conducted 5 studies. Those studies addressed the assessment of measurement properties of the OMQ-scale and the development of observational skills. This thesis is divided into 2 parts. The first part describes the validation of the OMQ-scale. The second part concerns learning of observational skills for students of physiotherapy, including recommendations for the design of an educational program. The thesis ends with a general discussion of the findings. In **Chapter 1**, an overview and definitions of terminology used in this thesis are given. **Chapter 2** presents the determination of the construct validity of the OMQ-scale using 7 hypotheses, defined to conform to the COSMIN standards. **Chapter 3**, studies the concurrent and predictive validity for movement quality outcomes, using the OMQ-scale and general movements assessment at 3 months of age. The study is based on data collected in a prospective, longitudinal cohort study for induvial neurodevelopmental trajectories over 5 years in children with perinatal asphyxia treated with hypothermia. **Chapter 4** focusses on the interrater reliability of the OMQ-scale in a cross-sectional study with a stratified sample of paediatric physiotherapists with a variety of clinical expertise based on work setting and work experiences. In **Chapter 5**, the reliability and responsiveness of the OMQ-scale are investigated in a prospective intervention study, with a pre-post design, in centres for paediatric physiotherapy practice. In **Chapter 6**, we examine the development of a proto-theory on the development of observational skills for physiotherapy students, which can be used to design an educational program for said cohort. To develop this proto-theory, we derived design principles from students, teachers, practitioners and researchers using a design-based methodology. **Chapter 7** summarises the main findings of this thesis. Finally, in **Chapter 8** the most important findings of the studies are discussed in the general discussion.

Because this thesis is based on published journal articles, some overlap will be inevitable.

References

1. Magill RA, Hall KG. A review of the contextual interference effect in motor skill acquisition. *Hum Mov Sci.* 1990;9(3-5):241-289.
2. Guthrie ER. *The Psychology of Learning.* revised. New York: Haper & Row; 1952.
3. Hadders-Algra M. Early human motor development: From variation to the ability to vary and adapt. *Neurosci Biobehav Rev.* 2018;90:411-427.
4. Boyce WF, Gowland C, Hardy S, et al. Development of a quality-of-movement measure for children with cerebral palsy. *Phys Ther.* 1991;71(11):820-828.
5. Schmidt RA, Wrisberg CA. Getting started. In: *Motor Learning and Performance, a Situation-Based Learning Approach.* 4th editio. Leeds: Human Kinetics; 2008:3-23.
6. Magill R, Anderson D. *Motor Learing and Control, Concepts and Applications.* 10th editi. Singapore; 2014.
7. Skjaerven LH, Kristoffersen K, Gard G. An eye for movement quality: a phenomenological study of movement quality reflecting a group of physiotherapists' understanding of the phenomenon. *Physiother Theory Pract.* 2008;24(1):13-27.
8. Knudson D V, Morrison CS. Interdisciplinary nature of qualitative analysis. In: *Qualitative Analysis of Human Movement*. 2nd editio. Leeds: Human Kinetics; 2002.
9. Thelen E. Motor development. A new synthesis. *Am Psychol*. 1995;50:79-95.
10. Edelman GM. Neural Darwinism: selection and reentrant signaling in higher brain function. *Neuron*. 1993;10(2):115-125.
11. Hadders-Algra M. The neuronal group selection theory: an attractive framework to explain variation in normal motor development. *Dev Med Child Neurol*. 2000;42:566-572.
12. Hadders-Algra M. Variation and variability: Key words in human motor development. *Phys Ther*. 2010;90(12):1823-1837.
13. Prechtl HFR. Qualitative changes of spontaneous movements in fetus and preterm infants are a marker of neurlogical dysfunction. *Early Hum Dev*. 1990;23:151-158.
14. Prechtl HFR. General movement assessment as a method of developmental neurology: new paradigms and their consequences. *Dev Med Child Neurol*. 2001;43:836-842.
15. Hadders-Algra M. The neuromotor examination of the preschool child and its prognostic significance. *Ment Retard Dev Disabil Res Rev*. 2005;11:180-188.
16. Shumway-Cook A, Woollacott MH. *Motor Control - Translating Research into Clinical Practice*. Fourth. London: Lippincott Williams & Wilkins; 2012.
17. Schmidt RA, Lee TD. *Motor Learning and Performance*. 5th ed. Leeds, UK: Human Kinetics; 2014.
18. Shumway-Cook A, Woollacott MH. *Motor Control: Translating Research into Clinical Practice*. 5th ed. London: Wolters Kluwer; 2017.
19. Willingham DB. A neuropsychollogical theory of motor skill learning. *Psychol Rev*. 1998;105(3):558-584.
20. Wallbott HG. Hand movement quality: a neglected aspect of nonverbal behavior in clinical judgement and person perception. *J Clin Psychol*. 1985;41(3):345-359.

21. Hedman LD, Quinn L, Gill-Body K, et al. White Paper: Movement system diagnoses in neurologic physiotherapy. *J Neurol Phys Ther*. 2018;42(2):110-117.
22. Sahrmann SA. The human movement system: our professional identity. *Phys Ther*. 2014;94(7):1034-1042.
23. Edwards I, Jones M, Carr J, Braunack-Mayer A, Jensen GM. Clinical reasoning strategies in physiotherapy. *Phys Ther*. 2004;84(4):312-330; discussion 331-335.
24. Wainwright SF, Shepard KF, Harman LB, Stephens J. Novice and experienced physiotherapist clinicians: a comparison of how reflection is used to inform the clinical decision-making process. *Phys Ther*. 2010;90(1):75-88.
25. Janssen AJWM, Diekema ETW, Dolder van R, Kollée LAA, Oostendorp RAB, Nijhuis-van derSanden MWG. Development of a movement quality measurement tool for children. *Phys* Ther. 2012;92:574-594.
26. Wright FV, Rosenbaum P, Fehlings D, Mesterman R, Breuer U, Kim M. The Quality Function Measure: Reliability and discriminant validity of a new measure of quality of gross motor movement in ambulatory children with cerebral palsy. *Dev Med Child Neurol*. 2014;56(8):770-778.
27. Sorsdahl AB, Moe-Nilssen R, Kaale HK, Rieber J, Strand LI. Change in basic motor abilities, quality of movement and everyday activities following intensive, goal-directed, activity-focused physiotherapy in a group setting for children with cerebral palsy. *BMC Pediatr*. 2010;10:26.
28. DeMatteao C, Law M, Russell D, Pollock N, Rosenbaum P, Walter S. The relialbilty and validity of the Quality of Upper Extremity Skills Test. *Phys Occup Ther Pediatr*. 1993;13(2):1-18.
29. Hadders-Algra M. General movements: A window for early identification of children at high risk for developmental disorders. *J Pediatr*. 2004;145(2 SUPPL.):12-18.
30. Hecker E, Baer GD, Stark C, Herkenrath P, Hadders-Algra M. Inter- and intrarater Reliability of the Infant Motor Profile in 3- to 18-months-old infants. *Pediatr Phys Ther*. 2016;28(2):217-222.
31. Soanes C, Stevenson A. *Oxford Dictionary of English*. 2nd ed. (Soenes C, Stevenson A, eds.). Oxford: Oxford University Press; 2005.
32. Boudreau JD, Cassell EJ, Fuks A. Preparing medical students to become skilled at clinical observation. *Med Teach*. 2008;30(9-10):857-862.
33. Berger L. A hierachy of clinical observations. *Pediatrics*. 1980;65(2):357-358.
34. Al-Moteri MO, Plummer V, Cooper S, Symmons M. Training paradigms to enhance clinical observational skills in clinical practice : A scoping review. *J Nurs Educ Pract*. 2015;5(12):96-103.
35. Jarodzka H, Scheiter K, Gerjets P, van Gog T. In the eyes of the beholder: How experts and novices interpret dynamic stimuli. *Learn Instr*. 2010;20(2):146-154.
36. Shapiro J, Rucker L, Beck J. Training the clinical eye and mind: using the arts to develop medical students' observational and pattern recognition skills. *Med Educ*. 2006;40(3):263-268.
37. Wellbery C, McAteer RA. The art of observation: a pedagogical framework. *Acad Med*. 2015;90(12):1624-1630.
38. Maloney S, Storr M, Paynter S, Morgan P, Ilic D. Investigating the efficacy of practical skill teaching: A pilot-study comparing three educational methods. *Adv Heal Sci Educ*. 2013;18(1):71-80.

39. Gurwin J, Revere KE, Niepold S, et al. A Randomized Controlled Study of Art Observation Training to Improve Medical Student Ophthalmology Skills. *Ophthalmology*. 2017;125(1):8-14.
40. Bardes CL, Gillers D, Herman AE. Learning to look: Developing clinical observational skills at an art museum. *Med Educ*. 2001;35(12):1157-1161.
41. Bleakley, A; Farrow, R; Gould DE Al. Making sense of clinical reasoning: judgment and the evidence of the senses. *Med Educ*. 2003;37(6):544-552.
42. Flexner A. *Medical Eductaion in the United States and Canada. A Report to the Carnegie Foundation for the Advancement of Teaching*. Bull. four. New York: Updyke; 1910.
43. Cabot RC, Locke EA. The organization of a department of clinical medicine. *Bost Med Surg J.* 1905;73(17):461-465.
44. Kluger AN, van Dijk D. Feedback, the various tasks of the doctor, and the feedforward alternative. *Med Educ.* 2010;44(12):1166-1174.
45. Norman G, Bordage G, Page G, Keane D. How specific is case specificity? *Med Educ*. 2006;40(7):618-623.
46. Omrod J. *Human Learning*. 4th ed. Upper Saddle River: Pearson Prentice Hall; 2004.
47. Brown AL, Ash D, Rutherford M, Nakagawa K, Gordon A, Campione JC. Distributed expertise in the classroom. In: Salomon G, ed. *Distributed Cognitions; Psychological and Educational Considerations*. Cambridge: Cambridge University Press; 1993:188-228.
48. Rasmussen J. The importance of communication in teaching: A systems-theory approach to the scaffolding metaphor. *J Curric Stud*. 2001;33(5):569-582.
49. Ryan R, Deci E. Self-determination theory and the facilitation of intrinsic motivation. *Am Psychol*. 2000;55(1):68-78.
50. Zimmerman BJ. Becoming a self-regulated learner: an overview. *Theory Pract*. 2002;41(2):64-70.
51. Greene JA, Azevedo R. A Theoretical Review of Winne and Hadwin's Model of Self-Regulated Learning: New Perspectives and Directions. *Rev Educ Res*. 2007;77(3):334-372.
52. Schön D. The reflective Practitioner. In: *How Professionals Think in Action*. San Fransisco: Jossey-Bass Inc.; 1983.
53. Potter M, Gordon S, Hamer P. The Nominal Group Technique: A useful consensus methodology in physiotherapy research. *NZ J Physiother*. 2004;32(3):126-130.
54. van de Ven A, Delbecq A. The Nominal Group as a Research Instrument for Exploratory Health Studies. *Am J Public Health*. 1972;March:337-342.
55. Mokkink LB, de Vet HCW, Prinsen CAC, et al. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. *Qual Life Res*. 2018;27(5):1171-1179.
56. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*. 2010;63(7):737-745.
57. Edwards JR, Bagozzi RP. On the nature and direction of relationships between constructs and measures. *Psychol Methods*. 2000;5(2):155-174.
58. Fayers PM, Hand DJ, Bjordal K, Groenvold M. Causal indicators in quality of life research. *Qual life Res*. 1997;6(5):393-406.

59. Mokkink LB, Terwee CB, Knol DL, et al. Protocol of the COSMIN study: Consensus-based Standards for the selection of health Measurement INstruments. *BMC Med Res Methodol*. 2006;6:2.
60. Mokkink LB, Prinsen CAC, Bouter LM, de Vet HCW, Terwee CB. The COnsensus-based Standards for the selection of health measurement INstruments (COSMIN) and how to select an outcome measurement instrument. *Brazilian J Phys Ther*. 2016;20(2):105-113.
61. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res*. 2010;19(4):539-549.
62. Portney LG, Watkins MP. *Foundations of Clinical Research, Applications to Practice*. Third Edit. London: Pearson Education International; 2009.
63. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials*. 1989;10(4):407-415.
64. Stratford PW, Binkley JM, Riddle DL, Guyatt GH. Sensitivity to change of the Roland- Morris Back Pain Questionnaire : Part 1. *Phys Ther*. 1998;78(11):1186-1196.
65. Ostelo RWJG, de Vet HCW. Clinically important outcomes in low back pain. *Best Pract Res Clin Rheumatol*. 2005;19(4):593-607.
66. Revicki DA, Cella D, Hays RD, Sloan JA, Lenderking WR, Aaronson NK. Responsiveness and minimal important differences for patient reported outcomes. *Health Qual Life Outcomes*. 2006;4:1-5.

Part I

Measurement properties of the OMQ scale

Chapter 2

Construct validity of the Observable Movement quality scale in paediatrics – hypotheses testing of a formative measurement model

Lieke MA Dekkers

Anjo JWM Janssen

A Rogier T Donders

Maria WG Nijhuis-van der Sanden

Bert JM de Swart

Published in: Physical Therapy Journal 2020; 100(2);346-358

Background

The Observable Movement Quality (OMQ) scale measures generic movement quality. Each item of the OMQ Scale focuses on a different element; together, the 15 items assess the whole construct of movement quality.

Objective

The aim of this study was to determine the construct validity of the OMQ scale using 7 hypotheses, defined to conform to the Consensus-Based Standards for the Selection of Health Measurement Instruments.

Design

This was an exploratory validation study.

Methods

A paediatric physiotherapist assessed motor performance in 101 children using an age-specific motor test and the OMQ scale. The direction, magnitude, and rationale for 7 hypotheses, which concerned relationships (n = 2), probability of low scores (n = 4), and difference between diagnosis subgroups (n = 1), were defined.

Results

The results confirmed 6 of the 7 hypotheses, indicating sufficient construct validity. Significant positive relationships were found between OMQ scale total scores and the severity of motor disabilities (r = 0.72) and z scores on motor tests (r = 0.60). Probabilities for low scores on OMQ scale items—exceeding the chi-square critical value—were confirmed for children diagnosed with spasticity, psychomotor retardation, mitochondrial diseases, and ataxia; however, probabilities for low strength regulation in children with ataxia were not confirmed. OMQ scale total scores for children who were not ambulatory because of neurological conditions were significantly different from those for children who were not ambulatory because of fatigue (r = 0.66).

Limitations

The sample of children was based on theoretical assumptions about relevant variations in clinical representations; on the basis of the results, it appears that children with low strength regulation were underrepresented.

Conclusion

The confirmation of nearly all hypotheses supported the validity of the OMQ scale for measuring movement quality in clinical practice in addition to standardized age-adequate motor performance tests.

Introduction

The assessment of movement quality is relevant for recognizing motor problems, evaluating interventions and predicting recovery.[1–5] Movement quality is represented by an interaction of personal characteristics and experiences, the task difficulty and environmental condition, and gives an impression on how movements are controlled and coordinated.[6–8] Furthermore, movement quality gives insights in the potential of the movement system to react or adapt to changing conditions.[6–10]

During motor development and motor learning, new movements and skills are mastered, showing both quantitative and qualitative changes.[11] Quantitative changes can be seen in the acquisition of new and more complex skills, whereas changes in the quality of movements are demonstrated by more subtle characteristics (e.g., accuracy, fluency, and automatization of movements).[12] Available and commonly used discriminative motor tests in paediatric physiotherapy specifically assess quantitative aspects by comparison with peers. These motor tests are validated, and norm referenced. For movement quality, however, available and commonly used measurement instruments were designed for particular diagnostic groups, for children in a specific age frame, or to assess the functioning of extremities.[4,5,13–15] Despite the frequently used descriptions of movement quality,[7,16–18] a standardized, generic instrument was not available to assess movement quality in children over time and for all age categories.[19] This hinders both comparability and longitudinal evaluation in clinical practice, and the education of observational skills for students in paediatric physiotherapy.[10,17,18,20,21]

The Observable Movement Quality (OMQ) scale[16] was designed to assess movement quality in children, over time, for all age categories and diagnostic groups, as a generic evaluative measurement instrument. The OMQ scale includes 15 items on aspects of movement quality, which can be filled in by a paediatric physiotherapist after the assessment with an age-specific discriminative motor test or disease-specific test. The development of the OMQ scale was based on a 3-phase study involving paediatric physiotherapists. The study started with semistructured interviews, followed by a structured meeting—using a nominal group technique[22]—to explore existing perspectives on the complex phenomenon of quality of movement and to identify all relevant conceptual aspects to be included in the OMQ scale.[16] During the following Delphi rounds, the final selection of items was made and their definitions were set, resulting in a 15-item scale.[16]

The selection of OMQ scale items was based on a conceptual construct. Each individual item of the OMQ scale focuses on a different element of observable movement quality; together, the items form the whole construct—quality of movement—creating what is defined as a formative measurement model.[23–25] The challenge for such a formative measurement model is to identify all items contributing to the construct, ensuring a comprehensive measurement of the construct. The individual items in such a measurement model do not necessarily correlate with each other and, thus, are not interchangeable.[23,24] This is in contrast to a

reflective measurement model, in which items are manifestations of the construct—implying correlations of the items—and the possibility that they may replace each other.[23–25] During development of the OMQ scale, content validity was established that showed that the content of the OMQ scale is an adequate reflection of the construct movement quality.[16] Recently, a first study was published on the interrater reliability of the OMQ scale for children between 6 months and 6 years of age, showing a moderate interrater reliability when being used by paediatric physiotherapists unfamiliar with the scale.[19] The next step of the validation of the OMQ scale is to provide evidence of validity. Considering the lack of a gold standard, construct validation—or hypotheses testing—is the most adequate method to provide this evidence.[25]

At the start of our research, and before examining the data, we formulated hypotheses using the Consensus-Based Standards for the Selection of Health Measurement Instruments (COSMIN) guidelines.[26,27] Two independent paediatric physiotherapists, not involved in this research, contributed to hypothesis formation during a meeting. Based on this meeting, 3 authors (LD, AJ, and MN) all with adequate clinical and research expertise in paediatric physiotherapy—came to a consensus on 7 independent hypotheses. The hypotheses, which were specific and clearly defined, indicated the direction, magnitude, and rationale (Table 1). Therefore, the aim of this research was to determine the construct validity of the OMQ scale by investigating the degree to which the scores of the OMQ scale are consistent with hypotheses.[25,26]

Table 1. Hypotheses for testing construct validity of the OMQ scale.

Hypotheses	Rationale	Expected value
1) There will be a *positive relationship* between the severity of motor disabilities—as classified by paediatric physiotherapists—and OMQ total scores.	Paediatric physiotherapists classify motor impairments based on the child's performance; a relationship between the severity of motor impairments and quality of movements has been described	Correlation strength*: $0.50 < r < 0.75$
2) There will be at least a fair *significant positive* correlation with z-scores on motor tests [BSID-III-NL or MABC-2-NL] and OMQ total scores	Results on quantitative motor tests are related to quality of movement in children with known pathologies or developmental delays; however, some children show a normal quantitative development and show simultaneously deviant quality of movements or vice versa	Correlation strength: $0.40 < r_s < 0.70$
3) The *probability*# of low OMQ item scores will exceed the critical value in children diagnose with spasticity in both increased muscle tone [item 5], and variation in movements [item 12]	Increased muscle tone and reduced variation in movements have been identified in literature as a sign or symptom for spasticity	Critical value: ≥3.84 for $p = 0.05$
4) The *probability*# of low OMQ item scores will exceed the critical value in children diagnose with psychomotor retardation on stereotype movements [item 15]	Stereotype movements, defined as movements that are both aimless and repetitive, have been identified in literature as a sign related to psychomotor retardation	Critical value: ≥3.84 for $p = 0.05$
5) The *probability*# of low OMQ item scores will exceed the critical value in children diagnose with mitochondrial disease confirmed by mitochondrial DNA on strength regulation [item 11]	Strength regulation has been identified in literature as problematic for children with a predominantly muscular mitochondrial disease	Critical value: ≥3.84 for $p = 0.05$
6) The *probability*# of low OMQ item scores will exceed the critical value in children diagnose with ataxia on tremors [item 6], accuracy [item 10], and strength regulation [item 11]	The presents of tremors and reduced accuracy of movements and strength regulation have been identified in literature as signs or symptoms for ataxia	Critical value: ≥3.84 for $p = 0.05$
7) There will be a statistically *significant difference* between OMQ scores for children who are not ambulatory because of a neurological condition and OMQ scores for children who are not ambulatory because of fatigue, caused by, e.g., a mitochondrial disease	The reason for wheelchair use in children with neurological conditions (i.e., they cannot walk) differs from that in children with fatigue (i.e., they lack endurance). Wheelchair users with neurological conditions are expected to show poorer movement quality than those affected by fatigue.	Correlation strength: $0.50 < r < 0.75$

OMQ scale = Observable Movement Quality scale; BSID-III BSID-III-NL = Bayley Scales of Infant and Toddler Development, 3rd Edition, Dutch version; MABC-2-NL = Movement Assessment Battery, 2nd edition, Dutch version. #Probability = the likelihood that any one event will occur, given all the possible[40, pg. 405], The expected probability of low scores in children with the diagnosis was <20% higher than in children without the diagnosis. *Correlation strength: 0.00 to 0.25, little or no relationship; 0.25 to 0.50, fair relationship; 0.50 to 0.75, moderate to good relationship; above 0.75 good to excellent relationship[40]

Methods

In an explorative validation study, the construct validity of the OMQ scale for children is determined. The study used anonymized data sampled in daily clinical practice from 2013 until 2017. The regional medical ethics committee of the Radboud University Medical Center (Radboudumc) agreed that the study conformed to the Declaration of Helsinki and that approval was not required. This committee waived the requirement to obtain informed consent (registration no.: 2018–4842).

The data from a convenience sample of 101 children were available for analyses. All data were retrospectively extracted from patient files—by the paediatric physiotherapist who performed the assessments—and anonymously sampled in a database. No repeated measurements of children were included. Data collection primarily took place as part of a multidisciplinary assessment during diagnostic trajectories for children with suspected mitochondrial dysfunction or disease. This trajectory was chosen because these children are known to show both a wide range of motor problems—with additional signs and symptoms—or an almost normal development. To ensure even sample sizes per age group, for sex, and for a diversity in diagnosis, we added data of cases of outpatient multidisciplinary evaluations from other trajectories (e.g., children preterm born or diagnosed with ataxia telangiectasia). All data were collected at the Radboudumc, Nijmegen, the Netherlands.

Motor Performance Assessment

As part of a multidisciplinary assessment during diagnostic trajectories or outpatients' evaluations, 1 certified paediatric physiotherapist (AJ)—with over 30 years of clinical experience and involved in the development and education of the OMQ scale—assessed motor performance. Motor performance was assessed by an age-appropriate motor test: the motor scales of the Bayley Scales of Infant and Toddler Development, 3rd Edition, Dutch version (BSID-III-NL),[28–30] for infants and children between 0 and 3 years old and the Movement Assessment Battery for children, 2nd edition, Dutch version (MABC-2-NL), for children between 3 and 16 years old.[31,32] The Gross Motor Function Measure and the Gross Motor Function Classification System[33,34] were used to measure and classify the subgroup of children with spasticity in both age-groups. Furthermore, the Gross Motor Function Measure was used in children diagnosed with psychomotor retardation when assessment with BSID-III-NL or MABC-2-NL was not possible because of cognitive disabilities.[35] The Scale for the Assessment and Rating of Ataxia[36,37] was used for children diagnosed with ataxia telangiectasia. The Motor Function Measure[38] was used to assess a child with suspected neuromuscular disease, and the Pediatric Balance Scale[39] was used in a child with mild ataxia.

Movement Quality

Movement quality was assessed by 1 paediatric physiotherapist (AJ), using the OMQ scale,[16] which was designed for children from 3 months to 16 years of age. The 15-item scale needs

to be filled in after the assessment of an age-specific discriminative motor test or disease-specific test. The OMQ scale scores movement quality relative to what is expected for a child's age. The 15 items are scored on a 5-point Likert scale (Appendix).

Data Methods

The descriptive data are presented as numbers and percentages for categorical variables and as median and interquartile range (IQR) for ordinal variables. For continuous data, means and SDs are reported.

Based on the clinical presentation and results of motor performance assessments, children were classified for motor disabilities by the paediatric physiotherapist into severe motor disabilities, mild motor disabilities, and no motor disabilities. Criteria for severe motor disabilities and mild motor disabilities are provided in Table 2. To make the outcomes of BSID-III-NL and MABC-2-NL comparable, the scores of motor tests were converted to *z* scores using an algorithm derived from the literature.[40] *Z* scores of ≤−2 indicated significantly delayed performance; scores of ≤−1 but >−2 indicated mildly delayed performance; and scores of >−1 indicated performance within normal limits.

Table 2. Criteria for outcome classification based on clinical performance and motor performance assessment.

Hypotheses
Severe motor disability classification Must meet one or more of the following criteria: • z-score on BSID-III-NL or MABC-2-NL less than or equal to –2 • GMFSC level III, IV or V • GMFM score below 55th percentile on motor growth curves • assessment with SARA • assessment with MFM • PBS total score below 35 (range 0 [no balance] to 56 [functional balance])
Mild motor disability classification: Must meet one or more of the following criteria: • z-score BSID-III-NL or MABC-2-NL less than or equal to –1 but greater than –2 • GMFCS level I or II • GMFM score above 56th percentile on motor growth curves • PBS total score between 35–44

BSID-III-NL = Bayley Scales of Infant and Toddler Development, 3rd edition, Dutch version, MABC-2-NL = Movement Assessment Battery, 2nd edition, Dutch version, GMFCS = Gross Motor Function Classification Scale, GMFM = Gross Motor Function Measure, SARA = Scale for rating and assessing Ataxia, MFM = Motor Function Measure, PBS = Pediatric Balance Scale.

Nonparametric tests were used to test hypotheses for comparisons without normality or variance assumptions (Table 1). For group comparisons, the Jonckheere-Terpstra test[41,42] for ordered alternatives was used when 3 independent groups were compared (hypothesis 1), and the Mann-Whitney *U* test[43] was used when 2 independent groups were compared (hypothesis 7); in addition, the effect size estimate, *r*—as described by Rosenthal[44]—was calculated to increase possibilities for interpretation. Spearman rank correlation coefficients were used to test correlation between *z* scores for motor tests and OMQ scale total scores (hypothesis 2).[40] Correlations were considered as follows: 0.00–0.25 = little or no

relationship; 0.25–0.50 = fair relationship; 0.50–0.75 = moderate to good relationship; and >0.75 = good to excellent relationship.[40] Furthermore, cross-tabulations and Pearson chi-square tests were used to test relationships within hypotheses for diagnoses and OMQ scale item scores (hypotheses 3–6).[40] Statistical analyses were performed with the IBM Statistical Package for the Social Sciences (IBM SPSS Statistics), version 25 (IBM Corp, Armonk, NY, USA). All statistical tests were 2-tailed, and a *P* value of <.05 was considered significant.

Role of the Funding Source
The Netherlands Organization for Scientific Research (NWO), which provided support to L.M.A. Dekkers (grant no. 023.004.037), had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Results

This study reviewed data of 101 children, including 51 boys (50.5%), with a mean age of 8 years and 6 months (SD = 5 years 4 months) (Table 3). For all but 2 children, a classification for motor disabilities by the paediatric physiotherapist based on clinical presentations was possible. The 2 missing classifications were due to difficulties in interpretation because of severe cognitive disabilities—e.g., low test scores but able to walk independently and to climb stairs. Fifty-six children (55.4%) were classified has having severe motor disabilities, 17 children (16.8%) were classified as having mild motor disabilities, and 26 children (25.7%) were classified as having no motor disabilities.

Motor performance was successfully assessed in 82 children (81.2%) with the BSID-III-NL, MABC-2-NL, or Gross Motor Function Measure. In 10 children (9.9%), only partial assessment with MABC-2-NL was possible because of the severity of their disabilities. Nine children (8.9%) were assessed with disease-specific tests (e.g., Scale for the Assessment and Rating of Ataxia). Outcomes on the BSID-III-NL (n = 21) showed a mean z score of −0.7 (SD = 1.6), and outcomes on the MABC-2-NL (n = 42) showed a mean z score of –1.6 (SD = 1.3). Data from the OMQ scale were available for all 101 children; outcomes on the OMQ scale showed a mean total score of 61.7 (SD = 11.0). Frequencies and percentages, means and SDs, and medians and IQRs for each item are presented in Table 4.

The relationship between the severity of motor disabilities and OMQ scale total scores—hypothesis 1 (n = 99)—was tested by a Jonckheere-Terpstra test, which showed a significant trend in the data: the more severe the motor disabilities, the lower the OMQ scale total scores (J = 2474.00; z = 7.13; P < .001; r = 0.72). Children classified with severe motor disabilities (n = 56) showed a median OMQ scale total score of 56.5 (IQR = 15); children classified with mild motor disabilities (n = 17) had a median OMQ scale total score of 68 (IQR = 9); and children classified with no motor disabilities (n = 26) had a median OMQ scale total score of 73 (IQR = 6.3). A box-plot for the relationship between severity of motor disabilities

and OMQ scale total scores showed larger variations in OMQ scale total scores for children with severe motor disabilities than for children in the other groups (Figure 1).

Table 3. Characteristic of included children (*n* = 101), reason for diagnostic trajectory or outpatient hospital visit, classification for motor disabilities, results on discriminative motor tests (z-scores), categorization of children with spasticity, and results on OMQ scale.

Characteristics	*n*	mean	(SD)
Male	51		
Female	50		
Age in years	101	8yr6mth	(5yr4mth)
Mitochondrial dysfunction			
Mitochondrial disease confirmed on (mt)DNA	41		
Suspected of mitochondrial dysfunction	39		
Psychomotor retardation	30		
Ataxia	23		
Spasticity	11		
Wheel chair use			
Due to neurological disorder	25		
Due to fatigue	20		
Reason for diagnostic trajectory or outpatient hospital visit			
Mitochondrial dysfunction	80		
Preterm born (<30 weeks GA)	7		
Ataxia Telangiectasia	7		
Perinatal asphyxia treated with hypothermia	4		
MAS or CHD needing ECMO	3		
Classification for motor disabilities by paediatric physiotherapist			
No motor disabilities	26		
Mild motor disabilities	17		
Severe motor disabilities	56		
Result on discriminative motor test (z-scores) or disease specific test			
BSID-III-NL	21	–0.73	(1.60)
MABC-2-NL	42	–1.64	(1.31)
GMFM#	19		
MFM	1		
PBS	1		
SARA	7		
Categorization of children with spasticity			
GMFCS level			
I	1		
III	1		
IV	4		
V	5		
Results on OMQ scale			
OMQ scale total scores	101	61.7	(10.99)

SD = Standard Deviation; yr = years; mth = months; GA = Gestational Age; (mt)DNA = mitochondrial deoxyribonucleic acid; MAS = Meconium Aspiration Syndrome; CHD = Congenital Hernia Diaphragmatic syndrome; ECMO = Extracorporeal Membrane Oxygenation; BSID-III-NL = Bayley Scales of Infant and Toddler Development, 3rd edition, Dutch version; MABC-2-NL = Movement Assessment Battery, 2nd edition, Dutch version; GMFM = Gross Motor Function Measure, #used to assess children with spasticity and psychomotor retardation; GMFCS = Gross Motor Function Classification System; OMQ scale = Observable Movement Quality Scale, MFM = Motor Function Measure, PBS = Pediatric Balance Scale, SARA = Scale for Assessment and Rating of Ataxia.

Table 4. Frequencies (percentages) of individual OMQ item scores, means (SD), and median (IQR) for each item (*n* = 101).

		Frequency (%) of Score of:						
	Item	**1**	**2**	**3**	**4**	**5**	**Mean (SD)**	**Median (IQR)**
1	Appropriate fine motor movements	6 (5.9)	21 (20.8)	12 (11.9)	26 (25.7)	36 (35.6)	3.6 (1.3)	4 (3)
2	Appropriate gross motor movements	6 (5.9)	17 (16.8)	22 (21.8)	28 (27.7)	28 (27.7)	3.5 (1.3)	4 (2)
3	Fluency of movements	8 (7.9)	14 (13.9)	16 (15.8)	30 (29.7)	33 (32.7)	3.7 (1.3)	4 (2)
4	Reduced muscle tone	2 (2.0)	7 (6.9)	12 (11.9)	28 (27.7)	52 (51.2)	4.2 (1.0)	5 (1)
5	Increased muscle tone	3 (3.0)	3 (3.0)	8 (7.9)	18 (17.8)	69 (68.3)	4.5 (1.0)	5 (1)
6	Tremors	3 (3.0)	11 (10.9)	5 (5.0)	10 (9.9)	72 (71.3)	4.4 (1.2)	5 (1)
7	Slow and/or delayed movements	2 (2.0)	13 (12.9)	22 (21.8)	15 (14.9)	49 (48.5)	4.0 (1.2)	4 (2)
8	Accelerated and/or abrupt movements	3 (3.0)	5 (5.0)	8 (7.9)	11 (10.9)	74 (73.3)	4.5 (1.0)	5 (1)
9	Asymmetry in movements	3 (3.0)	4 (4.0)	9 (8.9)	37 (36.6)	48 (47.5)	4.2 (1.0)	4 (1)
10	Accuracy (well-aimed)	5 (5.0)	11 (10.9)	21 (20.8)	29 (28.7)	35 (34.7)	3.8 (1.2)	4 (2)
11	Strength regulation	4 (4.0)	9 (8.9)	19 (18.8)	37 (36.6)	32 (31.7)	3.8 (1.1)	4 (2)
12	Variation in movements	2 (2.0)	4 (4.0)	12 (11.9)	25 (24.8)	58 (57.4)	4.3 (1.0)	5 (1)
13	Involuntary movements	1 (1.0)	3 (3.0)	2 (2.0)	16 (15.8)	79 (78.2)	4.7 (0.8)	5 (0)
14	Automated movements	1 (1.0)	10 (9.9)	15 (14.9)	33 (32.7)	42 (41.6)	4.0 (1.0)	4 (2)
15	Stereotype movements	1 (1.0)	5 (5.0)	6 (5.9)	3 (3.0)	86 (85.1)	4.7 (0.9)	5 (0)

OMQ = Observable Movement Quality Scale, SD = standard deviation, IQR = Inter Quartile Range

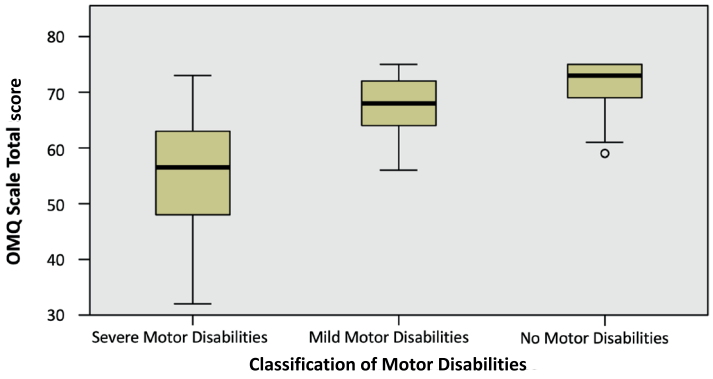


Figure 1. Box plot showing correlation between severity of motor disabilities and Observable Movement Quality (OMQ) Scale total scores (n = 99).

There was a significant positive relationship between OMQ scale total scores and z scores on motor tests—hypothesis 2 (n = 63)—as shown by the Spearman rank correlation (r_s = 0.595; 95% Bias-corrected and accelerated Confidence Interval = 0.381–0.750; P < .001). A scatterplot (see Figure 2) displays OMQ scale total scores and z scores on motor tests, showing more scattering of OMQ scale total scores in children with low z scores on motor tests than in children with high z scores.

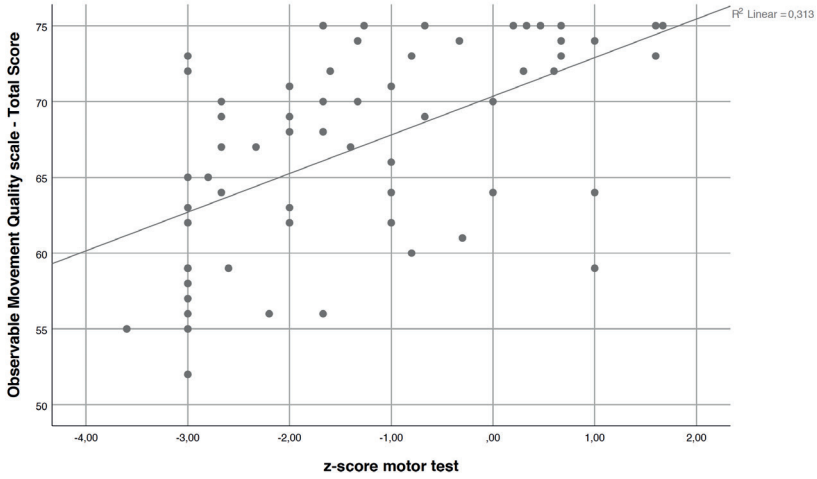


Figure 2. Scatterplot showing correlation of Observable Movement Quality scale total scores and z-scores on motor tests (n = 63).

To test the possible existence of relationships between diagnoses and OMQ scale items, cross-tabulation and Pearson chi-square tests were performed. The probabilities for scores on OMQ scale items in children diagnosed with spasticity, psychomotor retardation, mitochondrial diseases, and ataxia are presented in a contingency table (Table 5). All hypothesized relationships between diagnosis and low scores on OMQ scale items exceeded the critical value of 6.663 and were significant ($P < .005$), except for the diagnosis ataxia and the expected low scores on item 11 in which the Pearson chi-square test showed a value of 1.91 ($P = .13$).

Table 5. Contingency table showing the relationship between diagnoses and item scores on the OMQ scale ($n = 101$): hypotheses 3–6.

Diagnose		Item	X^2 (1, $n = 101$)	% of: Low score (score 1-2-3)	% of: Normal score (score 4-5)	Confirmed?
Spasticity	Yes	5	17.11, $p < 0.001$	54.5	45.5	Yes
	No			8.9	91.1	
	Yes	12	25.41, $p < 0.001$	72.7	27.3	Yes
	No			11.1	88.9	
Psychomotor retardation	Yes	15	25.04, $p < 0.001$	36.7	63.3	Yes
	No			1.4	98.6	Yes
Mitochondrial disease confirmed	Yes	11	6.85, $p < 0.01$	46.3	53.7	Yes
	No			21.7	78.3	Yes
Ataxia	Yes	6	50.23, $p < 0.001$	69.6	30.4	Yes
	No			3.8	96.2	
	Yes	10	22.23, $p < 0.001$	78.3	21.7	Yes
	No			24.4	75.6	
	Yes	11	1.91, $p = 0.13$	43.5	56.5	Yes
	No			28.2	71.8	

OMQ scale=observable Movement Quality scale, X^2= Pearson's Chi-Square

A Mann-Whitney test indicated a statistically significant difference ($U = 56.5$; $z = –4.42$; $P < .001$; $r = 0.66$) between OMQ scale total scores for children who are not ambulatory because of a neurological condition and OMQ scores for children who are not ambulatory because of fatigue caused, for example, by a mitochondrial disease (Figure 3). Thus, results for the construct validity testing were that 6 of the 7 hypotheses (85.7%) were confirmed, although hypothesis 6 was rejected for only 1 of the 3 items (item 11); these data were judged as sufficient construct validity.

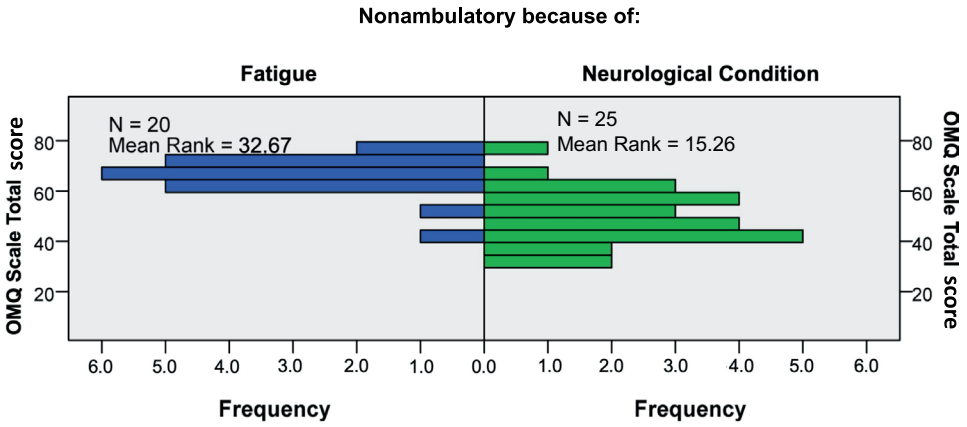


Figure 3. Comparison of total scores on Observable Movement Quality (OMQ) Scale for 2 groups of children who were not ambulatory (*n* = 45).

Discussion

This is the first study to assess the construct validity of the OMQ scale. According to the COSMIN criteria, the construct validity of an instrument is sufficient when 75% of the predefined hypotheses are confirmed in a sample of at least 50 patients.[45] With an 86% rate of confirmed predefined hypotheses in this study of 101 children, the COSMIN quality criteria were met. We used the overall quality criterion for adequate sample sizes for validation studies of measurement instruments, which were set by the COSMIN group,[46] as a guidance for this study. Furthermore, in addition to needing an appropriate sample size, even sample sizes per age group and sex were needed because the OMQ scale was designed for children in the broad age range of 3 months to 16 years. Therefore, data from other groups, also assessed during a multidisciplinary evaluation, were added to the data from children with suspected mitochondrial dysfunction. This was done to ensure that a diversity of motor problems was included in the sample.

Assessing the relationship between the severity of motor disabilities and OMQ scale total scores revealed a greater range of OMQ scale total scores in children with severe motor disabilities than in children with mild or no motor disabilities. A similar pattern was noticed when assessing the relationship between *z* scores on motor tests and OMQ scale total scores. A closer examination of the data showed that 3 children with high OMQ scale total scores were classified with severe motor disabilities—or a *z* score of ≤−2, indicating a significant delay for motor performances. However, these children scored more than 1 IQR above the subgroups' median for OMQ scale total scores. These 3 children were all suspected to have mitochondrial dysfunction. This indicates that the classification of severe disabilities in these children with good quality of movement was based on their limited functioning in daily life because of other aspects, such as severe fatigue. A contrasting

outcome can be noticed in children classified as having no motor disabilities—or a *z* score indicating motor performance within normal limits—who scored more than 1 IQR below the subgroups' median on OMQ scale total scores. These 3 children—including the outlier shown in the box plot in Figure 1—were children born preterm. Those 3 preterm-born children were assessed at, respectively, 5, 11, and 27 months of age, and showed no delay in motor performance. However, they showed a low score for movement quality. At a young age, the spontaneous recovery of abnormal motor signs, and thereby quality of movement, was reported in literature,[47–49] indicating the potential for these children to catch up in their quality of movement. However, low scores for motor quality could also display that these children are at risk for future developmental delays. Both these clinical discrepancies were seen in longitudinal motor performance studies, using quantitative motor assessments, in which large variabilities in individual developmental trajectories were shown.[50,51] This confirms the necessity for paediatric physiotherapists to rely not only on the outcomes of the quantitative motor assessment or on the assessment of movement quality alone. Instead, physiotherapists should evaluate both factors and relate the outcomes to the functional capabilities of the child before developing a personalised therapy approach. However, in future studies, we need to test whether combined use of quantitative motor tests and the assessment of movement quality can refine prediction models for individual motor developmental trajectories. This can possibly contribute to the determination of the effects on motor development after developmental interventions.

The assessment of the relationship between diagnoses and OMQ scale item scores showed that the hypothesis of an expected probability for low scores on item 11 (strength regulation) in children with ataxia was not confirmed. Cerebellar damage that results in ataxia leads to the presence of tremors and increases the instability and poor accuracy of movements.[52–54] In the literature, the role of the cerebellum is mentioned in the regulation of muscle tone and its importance for balance control and the modulation of rhythmic agonist and antagonist muscle activity—necessary for adequate direction, timing, and amplitude of movements.[55–57] However, clear statements about muscle strength, or strength regulation, were not found. Even though in this study the presence of tremors and deviant outcomes for the level of accuracy of movements were scored, no deviant outcomes on strength regulation were scored, which agrees with literature but contradicts the predefined hypothesis. This hypothesis was possibly not formulated specifically enough.

The COSMIN taxonomy of measurement properties[25] states that all measurement properties included in their taxonomy are relevant and should be evaluated for each measurement instrument. Three quality domains are thereby distinguished: reliability, validity, and responsiveness.[25] Content validity was considered as the most important measurement property; all content (e.g., items) should be relevant, comprehensible, and comprehensive with respect to the construct of interest and target population.[25] During the developmental process of the OMQ scale, the selection of OMQ scale items was based on a conceptual construct of quality of movement (based on expert opinions), and content validity was

established.[16] Next, the internal structure (which includes structural validity and internal consistency and is crucial for item reduction and selection of subscales) is considered important. This is true, however, only for measurement instruments based on a reflective model.[25,27] Because the OMQ scale is a formative measurement model, an analysis evaluating the internal structure was not relevant.[27] We continued to evaluate the measurement property of construct validity in this study, which we showed to be sufficient. A first study on reliability of the OMQ scale—including first reports on measurement error—was published, reporting moderate interrater reliability.[19] A future study on responsiveness of the OMQ scale and remaining items of reliability (such as intrarater reliability) will complete the validation of the OMQ scale and will provide further evidence for the use of the OMQ scale in clinical practice.

A limitation of our study was that the inclusion of children with mitochondrial diseases was based on the theoretical assumption. Assumed was that the children would show a relevant variation in clinical representations and that they would show deviant outcomes in all aspects of quality of movement. Based on our results it appears that children with low strength regulation were underrepresented. Therefore, it will be beneficial for the validation of the OMQ scale to assess the hypotheses also within other subgroups such as neuromuscular diseases or syndromes. Moreover, additional hypotheses could have been formulated regarding muscle tone, strength regulation, and the timing of movements. Another limitation of this study was that this was a single centre study from 1 university hospital; a multicentre study would have been beneficial for the generalisability of the results.

Conclusion

The OMQ scale demonstrates a sufficient construct validity in children assessed for motor performance as part of a multidisciplinary assessment. Our findings indicate that the construct of the OMQ scale—based on expert opinion in the developmental phase—is valid and can be used in clinical practice. However, a future study for additional hypotheses testing in subgroups of children diagnosed with neuromuscular diseases or syndromes will be beneficial for increased statements on validity. Furthermore, a future study on responsiveness of the OMQ scale and remaining items of reliability will complete the validation of the OMQ scale and will provide further evidence for the use of the OMQ scale in clinical practice.

Disclosures

The authors completed the ICJME Form for Disclosure of Potential Conflicts of Interest and reported no conflicts of interest.

References

1. Knudson DV, Morrison CS. *Qualitative Analysis of Human Movement.* 2nd ed. Leeds, United Kingdom: Human Kinetics; 2002.
2. Heineman KR, Hadders-Algra M. Evaluation of neuromotor function in infancy: a systematic review of available methods. *J Dev Behav Pediatr.* 2008;29:315–323.
3. Harris SR, Heriza CB. Measuring infant movement: clinical and technological assessment techniques. *Phys Ther.* 1987;67:1877–1880.
4. Wright FV, Rosenbaum P, Fehlings D, Mesterman R, Breuer U, Kim M. The Quality Function Measure: reliability and discriminant validity of a new measure of quality of gross motor movement in ambulatory children with cerebral palsy. *Dev Med Child Neurol.* 2014;56:770–778.
5. Sorsdahl AB, Moe-Nilssen R, Kaale HK, Rieber J, Strand LI. Change in basic motor abilities, quality of movement and everyday activities following intensive, goal-directed, activity-focused physiotherapy in a group setting for children with cerebral palsy. *BMC Pediatr.* 2010;10:26.
6. Magill R, Anderson D. Motor control theories. In: Magill R, Anderson D, eds. *Motor Learning and Control: Concepts and Applications.* 10th ed. Singapore, Singapore: McGraw-Hill Education; 2014:85–110.
7. Skjaerven LH, Kristoffersen K, Gard G. An eye for movement quality: a phenomenological study of movement quality reflecting a group of physiotherapists' understanding of the phenomenon. *Physiother Theory Pract.* 2008;24:13–27.
8. Shumway-Cook A, Woollacott MH. *Motor Control: Translating Research into Clinical Practice.* 5th ed. London, United Kingdom: Wolters Kluwer; 2017.
9. Hedman LD, Quinn L, Gill-Body K, et al. White Paper: movement system diagnoses in neurologic physiotherapy. *J Neurol Phys Ther.* 2018;42:110–117.
10. Sahrmann SA. The human movement system: our professional identity. *Phys Ther.* 2014;94:1034–1042.
11. Boyce WF, Gowland C, Hardy S, et al. Development of a quality-of-movement measure for children with cerebral palsy. *Phys Ther.* 1991;71:820–828; discussion 828–832.
12. Schmidt RA, Wrisberg CA. Getting started. In: Schmidt RA, Wrisberg CA, eds. *Motor Learning and Performance: A Situation-Based Learning Approach.* 4th ed. Champaign, IL: Human Kinetics; 2008:3–23.
13. Hadders-Algra M. General movements: a window for early identification of children at high risk for developmental disorders. *J Pediatr.* 2004;145(2 suppl):12–18.
14. Hecker E, Baer GD, Stark C, Herkenrath P, Hadders-Algra M. Inter- and intrarater reliability of the infant motor profile in 3- to 18-month-old infants. *Pediatr Phys Ther.* 2016;28:217–222.
15. DeMatteao C, Law M, Russell D, Pollock N, Rosenbaum P, Walter S. The reliability and validity of the Quality of Upper Extremity Skills Test. *Phys Occup Ther Pediatr.* 1993;13:1–18.
16. Janssen AJWM, Diekema ETW, van Dolder R, Kollée LAA, Oostendorp RAB, Nijhuis-van der Sanden MWG. Development of a movement quality measurement tool for children. *Phys Ther.* 2012;92:574–594.

17. Edwards I, Jones M, Carr J, Braunack-Mayer A, Jensen GM. Clinical reasoning strategies in physiotherapy. *Phys Ther.* 2004;84:312–330
18. Wainwright SF, Shepard KF, Harman LB, Stephens J. Novice and experienced physiotherapist clinicians: a comparison of how reflection is used to inform the clinical decision-making process. *Phys Ther.* 2010;90:75–88.
19. Dekkers LMA, Nijhuis-van der Sanden MWG, Jonker M, de Swart BJM, Janssen AJWM. Interrater reliability of the Observable Movement Quality scale for children. *Physiother Can.* 2018;70:113–119.
20. Ahola S, Piirainen A, Skjaerven LH. The phenomenon of movement quality: a phenomenographic study of physiotherapy students' movement experiences. *Eur J Physiother.* 2017;19:59–68.
21. Skjaerven LH, Kristoffersen K, Gard G. How can movement quality be promoted in clinical practice? A phenomenological study of physiotherapist experts. *Phys Ther.* 2010;90:1479–1492.
22. Potter M, Gordon S, Hamer P. The nominal group technique: a useful consensus methodolgy in physiotherapy research. *New Zealand Journal of Physiotherapy.* 2004;32:126–130.
23. Edwards JR, Bagozzi RP. On the nature and direction of relationships between constructs and measures. *Psychol Methods.* 2000;5:155–174.
24. Fayers PM, Hand DJ, Bjordal K, Groenvold M. Causal indicators in quality of life research. *Qual life Res.* 1997;6:393–406.
25. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J *Clin Epidemiol.* 2010;63:737–745.
26. Mokkink LB, de Vet HCW, Prinsen CAC, et al. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. *Qual Life Res.* 2018;27:1171–1179.
27. Mokkink LB, Terwee CB, Knol DL, et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. *BMC Med Res Methodol.* 2010;10:22.
28. Bayley N. *Bayley Scales of Infant and Toddler Development/Third Edition (Bayley-III).* San Antonio, TX: The Psychological Corp; 2006.
29. Van Baar A, Steenis LJ, Verhoeven M, Hessen D. *Bayley-III-NL: Handleiding Motororiekschaal.* Amsterdam, the Netherlands: Pearson Assessment and Information BV; 2014.
30. Steenis LJP, Verhoeven M, Hessen DJ, van Baar A l. Performance of Dutch children on the Bayley III: a comparison study of US and Dutch norms. *PLoS One.* 2015;10:e0132871.
31. Henderson SE, Sugden DA, Barnett LA,. *Movement Assessment Battery for Children, Second Edition (MABC-2).* London, United Kingdom: Pearson Education International; 2010.
32. Smits-Engelsman BCM. *Movement Assessment Battery for Children-2: Dutch Standardization.* Amsterdam, the Netherlands: Pearson Education International; 2010.
33. Wood E, Rosenbaum P. The Gross Motor Function Classification System for cerebral palsy: a study of reliability and stability over time. *Dev Med Child Neurol.* 2000;42:292–296.
34. Weis R. *Gross Motor Function Measure (GMFM-66 and GMFM-88) User's Manual.* Vol 8. London, United Kingdom: McKeith Press; 2004.

35. Russell D, Palisano R, Walter S, et al. Evaluating motor function in children with Down syndrome: validity of the GMFM. *Dev Med Child Neurol.* 1998;40:693–701.
36. Schmitz-Hübsch T, Du Montcel ST, Baliko L, et al. Scale for the assessment and rating of ataxia. *Neurology.* 2006;66:1717–1720.
37. Weyer A, Abele M, Schmitz-Hübsch T, et al. Reliability and validity of the scale for the assessment and rating of ataxia: a study in 64 ataxia patients. *Mov Disord.* 2007;22:1633–1637.
38. Bérard C, Payan C, Hodgkinson I, Fermanian J. A motor function measure scale for neuromuscular diseases: construction and validation study. *Neuromuscul Disord.* 2005;15:463–470.
39. Franjoine MR, Gunther JS, Taylor MJ. Pediatric balance scale: a modified version of the Berg Balance Scale for the school-age child with mild to moderate motor impairment. *Pediatr Phys Ther.* 2003;15:114–128.
40. Portney LG, Watkins MP. *Foundations of Clinical Research: Applications to Practice.* 3rd ed. London, United Kingdom: Pearson Education International; 2009.
41. Jonckheere AR. A distribution-free k-sample test against ordered alternatives. *Biometrika.* 1954;41:133–145.
42. Terpstra TJ. The asymptotic normality and consistency of Kendall's test against trend, when ties are present in one ranking. *Indagationes Mathematicae.* 1952;14:327–333.
43. Nachar N. The Mann-Whitney U: a test for assessing whether two independent samples come from the same distribution. *Tutor Quant Methods Psychol.* 2008;4:13–20.
44. Rosenthal R. *Meta-Analytic Procedures for Social Research.* 2nd ed. Newbury Park, CA: Sage; 1991.
45. Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60:34–42.
46. Terwee CB, Mokkink LB, Knol DL, Ostelo RWJG, Bouter LM, de Vet HCW. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res.* 2012;21:651–657.
47. Piper MC, Mazer B, Silver KM, Ramsay M. Resolutions of neurological symptoms in high-risk infants during the first two years of life. *Dev Med Child Neurol.* 1988;30:26–35.
48. De Vries AM, De Groot L. Transient dystonias revisited: a comparative study of preterm and term children at 2 1/2 years of age. *Dev Med Child Neurol.* 2002;44:415–421.
49. Coolman RB, Bennet FC, Sells CJ, Swanson MW, Andrews MS, Robinson NM. Neuromotor development of graduates of the neonatal intensive care unit: patterns encountered in the first two years of life. *Dev Behav Pediatr.* 1985;6:327–333.
50. Janssen AJWM, Nijhuis-van der Sanden MWG, Akkermans RP, Tissingh J, Oostendorp RAB, Kollée LAA. A model to predict motor performance in preterm infants at 5 years. *Early Hum Dev.* 2009;85:599–604.
51. Janssen AJWM, Oostendorp RAB, Akkermans RP, Steiner K, Kollée LAA, Nijhuis-van der Sanden MWG. High variability of individual longitudinal motor performance over five years in very preterm infants. *Res Dev Disabil.* 2016;59:306–317.

52. Bastian AJ, Martin TA, Keating JG, Thach WT. Cerebellar ataxia: abnormal control of interaction torques across multiple joints. *J Neurophysiol.* 1996;76:492–509.
53. Synofzik M, Ilg W. Motor training in degenerative spinocerebellar disease: ataxia-specific improvements by intensive physiotherapy and exergames. *Biomed Res Int.* 2014;2014.
54. Vilis T, Hore J. Central neural mechanisms contributing to cerebellar tremor produced by limb perturbations. *J Neurophysiol.* 1980;43:279–291.
55. Morton SM, Bastian AJ. Mechanisms of cerebellar gait ataxia. Cerebellum. 2007;6:79–86.
56. Bastian AJ. Learning to predict the future: the cerebellum adapts feedforward movement control. *Curr Opin Neurobiol.* 2006;16:645–649.
57. Salman MS, Tsai P. The role of the pediatric cerebellum in motor functions, cognition, and behavior: a clinical perspective. *Neuroimaging Clin N Am.* 2016;26:317–329.

Appendix.

Observable Movement Quality (OMQ) scale terminology, definitions, and 5-point Likert scale with description of aspects

1. Appropriate fine motor movements

The child adapts its postures and movements to the demands of the fine motor tasks and the environment.

1. Consistently inappropriate
2. Typically inappropriate; one or two instances of appropriate fine motor movements
3. Inappropriate half of the time and appropriate other half of the time
4. Typically appropriate; one or two instances of inappropriate fine motor movements
5. Consistently appropriate

2. Appropriate gross motor movements

The child adapts its postures and movements to the demands of the gross motor tasks and the environment.

1. Consistently inappropriate
2. Typically inappropriate; one or two instances of appropriate gross motor movements
3. Inappropriate half of the time and appropriate other half of the time
4. Typically appropriate; one or two instances of inappropriate gross motor movements
5. Consistently appropriate

3. Fluency of movements

The movements of the child are controlled in such a manner that they are adapted to the tasks and the environment in a fluent manner, without faltering or stumbling.

1. Consistently not fluent
2. Typically not fluent; one or two instances of fluent movements
3. Not fluent half of the time and fluent other half of the time
4. Typically fluent; one or two instances of not fluent movements
5. Consistently fluent

4. Reduced muscle tone

The movements and/or maintenance of posture of the child give the impression of being slack and not adapted to the tasks and environment.

1. Consistently low muscle tone: like a rag doll
2. Typically low muscle tone; one or two instances without low[a] muscle tone
3. Low muscle tone half of the time and without lowa muscle tone other half of the time
4. Typically without low[a] muscle tone; one or two instances of low muscle tone
5. Absence of low muscle tone

5. Increased muscle tone

The movements and/or maintenance of posture give the impression of being stiff/rigid and not adapted to the tasks and the environment.

1. Consistently high muscle tone: muscles are rigid and tight
2. Typically high muscle tone; one or two instances without high[b] muscle tone
3. High muscle tone half of the time and without high[b] muscle tone other half of the time
4. Typically without high[b] muscle tone; one or two instances of high muscle tone
5. Absence of high muscle tone

6. Tremors

There is an involuntary, rhythmic, periodic, uncontrollable trembling of a body part or body parts during the child's movements, which can vary in amplitude from barely observable to clearly visible or in frequency from low to high.

1. Constantly
2. Frequently
3. Occasionally
4. Infrequently
5. None

7. Slow/delayed movements

The child moves the body or a part at a lower speed than is suitable for the task and can, despite instruction, not accelerate sufficiently.

1. Consistently slow and delayed movements
2. Typically slow and delayed; one or two instances of appropriate timing and pacing
3. Slow and delayed half of the time and appropriately timed and paced other half of the time
4. Typically appropriate timing and pacing; one or two instances of slow and delayed movements
5. Absence of slow and delayed movements

8. Accelerated/abrupt movements

The child moves the body or a part at a higher speed or abruptly at a higher speed than is suitable for the task and can, despite instruction, not slow down sufficiently.

1. Consistently accelerated and abrupt movements
2. Typically accelerated and abrupt; one or two instances of appropriate timing and pacing
3. Accelerated and abrupt half of the time and appropriately timed and paced other half of the time
4. Typically appropriate timing and pacing; one or two instances of accelerated and abrupt movements
5. Absence of accelerated and abrupt movements

9. Asymmetry in movements

In movements and/or maintenance of posture, a body half or part of a body half inadequately participates in the task. The difference in the use of body parts does not fit with the age of the child and with the demands of the tasks and the environment.

1. Consistently asymmetric
2. Typically asymmetric; one or two instances of symmetry
3. Asymmetric half of the time and symmetric other half of the time
4. Typically symmetric; one or two instances of asymmetry
5. Consistently symmetric

10. Accuracy (well-aimed)

The child moves the body parts in such a way that the target is reached accurately and immediately.

1. Target consistently not reached
2. Target typically not reached; one or two instances of target being reached
3. Target not reached half of the time and target reached other half of the time
4. Target typically reached; one or two instances of target not being reached
5. Target consistently reached

11. Strength regulation

The movements of the child are in terms of force/strength well suited to the task and the environment.

1. Strength consistently not adapted to tasks
2. Strength typically not adapted to tasks; one or two instances of strength adapted to tasks
3. Strength not adapted to tasks half of the time and strength adapted to tasks other half of the time
4. Strength typically adapted to tasks; one or two instances of strength not adapted to tasks
5. Strength consistently adapted to tasks

12. Variation in movements

The childc can move body parts relatively independently from each other in different directions, so the necessary degrees of freedom are used to match the demands of the tasks and the environment.

1. Consistently no variation in movements
2. Typically no variation in movements; 1 or 2 instances of variation in movements
3. No variation in movements half of the time; variation in movements half of the time
4. Typically variation in movements; 1 or 2 instances of no variation in movements
5. Consistently variation in movements

13. Involuntary movements

While moving, parts of the child's body and/or those parts not directly involved in the movements show unconscious movements[d] not appropriate to the child's age.

1. Consistently involuntary movements
2. Typically involuntary movements; one or two instances without involuntary[e] movements
3. Involuntary movements half of the time and without involuntary[e] movements other half of the time
4. Typically without involuntary[e] movements; one or two instances of involuntary movements
5. Absence of involuntary movements

14. Automated movements

The child has mastered the skills appropriate for the age in such a way that these are consistent and can be executed without much attention, and, if necessary, in combination with another task or tasks.

1. Movements consistently not automated
2. Typically movements not automated; one or two instances of automated movements
3. No automated movements half of the time and automated movements other half of the time
4. Typically automated movements; one or two instances of no automated movements
5. Consistently automated movements

15. Stereotyped movements

The child shows spontaneous, repetitive, purposeless movements (examples include repeatedly turning and shaking of the head and/or rocking or flutter of body parts).

1. Consistently stereotyped movements
2. Typically stereotyped movements; one or two instances without of stereotypedf movements
3. Stereotyped movements half of the time and without stereotypedf movements half of the time
4. Typically without stereotypedf movements; one or two instances of stereotyped movements
5. Absence of stereotyped movements

---

[a] In a previous version (2012)[16], "normal" was used instead of "without low." (Adapted in 2015.).

[b] In a previous version (2012)[16], "normal" was used instead of "without high." (Adapted in 2015.)

[c] Deleted from the description: "has an extensive repertoire of." (Adapted in 2015.)

[d] Deleted from the description: "that disrupt the proper execution of the task." (Adapted in 2015.)

[e] In a previous version (2012)[16], "normal" was used instead of "without involuntary." (Adapted in 2015.)

[f] In a previous version (2012)[16], "normal" was used instead

Chapter 3

Individual neurodevelopmental trajectories of children treated with hypothermia for perinatal asphyxia from 3 months to 5 years of age

Lieke MA Dekkers

Anjo JWM Janssen

Katerina Steiner

Nienke Maas-van Schaaijk

Reinier P Akkermans

Bert JM de Swart

Maria WG Nijhuis-van der Sanden

Revised version in preparation for: Research in Developmental Disabilities

Abstract

Background

Hypothermia for perinatal asphyxia is a common treatment to decrease morbidity. This study aims to describe a) individual neurodevelopmental trajectories over 5 years in children with perinatal asphyxia treated with hypothermia and b) the correlation between movement quality at 3 months and motor developmental outcomes at 5 years of age.

Methods

In this prospective longitudinal cohort study, 18 children (12 male) were assessed at 3 (t_1), 6 (t_2), 12 (t_3), and 24 (t_4) months, and at the age of 5 (t_5) years, with standardized norm-referenced tests.

Results

Six children showed abnormal movement quality assessed with General Movements (t1) and all showed severe neurodevelopmental disabilities at t_5. The 12 children without severe disabilities, showed a significant normalization of z-scores over the five assessment points (linear mixed model analysis). At t_5, four of these children scored mildly delayed motor or cognitive development.

Conclusion and implications

Children without anomalies on the MRI before hospital discharge and normal movement quality at 3 months of age showed normal neurodevelopment at the age of 5, however, individual motor trajectories showed variability over time. Presents of abnormal GMs tend to detect CP and developmental problems, advocating a developmental surveillance to determine need for early intervention.

Introduction

Perinatal asphyxia is the second-most important cause of morbidity for children in Europe.[1] The annual incidence for children suffering from perinatal asphyxia per 10,000 births was 6.9 for boys and 6.1 for girls in the Netherlands in 2012.[2] Children suffering from perinatal asphyxia have high risk for severe neurologic disabilities. Even in children with mild perinatal asphyxia, the risk for long-term developmental delays in motor, verbal, or intellectual performances is increased.[3–5]

Treatment with hypothermia after perinatal asphyxia has been shown to reduce mortality and to decrease major neurodevelopmental disabilities, such as cerebral palsy (CP), severe developmental delay, intellectual impairment, and blindness,[6,7] although a risk remains for neurological and developmental disabilities in treated children.[4,8] The decision to introduce hypothermia into the Netherlands and Flanders was based on the Cochrane Review in 2007,[9] which indicated neuroprotective effects of moderate hypothermia in full-term neonates with perinatal asphyxia.[10] In 2008, hypothermia was introduced in the Netherlands and Flanders,[10] with comparable outcomes to those of the large international trials (TOBY Trial,[11] CoolCap Trial Group,[12] China Study Group,[13] Neonatal Research Network,[14] NICHD Infant Cooling Evaluation Collaboration,[15] and the neo.nEURO.network[16]).

The reviews of Jacobs[6] and Edwards[7] discussed the effects of therapeutic hypothermia for perinatal asphyxia on mortality and long-term neurodevelopmental disabilities. The included trials present data on neurological outcomes after treatment with hypothermia for perinatal asphyxia after assessment of, at least, 18 months of age, using the primary outcomes: death or major neurodevelopmental disabilities. The homogeneity of trial results was high and provides clear evidence of the therapeutic benefit of hypothermia treatment.[6,7] However, results on individual longitudinal neurodevelopmental trajectories after hypothermia treatment for children suffering from perinatal asphyxia have not yet been published.

Neurodevelopment is a complex process with a continuous interaction between genetically based and environmentally driven processes.[17,18] In this development, the quality of spontaneous activity is a fundamental characteristic of the development of neuronal networks.[19,20] In the young child, the observation of the quality of spontaneous movements, especially the quality of general movements (GMs), accurately provides information on the condition of the young nervous system and is a marker of neurological dysfunction.[21,22] In high-risk children, the assessment of GMs at 3–4 months of age—in the so called fidgety GM period—is the most sensitive and specific test to allow early prediction of spastic CP.[23–25] After 3–4 months of age, GMs disappear,[26] and movement quality can be assessed only by observing spontaneous or elicited skill performances and goal-directed movements.[27]

The first aim of this study was to describe the individual neurodevelopmental trajectories over 5 years in term children treated with hypothermia for perinatal asphyxia and followed in a standardized multidisciplinary follow-up programme of five appointments. The second

aim was to examine the correlation between movement quality at 3 months of age and motor development at 5 years of age.

Methods

This was a prospective longitudinal cohort study in which children with perinatal asphyxia who were treated with hypothermia at the Department of Neonatology, Radboudumc, Nijmegen, the Netherlands, were all followed in a standardized multidisciplinary follow-up programme. The children were assessed at 3 (t_1), 6 (t_2), 12 (t_3), and 24 (t_4) months and 5 years (t_5) of age. The medical ethical committee agreed that approval was not required (file number: 2017-3344) because the protocol is part of accepted medical practice and conforms to the principles of the Declaration of Helsinki.

Treatment protocol hypothermia

The treatment protocol was the same as described by Groenendaal.[10,11] Inclusion criteria for hypothermia treatment were a gestational age of at least 36 weeks and a clear presence of perinatal asphyxia followed by hypoxic ischemic encephalopathy (HIE). Cut-off for therapeutic hypothermia was a Thompson score of >7, between 1 and 3 hours after birth, indicating moderate to severe encephalopathy[28] and the possibility of starting hypothermia treatment within 6 hours of birth. The presence of congenital malformations were not an exclusion criterion for hypothermia. Total-body hypothermia was applied using CritiCool® (the Surgical Company, Amersfoort, the Netherlands). Total duration of therapy was 72 hours; depth of hypothermia was set at a core temperature of 33.5°C, using a rectal temperature probe. As the protocol strongly advised, morphine was used to reduce stress.[29] A cranial ultra sound was performed daily, and Magnetic Resonance Imaging (MRI)—using diffusion-weighted imaging—was performed between 4 and 8 days after birth.

Participants

This study included surviving children born between January 2009 and December 2010 suffering from HIE after perinatal asphyxia admitted to the Radboudumc Neonatal Intensive Care Unit (NICU) for hypothermia treatment.

Baseline characteristics were collected (t_0): gestational age, sex, birth weight, Apgar score at 5 minutes, and Thompson score; MRI, performed between 4 and 8 days after birth.

Parents and/or caregivers were informed about the 5-year follow-up programme as part of the standard care, following national guidelines and data collection for research during NICU admission, and refusal to participate or withdraw without given reason was always possible. At least one of the parents and/or caregivers was present throughout the test procedures, except for the psychological tests at 5 years.

Follow-up assessment protocol

During single outpatient clinical visits, children were assessed at t_1, t_2, and t_3 by a paediatrician, a paediatric physiotherapist (PPT), and a speech and language therapist (SLT). A child psychologist was added to the team at t_4 and t_5. All assessments were performed using the standardised procedures for administration and instructions for calculation of the test scores—as specified in the respective test manuals—and were conducted by experienced examiners as part of multidisciplinary outpatient evaluations.

Paediatrician's assessment

The paediatrician performed a physical examination at all assessment time points, including a structured neurological examination as standardised at Radboudumc; including examination of reflexes and muscle tone, and the registration and classification of deviant motor behaviour.

Motor performance assessment

The PPT assessed motor performance by an age-appropriate test: at t_1, t_2, t_3, and t_4, the Motor Scales of the Bayley Scales of Infant and Toddler Development, 3rd edition, Dutch version (BSID-III-NL)[30–32], and at t_5, the Movement Assessment Battery for Children, 2nd edition, Dutch version (MABC-2-NL).[33,34] The Gross Motor Function Classification Scale (GMFCS)[35,36] was used to classify the subgroup of children with spasticity at t_3, t_4, or t_5. The PPT registered physiotherapy treatment (yes/no and frequency).

Movement quality assessment

The PPT assessed the quality of movement using the assessment of fidgety GMs[20] at t_1, and the Observable Movement Quality (OMQ) scale[27] at all follow-up assessments. To classify the quality of the fidgety GMs, the complexity, variation and fluency of spontaneous movements were classified—as described by Hadders-Algra[20]—using four classes; two forms of normal GMs, normal-optimal and normal-suboptimal GMs; and two forms of abnormal GMs, mildly and definitely abnormal GMs. All involved PPTs conducted a two-day training in GMs assessments.

The OMQ scale is a new generic evaluative measurement instrument, designed to assess movement quality in children from 3 months to 16 years of age.[27] In the OMQ scale are 15 items on movement quality included, which can be filled in by a paediatric physiotherapist after the assessment of an age-specific, discriminative or disease specific motor test. Using the OMQ scale, movement quality is observed and scored relative to what is expected for a child's age. Each OMQ item focuses on an element of observable movement quality and is scored on a 5-point Likert scale; total scores range from 15 to 75. Lower scores indicate lower movement quality.[27]

Language and communication assessment

A SLT assessed speech production and speech development[37] at t_1 and t_2. The Dutch Communicative Development Inventory (N-CDI)[38] was administered at t_3. The Reynell test for language comprehension[39] and the Schlichting test for language performance[40] were used at t_4 and t_5.

Cognitive and behavioural assessment

Cognitive development was assessed by a psychologist (assistant), using the cognitive scale of the BSID-III-NL[30] at t_4. At t_5, cognition was assessed with the short version of the revised Amsterdam Intelligence Test (RAKIT)[41] and the Beery-Buktenica Developmental Test of Visual-Motor Integration, 6th edition (Beery™ VMI).[42,43] Behavioural outcomes were assessed at t_4 and t_5, using the Dutch version of the Child Behavior Checklist (CBCL)[44] for children ages 6 to 18 years as reported by both parents.

Statistical analysis

Descriptive analyses were performed to describe outcome variables. Based on distribution of the variables, median and interquartile ranges (IQR) were used.

To compare outcomes of different tests, *z*-scores were converted from the scores of BSID-III-NL Motor and cognitive scales, MABC-2-NL, and RAKIT[45] *Z*-scores less than or equal to (≤) −2, indicating 'significantly delayed'; scores less than or equal to (≤) −1 but greater than (>) −2, indicating 'mildly delayed'; scores greater than (>) −1 but less than or equal to (≤) 1, indicating 'within normal limits'; and scores greater than (>) 1, indicating 'accelerated' performances.

An LMM analysis for repeated measures was used to test longitudinal mean differences in motor outcome from 3 months to 5 years of age. In these analyses, we used a random intercept model with a fixed factor time (i.e., assessment time point [five levels] and a random factor child [see Table 2 for the number of included children]). The dependent variable was motor performance (*z*-scores of motor tests calculated from BSID-III-NL and the standard score of the MABC-2-NL).

To examine the correlation between movement quality and motor development, concurrent validity and predictive validity between the OMQ scale scores and the GMs outcomes in two categories were analysed with the Spearman's rank correlation coefficient (r_s)[45]. The categories, severe and mild abnormal GMs, were combined and used as the main outcome or reference standard (1 = 'definite abnormal' or 'mildly abnormal'; 2 = 'normal/subnormal'or 'normal/optimal'). For the OMQ scale, a cut-off point of 65 points on the total score was used. Outcomes at 5 years were dichotomized in 'normal' and 'delayed' motor development. Delayed development was indicated by MABC-2-NL total *z*-score and/or RAKIT score ≤ −2 and/or diagnosed CP. The outcome of the MABC-2-NL was the dependent variable, and cross-tabulations were generated. Correlations were considered as follows: 0.00–0.25, little

or no relationship; 0.25–0.50, fair relationship; 0.50–0.75, moderate to good relationship; above 0.75, good to excellent relationship.[45]
Statistical analyses were performed in IBM Statistical Package for the Social Sciences (IBM SPSS Statistics), version 22 (IBM Corporation, Armonk, NY). All statistical tests were two-tailed, and $p < 0.05$ was considered significant.

Results

A total of 23 children were eligible for hypothermia treatment; the surviving 18 children were included in the follow-up, of which 12 (67%) were male. Figure 1 shows a flow chart of the included children. Eleven out of 18 children participated in all five follow-up assessments, one child in three, and one child in two appointments. Four children (22%) were not scheduled for appointments at t_{4-5} due to participation in rehabilitation programmes for children with severe disabilities elsewhere. One child was lost to follow-up after the first appointment. In total, 75 out of 90 assessments were completed.

Characteristics and baseline data (t_0) for the included children are shown in Table 1. MRI was performed between 4 and 8 days after birth in 16 children (median: 6 days). In one child, MRI was performed at 3 and 14 days after birth due to very deviant cranial ultra sound images; while MRI with another child was not possible before the age of 14 days because of Extracorporeal Membrane Oxygenation treatment. MRI assessment showed no anomalies in 14 children (78%) and anomalies in basal ganglia and thalami (BGT) or BGT with anomalies of posterior limb of internal capsule (PLIC) in 3 children (17%). In one child, MRI showed parieto occipital ischemia. Four children (22%) were diagnosed with medical conditions shortly after birth, including two children with an obstetric plexus brachialis lesion and another two with congenital anomalies (i.e., cheiloschisis and hydronephrosis). During the follow-up period, two children (11%) were diagnosed with a syndrome (Greig cephalopolysyndactyly syndrome and, 14q32.31q32.33 deletion), and one child with a metabolic disorder (Acyl-Coenzyme A oxidase-1 deficit [ACOX1]).
For follow-up assessments, the mean age (Standard deviation [SD], range) at t_1 was 3.4 months (SD 0.5, 2.3−4.1); at t_2, 6.6 months (SD 0.6, 5.6−8.2); at t_3, 13 months (SD 0.9, 11.53−14.7); at t_4, 25.7 months (SD ±1.1, 24.3−27.8); and at t_5, 5 years and 5 months (SD 5.1, 57.4−79.18).

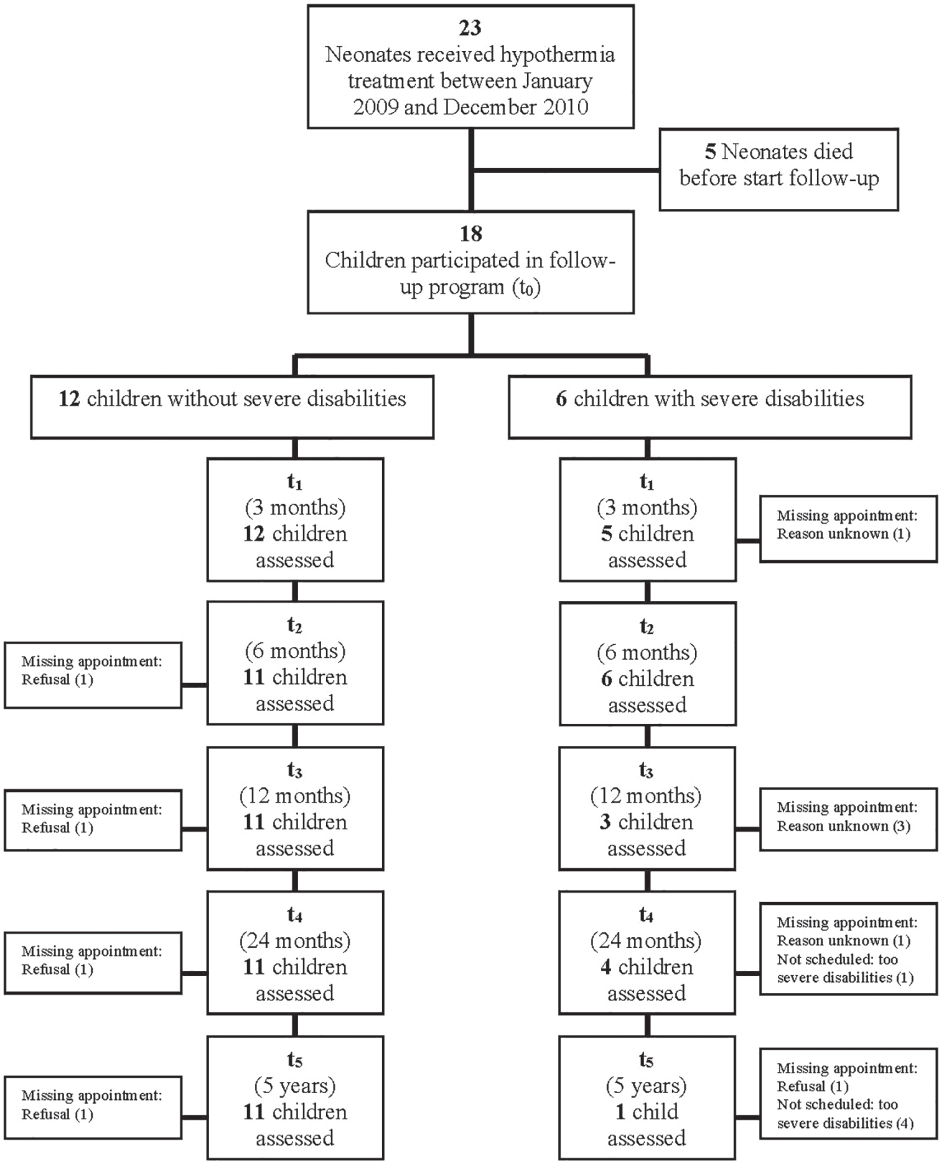


Figure 1. Flow chart of children included in the follow-up programme for children treated with hypothermia for perinatal asphyxia.

Table 1. Baseline data at birth (t_0) and medical condition for children included in study

Pt	Sex	GA	BW	TS	AS 5'	MRI	Medical condition
1	M	40	3345	10	1	No Anomalies	Obstetric plexus brachialis lesion
2	M	38	3750	10	4	No Anomalies	Obstetric plexus brachialis lesion
3	M	41	4200	8	1	No Anomalies	Cheiloschisis, Autism
4	M	40	3500	11	0	BGT/IC	CP (GMFCS level V)
5	M	42	2800	11	2	No Anomalies	Syndrome: 14q32.31q32.33 deletion
6	M	41	4400	< 7	5	No Anomalies	
7	M	41	3196	8	3	No Anomalies	Metabolic disorder: ACOX1 deficit
8	M	39	4400	< 7	5	No Anomalies	
9	M	40	2690	9	7	No Anomalies	
10	M	39	2670	12	5	No Anomalies	Autism, Hearing loss
11	M	41	4100	11	6	No Anomalies	
12	M	40	3780	9	5	Stroke*	
13	F	42	4100	11	5	No Anomalies	Hydronephrosis
14	F	38	2700	-	10#	No Anomalies	
15	F	39	2650	8	0	No Anomalies	
16	F	36	2100	13	4	No Anomalies	
17	F	41	4140	14	1	BGT	CP (GMFCS level V) and Greig cephalopolysyndactyly-syndrome
18	F	41	3100	12	1	BGT	CP (GMFCS level I) and epilepsy
Mean (SD)		40.3 (1.4)	3423 (725)	10 (3)	4 (4)		

Pt: patient, M: male; F: female, GA: gestational age in weeks, BW: birth weight in grams, TS: Thompson score, AS 5': Apgar scores at 5 minutes, #: incident with reanimation at 20 minutes postpartum, BGT: Basal Ganglia and thalami anomalies, IC: anomalies of internal capsule (PLIC), *parieto occipital ischemia, CP: Cerebral Palsy, GMFCS: Gross Motor Function Classification System, ACOX1: Acyl-Coenzyme A oxidase-1 deficit

Paediatrician's assessment

Sixteen children (89%) showed minor neurological symptoms at, at least, one of the assessments. Two children (11%) were diagnosed with CP at the age of 2 years: one with bilateral spastic CP with axial hypotonia and the other with dyskinetic CP. One child was diagnosed with CP with unilateral paresis and epilepsy at age 5. All children diagnosed with CP showed anomalies on the MRI before hospital discharge, as is shown in Table 1. The exception was the child with the ischemia on the MRI, which showed normal development. Furthermore, two children (11%) were diagnosed with autism: one child at age 2 and the other at age 3. One child was diagnosed with severe hearing loss at 3 years of age.

Table 2. Results of developmental outcomes (z-scores) from 3 months to 5 years of age.

		t_1				t_2			t_3			t_4				t_5				
Pt	**GMs**	**OMQ**	**FM**	**GM**	**MOT**	**FM**	**GM**	**MOT**	**FM**	**GM**	**MOT**	**FM**	**GM**	**MOT**	**COG**	**MAN**	**AIM**	**BAL**	**MOT**	**COG**
Children without severe disabilities																				
1	SO	65	- 2.33	- 1.33	- 1.40	0.33	- 0.67	- 0.13	- 0.33	- 0.33	- 0.33	0.33	- 0.67	- 0.13	1.00	- 0.33	- 1.67	- 0.67	-1.33	- 1.33
2	NA	NA	NA	NA	NA	0.33	- 0.33	0.07	- 1.00	- 0.67	- 0.87	0.33	0.67	0.60	- 0.33	- 1.00	- 1.00	- 0.67	-1.33	- 0.60
3	SO	65	- 1.67	- 0.33	- 1.07	0.33	0.33	0.47	- 0.67	- 0.33	- 0.87	- 0.33	- 0.33	- 0.33	0.00	- 0.33	- 1.33	- 0.67	-1.00	0.73
6	NA	70	- 1.00	- 1.67	- 1.47	0.67	- 0.33	0.27	0.00	- 0.33	- 0.13	0.33	- 0.33	0.70	0.33	0.33	0.00	1.33	1.00	1.40
8	SO	62	- 1.67	0.00	- 0.87	1.33	- 1.00	0.27	0.00	- 0.67	- 0.33	0.00	0.33	0.27	0.33	1.00	0.00	- 0.33	0.67	0.67
9	SO	63	1.00	0.33	0.80	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
11	SO	68	- 1.00	- 0.33	- 0.73	1.00	0.00	0.60	- 0.33	0.00	- 0.13	0.33	- 0.67	- 0.13	0.33	1.00	0.67	- 0.67	0.67	1.33
12	NO	66	- 1.33	- 1.00	- 1.27	- 0.67	0.00	- 0.33	- 1.00	0.67	- 0.13	- 1.33	- 0.33	- 0.87	-	- 1.00	0.00	- 0.33	- 0.67	0.00
13	SO	65	- 2.67	- 0.33	- 1.67	0.67	0.00	0.47	- 0.33	- 0.33	- 0.33	1.67	- 0.33	0.80	0.00	2.00	- 1.00	0.33	0.67	0.73
14	SO	62	0.00	- 2.00	- 0.33	- 1.00	- 1.00	- 1.07	- 0.67	- 2.00	- 1.67	0.67	- 1.00	- 0.13	0.67	0.33	0.00	0.67	0.67	1.60
15	NO	75	1.00	- 0.33	0.47	1.00	- 0.67	0.27	- 1.00	- 1.33	- 1.27	0.33	- 0.33	0.70	-	- 0.67	- 0.33	- 0.67	- 0.67	- 0.60
16	NO	-	- 0.33	0.00	- 0.13	1.67	0.33	1.20	- 0.33	- 2.67	- 1.67	1.00	- 1.00	0.70	- 0.33	0.33	- 1.33	0.67	0.00	0.13
Mean (SD)		66.1 (2.51)	- 0.91 (1.25)	- 0.64 (0.76)	- 0.70 (0.80)	0.51 (0.79)	- 0.30 (0.48)	0.19 (0.58)	- 0.51 (0.38)	- 0.72 (0.94)	- 0.70 (0.60)	0.30 (0.75)	- 0.36 (0.51)	0.20 (0.55)	0.20 (0.44)	0.15 (0.94)	- 0.54 (0.75)	- 0.09 (0.72)	- 0.12 (0.90)	0.37 (0.93)
Children with severe disabilities																				
4	DA	48	- 2.67	- 2.33	- 3.07	- 2.67	- 3.00	- 3.20	-	-	-	<-2.0	<-2.0	-	-	-	-	-	-	-
5	MA	46	- 0.33	- 2.67	- 1.67	- 2.00	- 1.00	- 2.40	- 1.33	- 2.67	- 2.20	-	-	-	-	-	-	-	-	-
7	MA	40	- 3.00	- 2.67	- 3.20	- 1.67	- 3.00	- 2.60	- 3.00	- 3.00	- 3.40	<-2.0	<-2.0	<-2.0	-	-	-	-	-	-
10	MA	59	- 2.00	- 2.00	- 2.20	- 0.67	- 0.33	- 0.53	-	-	-	- 2.33	- 2.00		-	-	-	-	-	-
17	MA	46	- 3.00	- 2.67	- 3.20	- 2.00	- 3.00	- 2.80	- 3.00	- 3.00	- 3.40	<-2.0	<-2.0	<-2.0	-	-	-	-	-	-
18	MA*	-	-	-		1.00	- 0.33	0.47	-	-	-	-	-	-	-	- 2.67	- 0.67	- 2.00	- 3.17	0.53
Mean (SD)		47.8 (6.94)	- 2.20 (1.12)	- 2.47 (0.30)	- 2.67 (0.70)	- 1.34 (1.32)	- 1.78 (1.36)	- 1.84 (1.46)	- 2.44 (0.96)	- 2.89 (0.19)	- 3.00 (0.69)									

t_1: follow-up at 3 months, t_2: follow-up at 6 months, t_3: follow-up at 12 months, t_4: follow-up at 24 months, t_5: follow-up at 5 years of age, Pt: patient, GMs: General Movements assessment, GM-classification: NO, normal-optimal; SO, normal-suboptimal; MA, mildly abnormal; DA definitely abnormal, *assessed at hospital discharge, OMQ: OMQ scale total score, NA: Not Assessed due to state, FM: fine motor skills on BSID-III-NL, GM: gross motor skills on BSID-III-NL, MOT: total motor score at BSID-III-NL(t_{1-4}) or M-ABC-2-NL (t_5), COG: cognitive development at BSID-III-NL (t_4) or RAKIT (t_5), MAN: Manual Dexterity subscale on M-ABC-2-NL, AIM: Aiming and Catching subscale on M-ABC-2-NL, BAL: Balance subscale on M-ABC-2-NL.

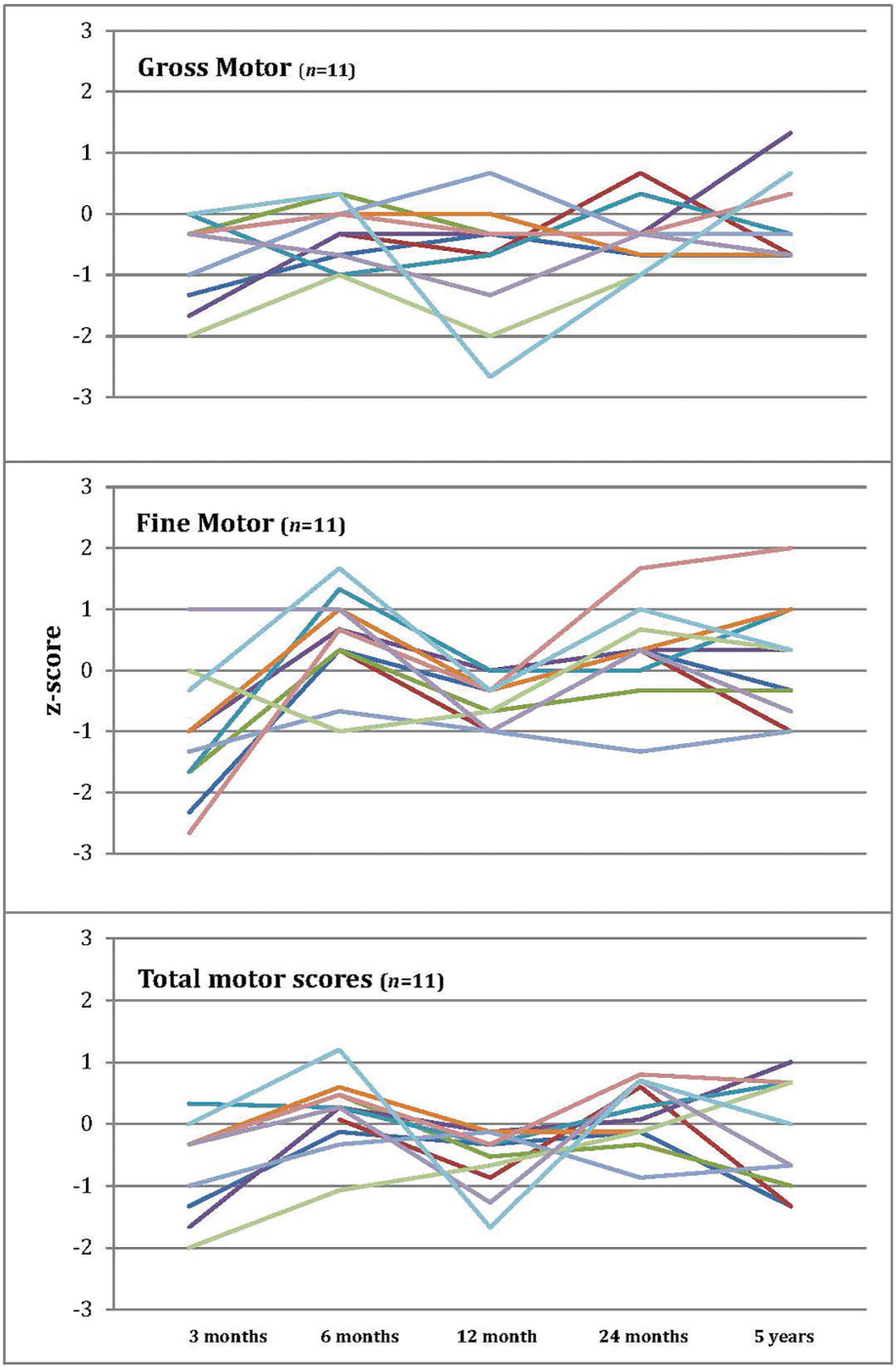


Figure 2. Individual *z*-scores over the follow-up assessments for children without severe developmental delay. Assessment at 3, 6, 12, and 24 months using BSID-III-NL, at 5 years using MABC-2-NL.

Analysis of neurodevelopmental outcomes

Multiple assessments were missing from children with severe disabilities who had not been scheduled for appointments at for t_{4-5} or at t_5 and for the one child who was diagnosed with CP during follow-up assessment at 5 years of age. Since outcomes for these children were not available, we made two groups: one of children 'without severe disabilities' and one of children 'with severe disabilities' (see Table 2), which enabled us to perform reliable analyses of neurodevelopmental trajectories.

Motor performance assessment

Table 2 shows the longitudinal results for motor performance outcomes for each child at all follow-up assessments. Individual longitudinal motor performance trajectories for the children without severe disabilities are shown in figure 2. The individual trajectories for these children show variability in the individual scores over time.

The LMM analyses (fixed factor time; random child) of the mean *z*-scores of the group without severe disabilities of the BSID-III-NL motor scores revealed a significant difference between five time points (t_1-t_5: F = 5.279, p = 0.002). At 3 months, the mean BSID-III-NL motor *z*-score was –0.70 (SD = 0.81), rising to a mean *z*-score of –0.12 (SD = 0.90) on the MABC-2-NL at 5 years of age.

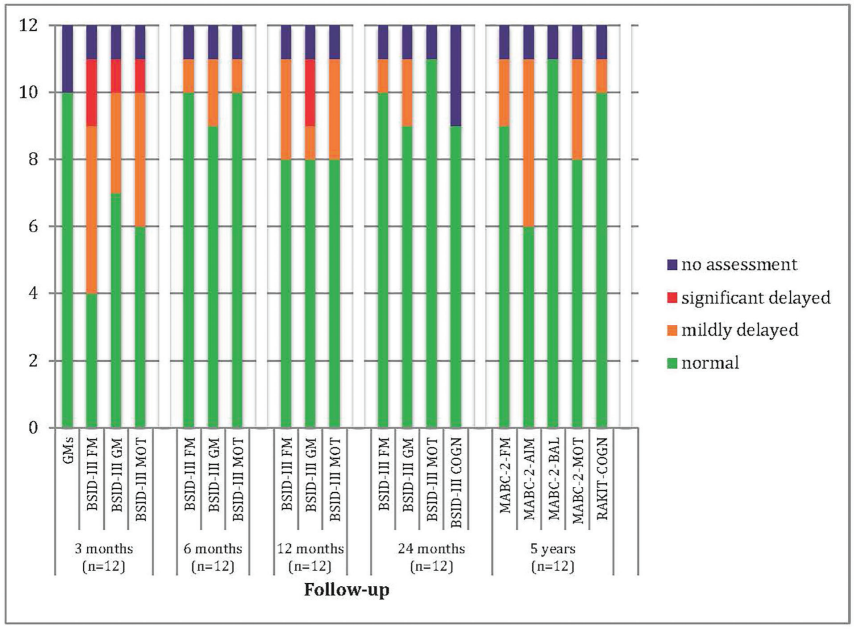


Figure 3. Cross-sectional *z*-scores for children without severe developmental delay. Assessment of general movements (GMs), Bayley Scales of Infant and Toddler Development (BSID-III); fine motor scale (FM), gross motor scale (GM), total motor scores (MOT) and cognition (COGN), Movement Assessment Battery for children: subscales manual dexterity (FM), aiming and caching (AIM), balance (BAL) and total scores (MOT), and the revised Amsterdam intelligence test (RAKIT-COGN)

Figure 3 presents an overview of cross-sectional outcomes for motor performance at all five follow-up moments for 12 children without severe disabilities. The number of children scoring a delayed motor performance from 3 months to 5 years of age decreased from five (42%) to three (25%).

All 18 included children received physiotherapy in the first year of life. At age 5, no children within the group without severe disabilities received physiotherapy; however, one child (6%) was recommended to start physiotherapy after follow-up assessment. All five children participating in rehabilitation programmes for children with severe disabilities or diagnosed with CP received physiotherapy at the age of 5.

Movement quality assessment

At t_1, 16 (89%) children's GMs were successfully assessed, of whom 10 (56%) scored within normal limits (see Table 2 for the GM outcomes). All six children (33%) with abnormal scores on GMs at t_1 showed significant motor developmental delays at t_5. An overview for median scores on the OMQ scale and total scores on the BSID-III-NL for each category of GMs is presented in Table 3.

Table 3. Overview mean score OMQ-scale and BSID-III for children categorized based on GMs at 3 months (t_1) (n = 15).

GMs category	*n*	Median (IQR) total score on OMQ-scale	Median (IQR) total score on BSID-III-NL
Normal/optimal	3	71 (9)	98 (26)
Normal/suboptimal	7	65 (3)	87 (20)
Mildly abnormal	4	46 (9)	60 (19)
Definitive abnormal	1	48*	58*

IQR: Inter Quartile Range, GMs: general movements, OMQ scale: Observable Movement Quality Scale, BSID-III-NL: Bayley Scales of Infant and Toddler Development, 3th edition, Dutch version, *: no IQR could be calculated (n = 1).

Language and communication assessment

All children scored within normal limits at t_1 and t_2 for speech production and speech development; the same scores were seen at t_3 on the N-CDI. At t_{4-5}, all children scored within normal limits for language, syntactic, and lexical development on the Reynell test for language comprehension and the Schlichting test for language performance.

Cognitive and behavioural assessment

At t_4, the cognitive scale on the BSID-III-NL was successfully assessed in nine children from the group without severe disabilities, as shown in figure 2. Assessment in one child was not possible due to behavioural problems. For another child, results are missing; the assessment

was mistakenly performed with a previous version of the BSID caused by accidental enrolment in another follow-up protocol. All children assessed scored within normal limits. At t_5, all children from the group without severe disabilities scored within normal limits on integration of visual and motor abilities on the Beery™ VMI. One of the 11 assessed children scored 'mildly delayed' on the intelligence quotient on the RAKIT. For behaviour rating scales on the CBCL at t_4, eleven scales were returned, showing scores within normal limits for all children. At t_5, 10 scales were returned, showing total scores 'within clinical range'—i.e. deviant, or problematic behaviour—for one child on the mother's checklist, with the emphasis on internalizing problems, whereas another child scored within clinical range on the father's checklist, both on internalizing and externalizing problems.

Correlation between movement quality and motor development

Table 4 shows the correlation between the OMQ scale and GM outcomes at t_1 (r_s = 0.65), which was moderate to good and significant. Also, the correlation between the MABC-2-NL at t_5 and OMQ scale total scores at t_1 (r_s = 0.75) was moderate to good and significant. The correlation between the MABC-2-NL total outcomes and GMs (r_s = 0.84) was good to excellent and significant. Note that the number of included children (n = 18) was small; outcomes for concurrent and predictive validity should be interpreted as a tendency.

Table 4. Concurrent and predictive validity of OMQ-scale and assessment of GMs (t_1).

Concurrent validity of OMQ scale and GMs at 3 months of age				
Concurrent validity		**GMs abnormal**		
		Spearman *r*	*p* value	*BCa* 95% CI
	OMQ < 65	0.65	0.01	(0.4–1.0)
Predictive validity between dichotomized quality of movements at 3 months of age and dichotomized motor outcome at 5 years of age on MABC-2-NL				
Predictive validity		**MABC-2-NL z-score ≤ -1 SD**		
		Spearman *r*	*p* value	*BCa* 95% CI
	GMs abnormal	0.84	< 0.01	(0.4–1.0)
	OMQ < 65	0.75	< 0.01	(0.5–1.0)

OMQ scale: Observable Movement Quality Scale, GMs: General Movements, MABC-2-NL: Movement Assessment Battery for Children, second Dutch version.

Discussion

This study described motor performance trajectories over 5 years for children treated with hypothermia after perinatal asphyxia of a single centre in the Netherlands. From the 18 included children, 12 children showed normal or mildly delayed development at age 5. Children with severe disabilities due to CP, syndromes, or metabolic disorders were correctly estimated with GM assessment and the OMQ scale, while MRI only correctly detected children with severe disabilities due to CP.

Research in groups with low incidence is difficult to perform, especially in children, where you need to control for developmental changes. We tried to assess the included children as often as feasible and succeeded in assessing them five times in the first five years of their lives. Hereby, we obtained the highest quantity of measurements allowing us to describe motor trajectories for children treated with hypothermia.

Hypothermia treatment was offered to all children meeting the inclusion criteria within six hours after birth. The clinical decision postnatal to start treatment depends on the described inclusion criteria for hypothermia; however, some possible underlying diseases or syndromes were not yet diagnosed at the start of hypothermia. This resulted in the inclusion of one child later diagnosed with a syndrome, which does not result in spasticity—therefore we concluded that CP actually was caused by asphyxia—one child with a syndrome, and one child with a metabolic disorder. This may have caused the high number of children with severe neurodevelopmental disabilities at 5 years of age. In our study, 33% of the children had severe neurodevelopmental disabilities; the percentages in other studies ranged from 19 to 27.[11,12,46] However, the percentage of those diagnosed with CP (27%) is comparable with those presented in the TOBY study.[11]

Results on the MRI can be used for prediction of outcome in children treated with hypothermia for perinatal asphyxia.[47] Furthermore, Novak et. al.[24] reported that GMs—at fidgety movements' period—plus neonatal MRI accurately diagnoses CP in more than 95% of the children. In our study, four children showed anomalies on the MRI before hospital discharge, BGT in two children, and PLIC in one child. Furthermore, MRI showed parieto occipital ischemia in one child. All children with anomalies on the MRI showed abnormal GMs, and developed CP, except for the child with the ischemia; this child scored normal GMs and 'within normal limits' onward from the second follow-up assessment. All children with a normal MRI showed normal neurodevelopmental trajectories. The six children in our study with severe neurodevelopmental disabilities at 5 years of age all showed an abnormal quality of GMs at 3 months of age. Our study showed a moderate to good and significant correlation between the assessments of GMs and the assessment of quality of voluntary movements, scored on the OMQ scale. The correlation between MABC-2-NL total outcomes and GMs was high and significant, as was the correlation between MABC-2-NL and OMQ scale total scores. These outcomes indicate that the OMQ scale could be used as an alternative for GMs

in children at risk for developmental delays; however, more research is needed with larger groups of children and with different diagnoses before final conclusions can be drawn.

Motor developmental trajectories for children without severe disabilities show results within normal limits for most of the assessments, however, changes of *z*-scores occur over time in the individual children. At 3 months, two children scored 'significantly delayed' on fine motor assessments. One of these children was diagnosed with a plexus brachialis lesion shortly after birth, and MRI showed ischemia in the other child. Both children caught up from their poor performances and scored within normal limits at the remaining follow-up assessments. Another child scored significantly delayed at 12 months of age on gross motor performances. This child used bottom shuffling as an alternative form of locomotion. At the next follow-up assessment, this child had learned how to walk independently, resulting in motor developmental scores within normal limits. Also, for the remaining children, *z*-scores show variability between follow-up assessments and reveals a necessity for repeated assessments over time; one assessment alone could misinform individualised intervention processes.

From the individual trajectories for motor development, it can be seen that for all included children without severe disabilities, *z*-scores were reported below the standardized mean (0 SD) at 12 months of age on the BSID-III-NL Total Motor Scale; however, these scores are within the normal range of 1 to –1 SD. It is known is that there are important differences in functioning and developmental levels of children in the Netherlands and the USA, which caused the BSID-III to be adapted for the Dutch population.[32] Despite using the Dutch reference norms, all scores were below the standardized mean; apparently, learning how to walk independently is a critical milestone that appears at an older age for this particular group of children. However, children who scored within normal limits at 3 months of age also scored within normal limits at 5 years of age. For the Fine Motor scale, almost all children without severe disabilities show a decrease in *z*-scores from 6 to 12 months followed by normalization for most children at 24 months of age. A similar trend was seen in the presented reference values for BSID-III-NL fine motor scale showing the average Dutch scaled scores in relation to age.[32] This could indicate a too strict standardization of the BSID-III-NL Fine Motor Scale, possible due to the choice for constructing Dutch norms using weighted samples based on age in days;[48] while for US norms, age groups varying from two weeks to three months were used.[30] The argumentation for this choice was based on the opinion that age in days seems to most precisely reflect the development of young children.[32,48] However, more research is needed on the relation between the swing in scores and norms, using longitudinal data and individual developmental trajectories, before a conclusion can be drawn.

On individual trajectories for language and communication and behaviour, all children scored within normal limits, except one child who scored within clinical range at the age of 5 years. For cognition, only one child without severe motor disabilities scored mildly delayed,

which is comparable to outcomes for the CoolCap follow-up study[46] in which 7- to 8-year-old children were assessed on neurodevelopmental outcomes.

Our study was a single-centre study with a small sample size, which calls for caution in interpretation of the results. Due to the nature of our study, in which all patients treated at Radboudumc were included, the number of patients could not be expanded. Other limitations were the impossibility for blinding the developmental assessments and the circumstances that children with severe impairments were not scheduled at 24 months and 5 years of age for follow-up because of participation elsewhere in rehabilitation programmes. This resulted in the absence of data on developmental trajectories over the last 3 years in these children. Furthermore, all children received physiotherapy treatment at some point in the first 5 years of their lives. However, it is not clear what physiotherapy treatment enhanced and how or if it influenced the motor performance trajectories for the individual child.

Conclusion

Our study is providing additional data about neurodevelopmental outcomes for children treated with hypothermia for perinatal asphyxia. With this limited data set, we demonstrated that children without anomalies on MRI before hospital discharge and with normal scores on GMs at 3 months of age show normal neurodevelopmental trajectories. The presence of anomalies on MRI tends to estimate CP, and the presence of abnormal scores on GMs tends to indicate both the presence of CP and developmental problems due to syndromes or metabolic disorders. However, the variability in intra-individual motor trajectories advocates a developmental surveillance to determine the need for early intervention instead of decisions on single-point assessments, especially for those children with anomalies on MRI and abnormal scores on quality of movement assessments.

Ethics approval and consent to participate

Medical ethical committee approval was not required, as was confirmed on behalf of the research ethics committee "Commissie Mensgebonden Onderzoek Regio Arnhem-Nijmegen, The Netherlands" (file number: 2017-334); since the protocol is part of accepted medical practice and is performed conform the principles of the Declaration of Helsinki. The medical ethical committee stated that the study could be carried out (in the Netherlands) without approval by an accredited research ethics committee and without explicit written informed consent of the participants.

Funding details

The Dutch Research Counc (NWO) supported Lieke Dekkers: grant 023.004.037. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Disclosure of interest

The authors report no conflict of interest.

Data availability statement

The data that support the findings of this study are available from the corresponding author, [LD], upon reasonable request.

References

1. Black RE, Cousens S, Johnson HL, et al. Global, regional, and national causes of child mortality in 2008: a systematic analysis. *Lancet*. 2010;375(9730):1969-1987.
2. StatLine. Ziekenhuisopnames: geslacht,leeftijd en diagnose-indeling ICD9, 768 Intra-uteriene hypoxie en asfyxie. Centraal Bureau voor de Statistiek.
3. Perez A, Ritter S, Brotschi B, et al. Longterm neurodevelopmental outcome with hypoxic-ischemic encephalopathy. *J Pediatr*. 2013;163(2):454-459.
4. Douglas-Escobar M, Weiss M. Hypoxic-Ischemic Encephalopathy - A review for the clinician. *JAMA Pediatr*. 2015;169(4):397-403.
5. Pappas A, Shankaran S, Mcdonald SA, et al. Congnitive outcomes after neonatal encephalopathy. *Pediatrics*. 2015;135(3):e264-e634.
6. Jacobs SE, Berg M, Hunt R, Wo T, Inder TE, Davis PG. Cooling for newborns with hypoxic ischaemic encephalopathy (Review). *Cochrane Libr*. 2013;(3):1-112.
7. Edwards AD, Brocklehust P, Gunn AJ, et al. Neurological outcomes at 18 months of age after moderate encephalopathy : synthesis and meta-analysis of trial data. *BMJ*. 2010;340:c363.
8. Shankaran S, Pappas A, McDonald S, et al. Childhood outcomes after hypothermia for neonatal encephalopathy. *N Engl J Med*. 2012;366(22):2085-2092.
9. Jacobs SE, Hunt R, Tarnow-Mordi WO, Inder TE, Davis PG. Cooling for newborn with hypoxic ischemic encephalopathy. *Cochrane Database Syst Rev*. 2007;(4):CD003311.
10. Groenendaal F, Casaer A, Kijkman KP, et al. Introduction of hypothermia for neonates with perinatal asphyxia in the Netherlands and Flanders. *Neonatology*. 2013;104:15-21.
11. Azzopardi D V, Strohm B, Edwards AD, et al. Moderate hypothermia to treat perinatal asphyxial encephalopathy. *N Engl J Med*. 2009;361(14):1349-1358.
12. Gluckman PD, Wyatt JS, Azzopardi D, et al. Selective head cooling with mild systemic hypothermia after neonatal encephalopathy: multicentre randomised trial. *Lancet*. 2005;365:663-670.
13. Zhou W, Cheng G, Shao X, et al. Selective head cooling with mild systemic hypothermia after neonatal hypoxic-ischemic encephalopathy: a multicenter randomized controlled trial in China. *J Pediatr*. 2010;157(3):367-372.e3.
14. Shankaran S, Laptook AR, Ehrenkranz RA, et al. Whole-body hypothermia for neonates with hypoxic–ischemic encephalopathy. *N Engl J Med*. 2005;353(15):1574-1584.
15. Jacobs SE, Morley CJ, Inder TE, et al. Whole-Body Hypothermia for Term and Near-Term Newborns With Hypoxic-Ischemic Encephalopathy. A rondomized Controlled Trial. *Arch Pediatr Adolesc Med*. 2011;165(8):692-700.
16. Simbruner G, Mittal RA, Rohlmann F, Muche R. Systemic hypothermia after neonatal encephalopathy: outcomes of neo.nEURO.network RCT. *Pediatrics*. 2010;126(4):e771-778.
17. Thelen E. Motor development. A new synthesis. *Am Psychol*. 1995;50:79-95.
18. Hadders-Algra M. The neuronal group selection theory: an attractive framework to explain variation in normal motor development. *Dev Med Child Neurol*. 2000;42:566-572.

19. Prechtl HFR. General movement assessment as a method of developmental neurology: new paradigms and their consequences. *Dev Med Child Neurol*. 2001;43:836-842.
20. Hadders-Algra M, Mavinkurve-Groothuis AMC, Groen SE, Stremmelaar EF, Martijn A, Butcher PR. Quality of general movements and the development of minor neurological dysfunction at toddler and school age. *Clin Rehabil.* 2004;18(3):287-299.
21. Einspieler C, Prechtl HFR. Prechtl's assessment of General Movements: a Diagnostic tool for the functional assessment of the young nervous system. *Ment Retard Dev Disabil Res Rev.* 2005;11(November 2004):61-67.
22. Hadders-Algra M. Putative neural substrate of normal and abnormal general movements. *Neurosci Biobehav Rev*. 2007;31(8):1181-1190.
23. Bosanquet M, Copeland L, Ware R, Boyd R. A systematic review of tests to predict cerebral palsy in young children. *Dev Med Child Neurol.* 2013;55:418-426.
24. Novak I, Morgan C, Adde L, et al. Early, accurate diagnosis and early intervention in Cerebral Palsy: advances in diagnosis and treatment. *JAMA Pediatr.* 2017;171(9):897-907.
25. Hadders-Algra M. Neural substrate and clinical significance of general movements: an update. *Dev Med Child Neurol*. 2018;60(1):39-46.
26. Prechtl HFR, Hopkins B. Developmental transformations of spontaneous movements in early infancy. *Early Hum Dev*. 1986;14:233-238.
27. Janssen AJWM, Diekema ETW, Dolder van R, Kollée LAA, Oostendorp RAB, Nijhuis-van derSanden MWG. Development of a movement quality measurement tool for children. *Phys Ther*. 2012;92(4):574-594.
28. Thompson CM, Puterman a S, Linley LL, et al. The value of a scoring system for hypoxic ischaemic encephalopathy in predicting neurodevelopmental outcome. *Acta Paediatr*. 1997;86(7):757-761.
29. Thoresen M, Satas S, Løberg EM, et al. Twenty-four hours of mild hypothermia in unsedated newborn pigs starting after a severe global hypoxic-ischemic insult is not neuroprotective. *Pediatr Res*. 2001;50(3):405-411.
30. Bayley N. *Bayley Scales of Infant and Toddler Development, 3 Edition (Bayley-III).* San Antonio: the psychological Cooperation; 2006.
31. Van Baar AL, Steenis LJP, Verhoeven M, Hessen DJ. *Bayley-III-NL, Handleiding Motororiekschaal*. Amsterdam: Pearson Assessment and Information B.V.; 2014.
32. Steenis LJP, Verhoeven M, Hessen DJ, van Baar AL. Performance of Dutch children on the Bayley III: A comparison study of US and Dutch norms. *PLoS One*. 2015;10(8):e0132871.
33. Henderson SE, Sugden DA, Barnett LA. *Movement Assessment Battery for Children- Second Edition (MABC-2)*. London: Pearson Education International; 2010.
34. Smits-Engelsman B. *Movement Assessment Battery for Children-2. Dutch Standardization*. Amsterdam: Pearson Education; 2010.
35. Wood E, Rosenbaum P. The Gross Motor Function Classification System for cerebral palsy: a study of reliability and stability over time. *Dev Med Child Neurol*. 2000;42(5):292-296.
36. Weis R. *Gross Motor Function Measure (GMFM-66 and GMFM-88) User's Manual*. Vol 8. London: Mc Keith Press; 2004.

37. Koopmans-van Beinum FJ, Os den EA. Spraak in ontwikkeling: perceptie, productie en interactie in het eerste levensjaar. *Stem-, Spraak- en Taalpathologie, 2.* 1993;3:137-161.
38. Zink I, Leajagere M. *Lijsten Voor Communicatieve Ontwikkeling; Aanpassing En Hernormering van de MacArthus CDIs van Fenson et Al.* Leuven/Leusden, Belgie: Acco; 2002.
39. Eldik MCM, et.al. *Handleiding Reynell Test Voor Taalbegrip.* Nijmegen, Nederland: Berkhout Nijmegen B.V.; 1995.
40. Schlichting JEP, et.al. *Handleiding Schlichting Test Voor Taalproductie Derde Gewijzigde Druk.* 3rd ed. Lisse, Nederland: Swets & Zeitlinger B.V.; 2003.
41. Bleichrodt N, Resing WC., Drenth PJ., Zaal J. *Intelligentie-Meting Bij Kinderen: Empirische En Methodologische Verantwoording van de GeReviseerde Amsterdamse Kinder Intelligentie Test.* Lisse: Swets&Zeitlinger; 1987.
42. Beery K, Buktenica N., Beery N. *Beery-Buktenica Developmental Test of Visual Motor Integration (Beery VMI)*. 6th ed. San Antonio: Pearson; 2010.
43. Beery K, Buktenica NA. *The Beery-Buktenica Developmental Test of Visual-Motor Integration: VMI with Supplemental Developmental Tests of Visual Perception and Motor Coordination: Administration, Scoring and Teaching Manual.* Parsippany, NJ: Modern Curriculum Press; 1997.
44. Achenbach MT. *Manual for the Child Behavior Checlist/4-18.* Burlington, VT: University of Vermeont, Department of Psychiatry; 1991.
45. Portney LG, Watkins MP. *Foundations of Clinical Research, Applications to Practice.* Third Edit. London: Pearson Education International; 2009.
46. Guillet R, Edwards AD, Thoresen M, et al. Seven-to eight-year follow-up of the CoolCap Trial of head cooling for neonatal encephalophay. *Pediatr Res.* 2012;71(2):205-209.
47. Laerhoven van H, Haan de TR, Offringa M, Post B, Lee van der JH. Prognostic tests in term neonates with hypoxic-ischemic encephalopaty: a systematic review. *Pediatrics.* 2013;131:88-98.
48. van Baar AL, Steenis LJP, Hessen DJ. *Bayley-III-NL; Technische Handleiding.* Amsterdam, the Netherlands: Pearson Assessment and Information B.V.; 2014.

Chapter 4

Interrater reliability of the Observable Movement Quality scale for children

Lieke MA Dekkers
Maria WG Nijhuis-van der Sanden
Marianne Jonker
Bert JM de Swart
Anjo JWM Janssen

Published in: Physiotherapy Canada 2018; 70(2);113-119

Abstract

Purpose

The authors investigated the interrater reliability, the standard deviation of he random measurement error, and the limits of agreement (LoA) of the Observable Movement Quality (OMQ) scale in children. Movement quality is important in the recognition of motor problems, and the OMQ scale, a questionnaire used by paediatric physiotherapists, has been developed for use with an age-specific motor test to observe movement quality and score relative to what is expected for a child's age.

Method

Paediatric physiotherapists (*n* = 28; 2 men, 26 women) observed video-recorded assessments of age-related motor tests in children (*n* = 9) aged 6 months to 6 years and filled in the OMQ scale (possible score range 15–75 points). For our analyses, we used linear mixed models without fixed effects.

Results

The interrater reliability was moderate (intra-class correlation coefficient [$ICC_{2,1}$]: 0.67, 95% CI: 0.47, 0.88); neither work setting nor work experience exerted any influence on it. The standard deviation of the random measurement error was 5.7, and the LoA was 31.5. Item agreement was good (proportion of observed agreement [P_o] total 0.82–0.99).

Conclusion

The OMQ scale showed moderate interrater reliability when being used by therapists who were unfamiliar with the questionnaire and who had received only 2 hours of training. Feedback from the participants suggested a need for more comprehensive training in using the OMQ scale in clinical practice.

Movement quality gives an impression of how movements are controlled and coordinated.[1] Thus, it represents the interaction between personal characteristics and experience, task difficulty, and environmental conditions, and it gives one an insight into the potential of the neurological system to react or adapt to changing conditions.[2] In physiotherapy, assessment of movement quality is relevant for recognizing motor problems, evaluating interventions, and predicting recovery.[3–7] To obtain information about movement quality, clinicians and researchers must rely on subjective observation—that is, the process of gathering, organizing, and giving meaning to visual, auditory, and sensory information obtained about a moving person.[2,3]

During the acquisition and re-acquisition of movement, clinicians can observe both quantitative and qualitative changes.[8] Quantitative changes can be seen in people's acquisition of new and more complex motor skills.[9] Currently available discriminative motor tests specifically assess quantitative aspects by comparing individuals with their peers; these tests are norm referenced and validated.

However, changes in the quality of movements demonstrate more subtle characteristics, such as velocity, fluency, accuracy, and automatism of movements.[9] Available and commonly used qualitative measurement instruments focus mostly on specific diagnostic groups, such as children with cerebral palsy (Quality of Function Measure [QFM]),[6] or are designed to assess the functioning of extremities (Quality of Upper Extremity Skills Test [QUEST])[7,10] or children in a specific age range (General Movements [GMs], Infant Motor Profile).[11,12] Earlier studies[2,13] found that descriptions of movement quality are frequently used but not standardized; such descriptions differ among therapists depending on the theoretical construct used in the clinical reasoning, which precludes comparability and longitudinal evaluation.[14,15] Currently, no generic instruments are available to assess movement quality in children over time for all age categories.

To fill this gap, we developed the Observable Movement Quality (OMQ) scale.[13] The OMQ scale is a questionnaire in which each item focuses on an element of observable movement quality (e.g., presence of tremors, fluency, speed of movements). While observing, the therapist is asked to take into account the expected level of performance for a child's age and developmental stage, the task performed, and the environmental circumstances. Therefore, scoring demands an introspective judgment of movement quality based on systematic observations and internal reflection, which incorporates the therapist's knowledge, reasoning, and specific experiences with the target group of children.[2,3] Although the development process established the OMQ scale's content validity,[13] studies on psychometric properties are needed to validate its use in clinical practice.

The aim of this study was to determine the OMQ scale's interrater reliability and standard deviation of the measurement error for paediatric physiotherapists who assessed children from ages 6 months to 6 years with different diagnoses. We decided to start with this age group because judging movement quality is more challenging in younger children because of the larger neurobiological changes that occur during early childhood.[11] Moreover, we chose

a design using more than two raters to increase its generalizability to clinical practice.[16] The OMQ scale's scoring is, as previously mentioned, based on the introspective judgment of movement quality, which will be influenced by knowledge, reasoning, and personal experiences with the target group of children;[17,18] therefore, we decided to perform two subgroup analyses based on therapists' work setting and years of work experience.

METHODS

This was a cross-sectional reliability study in which paediatric physiotherapists judged video recordings of assessments of norm-referenced motor tests of nine children. The medical ethical committee of Radboud University Medical Centre approved the study, which conforms to the principles of the Declaration of Helsinki (registration number 2011/370).

Paediatric physiotherapists

This study included a stratified sample of paediatric physiotherapists employed in a variety of work settings to guarantee that we included therapists with a variety of clinical expertise: private paediatric physiotherapists practices, general hospitals and medical day care centres, and university hospitals and rehabilitation centres. The work settings were located in the southern and central parts of the Netherlands. For each work setting, we included an equal distribution of novice and experienced paediatric physiotherapists. The categories of work experience were based on studies by Jensen and colleagues[19,20] and Wainwright and colleagues.[21] To obtain sufficient contrast between novice and experienced physiotherapists, we included experienced paediatric physiotherapists with 8 years or more years of work experience and novice paediatric physiotherapists with 5 or fewer years of work experience. Therapists were verbally informed by the researchers about the study, and those who were interested received an invitation letter in October or November 2011 explaining the study's aim and the total time investment (about 6 h over the course of 5 wk).

The participating therapists signed informed consent forms and received explanations of the privacy rules pertaining to the video recordings of the children. The therapists then received an invitation to a 2-hour training session on scoring the OMQ scale; sessions were organized at nine locations. None of the participants had previous experience with the OMQ scale. The training outlined the purpose of the scale and explained the definitions of the items; all participants received a manual. Participants then watched one video recording of a child with motor problems and filled in the OMQ scale individually. Finally, the scores were compared among the participants; differences and problems in scoring were discussed and unclear issues resolved.

After the training, each therapist received a DVD and numbered OMQ scale scoring sheets for each video recording. The numbers on the scoring sheets corresponded to a unique number for each therapist combined with a number for each child. We asked the therapists to observe the video recording of each of the nine children individually in the order

recorded on the DVD and to score each child's motor quality according to the OMQ scale. The therapists had a maximum of 5 weeks to return the DVDs and OMQ scale scoring sheets to the researchers, using the reply envelope included.

OMQ scale

The OMQ scale[13] was designed for children aged 3 months to 16 years. The 15-item questionnaire needs to be filled in against an age-specific, discriminative motor test to observe and score movement quality relative to what is expected for a child's age. The 15 items are scored on a 5-point Likert scale; thus, total scores range from 15 to 75 (see last presented Table for the 15 scale items). Lower scores indicate lower movement quality. Content validity was established during the development of the OMQ scale.[13]

Video-recorded children

For this study, we video recorded nine children; this enabled multiple paediatric physiotherapists to observe each child in the same condition. All parents signed informed consent forms for the recording and use of the video for this study.

Eight children were recruited through paediatric physiotherapy practices as a representative sample. The inclusion criteria were (1) aged 6 months to 6 years and (2) a diagnosis or indication for treatment by a paediatric physiotherapy. We also recruited one typically developing child to ensure that the video recordings included a representation of typical movement quality. We video recorded an age-appropriate motor test during a 1-hour session and used the Alberta Infant Motor Scale to assess children aged 6–13 months;[22] the Bayley Scales of Infant and Toddler Development, Third Edition, to assess children aged 15–23 months;[23] and the Movement Assessment Battery for Children, Second Edition, Dutch version, to assess children aged 3–6 years.[24]

One experienced paediatric physiotherapy performed all the motor tests, and another researcher video recorded all the motor tests using a pre-designed protocol. We edited the video recordings to be 15 minutes long per child, ensuring that they showed both fine and gross motor skills and that the aspects of the OMQ scale were observable. The nine video recordings were copied onto a DVD in a random order, using the random number generators menu in IBM SPSS Statistics, version 21 (IBM Corporation, Armonk, NY) to reduce the influence of learning during observation of the nine video cases on the outcome measures.

Statistical methods

We described the characteristics of the therapists and video recorded children to establish the median and range of the continuous variables and the number and percentage of the categorical data. We converted the motor test scores into *z* scores and calculated OMQ scale total scores as median and range for all therapists and for the two work experience subgroups.

To study the standard deviation of the random measurement error of the OMQ scale, we used a linear mixed model without fixed effects. The dependent variable was the total score on the OMQ scale. Therapists and videorecorded children were treated as random variables (Model A). To study the differences in random measurement error between the two subgroups of paediatric physiotherapists (novice, working ≤5 y; experienced, working ≥8 y) and the three work setting subgroups (paediatric physiotherapy practice, general hospitals and medical day care centres, and academic hospitals and rehabilitation centres), we used the same linear mixed model but in a manner (i.e., using a grouping statement in the random intercept statement) that allowed us to estimate a random measurement error per experience group (Model B) and per work setting (Model C).

Initially, we included experience and work setting as independent class variables in Models A, B, and C. However, these terms were always far from statistically significant ($p > 0.80$) and so were omitted from the final models. We calculated the OMQ scale scores obtained from the paediatric physiotherapists as a group and by subgroup to obtain the intra-class correlation coefficient type 2:1 ($ICC_{2,1}$), a two-way random effects single-measures model of absolute agreement, standard deviation of the random measurement error, repeatability coefficient (RC), and limits of agreement (LoA). Note that the last two calculations are specific interpretations of the standard deviation of the random measurement error. Furthermore, item agreement is presented as linear-weighted κ, the percentage of observed agreement (P_O), and P_O total, which includes the agreement of a 1-point scoring difference on the Likert scale.

For sample size calculation, we assumed an interrater ICC of 0.8 (i.e., good reliability) and more than 0.6 (i.e., moderate reliability). To obtain a power of 80% ($\alpha = 0.05$, F test), we needed a minimum of 23 observers observing nine different videos.[25] The data were checked for outliers. Statistical analyses were performed in IBM SPSS Statistics and SAS version 9.2 for Windows (SAS Institute, Cary, NC). Two-sided *ps* < 0.05 were considered statistically significant.

RESULTS

Paediatric physiotherapists

Thirty-one paediatric physiotherapists agreed to participate in this study. Three female therapists were excluded—one who failed to complete four of nine OMQ scales; a second who misinterpreted the Likert scale and scored inconsistently, as confirmed by outlier analysis; and a third who had technical problems playing the video recordings on the DVD. Table 1 shows the characteristics of the 28 paediatric physiotherapists by work setting.

Of the 28 participants, 26 (93%) women and 2 (7%) men had a median work experience of 11 years (range 1–29 y), and 12 (43%) worked in a paediatric physiotherapists practice, 8 (29%) worked in a general hospital or medical day care centre, and 8 (29%) worked in a university hospital or rehabilitation centre. These 28 paediatric physiotherapists returned

252 OMQ scale scoring sheets. Median OMQ scale total scores ranged from 43 to 67 for all patients (see Table 2).

Table 1. Characteristics of Paediatric physiotherapists by Work Setting.

Characteristic	Total (*n* = 28)		Paediatric physiotherapy practice		General hospital or medical day care centre		Academic hospital or rehabilitation centre	
	No. (%)	Median (range)	No. (%)	Median (range)	No. (%)	Median (range)	No. (%)	Median (range)
Work experience, y								
≤5	11 (39)	3 (1–5)	5 (42)	2 (1–3)	3 (38)	4 (1–5)	3 (38)	2 (2–4)
≥8	17 (61)	20 (8–29)	7 (58)	19 (10–29)	5 (63)	9 (8–29)	5 (63)	25 (12–28)
Sex								
Male	2 (7)	-	2 (17)	-	0	-	0	-
Female	26 (93)	-	10 (83)	-	8 (100)	-	8 (100)	-

Note: Percentages may not total 100 because of rounding.

Table 2. Characteristics of Video-Recorded Children (*n* = 9) and Results on Motor Test and OMQ Scale.

Characteristic			Motor test		OMQ scale total score, median (range)		
Age at video recording, mo	Sex	Diagnosis/indications for treatment	Instrument	*z*-score*	All therapists (*n* = 28)	Work experience, y ≤5 (n = 11)	Work experience, y ≥8 (*n* = 17)
13	F	Neuromuscular disorder with hypotonia	AIMS	-7.7	43 (32–51)	46 (35–51)	39 (32–51)
18	M	Trisomy 21	BSID–III	-3.2	51.5 (37–60)	52 (37–55)	50 (38–60)
64	M	Developmental coordination disorder	MABC–2–NL	-3.0	47 (36–60)	51 (44–59)	44 (36–60)
23	F	Trisomy 21	BSID–III	-2.6	53 (32–61)	53 (32–59)	53 (39–61)
14	F	Spastic cerebral palsy, unilateral	BSID–III	-2.2	51 (42–59)	53 (43–59)	50 (42–58)
8	M	Pre-term birth	AIMS	-1.7	65.5 (49–74)	66 (56–74)	64 (49–74)
54	M	Developmental coordination disorder	MABC–2–NL	-1.7	55.5 (47–69)	58 (48–64)	55 (47–69)
6	F	Idiopathic asymmetry†	AIMS	-0.6	67 (48–75)	67 (49–74)	67 (48–75)
38	M	Typical development	MABC–2–NL	1.7	64 (51–74)	62 (51–70)	66 (52–74)

*Standardized score, whereby the raw score is expressed in standard deviation units to compare it with norm scores from typically developing children of the same age (mean = 0; SD = 1).
†Seen in young infants with an asymmetrical head and/or body posture.
OMQ = Observable Movement Quality; F = female; AIMS = Alberta Infant Motor Scale; M = male; BSID–III = Bayley Scales of Infant Development, Third Edition; MABC–2–NL = Movement Assessment Battery for Children, Second Edition, Dutch version.

Video-recorded children

Table 2 also shows the characteristics of the children, five boys (56%) and four girls (44%), aged 6 months to 5 years, 4 months. As the table shows, diagnoses and indications for treatment by a paediatric physiotherapist were common except for the one typically developing child. Motor test *z* scores ranged from –7.7 to 1.7.

Interrater reliability

The interrater reliability was moderate ($ICC_{2,1}$: 0.67; 95% CI: 0.47, 0.88;[26] Table 3). The standard deviation of the random measurement error was 5.7, and no statistically significant differences (i.e., systematic measurement errors) were found among the paediatric physiotherapists. The RC was 15.7, representing the value below which the absolute difference between two measurements can be expected only in the presence of random measurement error. The ICC, RC, and LoA across the different subgroups (work experience and work setting) were similar to those for all therapists as a group.

Item agreement

Table 4 shows that the median score for all items on the OMQ scale varied between 3 and 5. For item agreement, κ values for each scale item were low to fair (0.07–0.54), the proportions of observed agreement were fair to good (0.42–0.94), and they improved to good (0.82–0.99) when a 1-point scoring difference on the Likert scale was accepted.

Table 3. Interrater Reliability of the OMQ Scale for the Paediatric Physiotherapists by Work Experience and Work Setting.

Paediatric physiotherapists	$ICC_{2,1}$	95% CI	SD			RC	LoA
			Random measurement error	Between subjects	Among therapists		
All (*n* = 28)*	0.67	0.47, 0.88	5.7	8.0	3.1	15.7	31.5
Work experience†							
≤5 y	0.60	0.37, 0.86	5.8	8.1	2.2	16.1	32.3
≥8 y	0.71	0.51, 0.90	5.6	8.1	3.5	15.4	30.9
Work setting‡							
Paediatric physiotherapy practice	0.70	0.49, 0.90	5.4	8.0	2.6	15.1	30.2
General hospital or medical daycare centre	0.62	0.38, 0.87	5.6	8.0	4.0	15.4	30.8
Academic hospital or rehabilitation centre	0.63	0.39, 0.87	6.1	8.0	2.6	16.9	33.9

*Based on Model A.
†Based on Model B.
‡Based on Model C.
OMQ = Observable Movement Quality; $ICC_{2,1}$ = intra-class correlation coefficient, a two-way random effects single-measures model of absolute agreement;
RC = repeatability coefficient; LoA = limits of agreement.

Table 4. Agreement of Each Item on the OMQ Scale, Scored by Paediatric Physiotherapists for Video-Recorded Children, Using a 5-Point Likert Scale.

Item	Median (range)	Weighted k* (95% CI)	Mean (range)	
			Po	*Po* total†
1. Appropriate fine motor movements	3 (1–5)	0.35 (0.23, 0.48)	0.52 (0.32–0.61)	0.88 (0.68–1.0)
2. Appropriate gross motor movements	3 (1–5)	0.34 (0.21, 0.46)	0.44 (0.39–0.57)	0.91 (0.75–1.0)
3. Fluency of movements	3 (1–5)	0.24 (0.12, 0.36)	0.45 (0.32–0.57)	0.88 (0.79–0.93)
4. Reduced muscle tone	3 (1–5)	0.54 (0.40, 0.68)	0.58 (0.32–0.82)	0.89 (0.50–1.0)
5. Increased muscle tone	5 (1–5)	0.31 (0.05, 0.57)	0.72 (0.36–0.1)	0.90 (0.68–1.0)
6. Tremors	5 (3–5)	0.07 (–0.07, 0.20)	0.94 (0.86–1.0)	0.99 (0.96–1.0)
7. Slow and/or delayed movements	4 (1–5)	0.42 (0.17, 0.67)	0.52 (0.36–0.89)	0.82 (0.64–1.0)
8. Accelerated and/or abrupt movements	5 (2–5)	0.14 (0.04, 0.24)	0.65 (0.43–1.0)	0.87 (0.75–1.0)
9. Asymmetry in movements	4 (1–5)	0.40 (0.09, 0.72)	0.56 (0.43–0.79)	0.90 (0.89–1.0)
10. Accuracy (well-aimed)	3 (1–5)	0.26 (0.17, 0.36)	0.53 (0.39–0.82)	0.94 (0.82–1.0)
11. Strength regulation	3 (1–5)	0.28 (0.19, 0.37)	0.45 (0.39–0.53)	0.93 (0.82–1.0)
12. Variation in movements	3 (1–5)	0.27 (0.13, 0.42)	0.42 (0.32–0.46)	0.91 (0.82–1.0)
13. Involuntary movements	4 (1–5)	0.19 (0.01, 0.37)	0.53 (0.39–0.85)	0.86 (0.75–0.96)
14. Automated movements	3 (1–5)	0.29 (0.17, 0.41)	0.49 (0.39–0.61)	0.90 (0.82–0.96)
15. Stereotype movements	5 (1–5)	0.31 (0.03, 0.59)	0.71 (0.50–0.92)	0.91 (0.81–1.0)

*Linear weighting.
†Agreement of a 1-point scoring difference on the Likert scale (a score of 1 point higher or lower).
OMQ = Observable Movement Quality; Po = proportion of observed agreement.

DISCUSSION

In this cross-sectional reliability study, we identified a moderate interrater reliability for the total score on the 15-item OMQ scale. We found no differences between the two groups of paediatric physiotherapists (≤5 or ≥8 y of experience) or among the therapists in the three types of work setting.

This study used video recordings instead of live assessments. Using video recordings both ensured that all therapists observed movement quality under the same circumstances and eliminated the need for multiple observers to examine the children at once. Given that the therapists were observing video recordings, they were unable to interact with the children as they would do in clinical practice, but they did not mention this as a problem. However, they recognized that using video recordings could lead to losing some information for items related to muscle tone.

The lack of difference in interrater reliability based on either work experience or work setting does not support the hypothesis for the expected differences in introspective judgment of movement quality on the basis of clinical experience. One explanation could be that paediatric physiotherapists in the Netherlands complete a master's programme in paediatric physiotherapists after receiving their bachelor's degree in physiotherapists. During this 3-year master's programme, physiotherapists work part time with children under

the supervision of an experienced colleague—in addition to completing their coursework—to develop clinical expertise by observing, treating, and evaluating interventions. This study focused only on years spent working as a certified paediatric physiotherapists and did not include the years spent working as a general physiotherapists. Thus, the differences in outcomes might have been higher if novices with 5 or fewer years of overall working experience had been included.

This study showed a wide range in OMQ scale total scores (17–29 points difference per video-recorded child) and a reasonably large RC (15.7) for the OMQ scale. This could indicate a variation in how the participating therapists interpreted the scoring options for the OMQ scale.[27] For example, as the best-fitting choice for item agreement, we used linear-weighted κ statistics[28]; however, the high level of agreement among the observers led to a low κ value[29,30] for item 6 (0.07), and multiple observers gave the exact same score (i.e., perfect agreement). Furthermore, the small sample size of included children prevented us from performing statistical correlations for the outcomes on the motor tests and OMQ scale. However, ranking the *z* scores for the motor tests showed that children with higher z scores also showed higher OMQ scale total scores. Only the pre-term infant showed a delay in motor performance (*z* score –1.7), with a high median OMQ scale total score; this score indicates good quality of movement, which can be observed in pre-term infants at this age,[31] and demonstrates the potential for this child to catch up in motor performance.

Reliability studies are often performed with two or three extensively trained, experienced raters. However, in clinical practice many therapists, both novice and experienced, use a measurement instrument. Reliability studies that use only two or three raters yield results with limited generalizability for the clinical setting.[16] By including 28 paediatric *z* employed all over the Netherlands in different work settings and taking into account their years of experience, we increased the generalizability of the results, and we further enhanced them by including children with a variety of diagnoses and who were representative of daily practice.

In this study, the focus was on detecting movement quality differences in clinical practice rather than on using the OMQ scale for evaluative purposes. In the future, evaluative and longitudinal studies in which intrarater reliability is more relevant will be necessary. The results of this study are motivating and illustrate how training in using the OMQ scale can be improved, including revising the scoring instructions.

Compared with the results of other measurement tools for movement quality,[6,8,10,12] the results of the OMQ scale for interrater reliability were lower. However, both the QFM[6,8] and the QUEST[10] were designed for the cerebral palsy diagnosis group and developed to describe impairment-related movement quality, whereas the OMQ scale was intended to be a generic measurement tool to assess movement quality of the entire body, for all age categories and all diagnoses. In addition, for this study, we developed a 2-hour training session for participating therapists to explain the scale and teach them how to use and interpret it. None of the participating therapists had used the OMQ scale before. In

comparison, training in using the GMs,[11] QUEST, and QFM takes 1–2 full days. These factors could have contributed to the lower interrater reliability outcomes in this study.

This study had one limitation: It included two children with Down syndrome for video recording and using a norm-referenced test. Conversations with the therapists revealed that they found it challenging to score these children, possibly because paediatric physiotherapists are trained to use the developmental trajectories for such children (as described by Palisano and colleagues[32]) as reference values while observing them. In this study, the therapists had to change perspective and compare their observations with typical development. As the scoring differences on the OMQ scale demonstrated, this perceptual shift proved difficult. During further development of the training for the OMQ scale, we will take these perceived difficulties into account by expanding the focus on observation, regardless of expected motor performance for certain diagnoses or syndromes, supported by videotaped examples.

CONCLUSION

The OMQ scale demonstrates moderate interrater reliability when used by paediatric physiotherapists to assess movement quality of children aged 6 months to 6 years. These therapists were unfamiliar with the questionnaire and attended a 2-hour training session on it. Our findings are motivating and indicate that the OMQ scale could be used reliably in clinical practice, although they suggest a need to improve the training. A future study may show that more intensive training can improve the OMQ scale's interrater reliability, a necessary step before determining responsiveness and interpretability. Future clinical cohort studies should also test the effect of the age of a child on interrater reliability and on differences between video and life scoring.

KEY MESSAGES

What is already known on this topic

The assessment of movement quality is relevant for recognizing motor problems, evaluating interventions, and predicting recovery. Currently, no generic instrument is available to assess movement quality over time for all age categories. The Observable Movement Quality (OMQ) scale was developed for this purpose; however, studies on its psychometric properties are needed.

What this study adds

This study demonstrates that it is feasible to rate movement quality using the OMQ scale; however, more comprehensive training is necessary to increase the moderate interrater reliability in therapists unfamiliar with the questionnaire.

REFERENCES

1. Magill R, Anderson D. *Motor learning and control: concepts and applications.* 10th ed. Singapore: McGraw Hill Education; 2014.
2. Skjaerven LH, Kristoffersen K, Gard G. An eye for movement quality: a phenomenological study of movement quality reflecting a group of physiotherapists' understanding of the phenomenon. Physiother Theory Pract. 2008;24(1):13–27.
3. Knudson DV, Morrison CS. *Qualitative analysis of human movement.* 2nd ed. Leeds, UK: Human Kinetics; 2002.
4. Heineman KR, Hadders-Algra M. Evaluation of neuromotor function in infancy—A systematic review of available methods. J Dev Behav Pediatr. 2008;29(4):315–23.
5. Harris SR, Heriza CB. Measuring infant movement: clinical and technological assessment techniques. Phys Ther. 1987;67(12):1877–80.
6. Wright FV, Rosenbaum P, Fehlings D, et al. The Quality Function Measure: reliability and discriminant validity of a new measure of quality of gross motor movement in ambulatory children with cerebral palsy. Dev Med Child Neurol. 2014;56(8):770–8.
7. Sorsdahl AB, Moe-Nilssen R, Kaale HK, et al. Change in basic motor abilities, quality of movement and everyday activities following intensive, goal-directed, activity-focused physiotherapy in a group setting for children with cerebral palsy. BMC Pediatr. 2010;10(1):26.
8. Boyce WF, Gowland C, Hardy S, et al. Development of a quality-ofmovement measure for children with cerebral palsy. Phys Ther. 1991;71(11):820–8, discussion 828–32.
9. Schmidt RA, Wrisberg CA. *Motor learning and performance, a situation-based learning approach.* 4th ed. Leeds, UK: Human Kinetics; 2008.
10. DeMatteo C, Law M, Russell D, et al. The reliability and validity of the Quality of Upper Extremity Skills Test. Phys Occup Ther Pediatr. 1993;13(2):1–18.
11. Hadders-Algra M. General movements: a window for early identification of children at high risk for developmental disorders. J Pediatr. 2004;145(2 Suppl.):S12–8.
12. Hecker E, Baer GD, Stark C, et al. Inter- and intrarater reliability of the Infant Motor Profile in 3- to 18-month-old infants. Pediatr Phys Ther. 2016;28(2):217–22.
13. Janssen AJWM, Diekema ETW, van Dolder R, et al. Development of a movement quality measurement tool for children. Phys Ther. 2012;92(4):574–94.
14. Edwards I, Jones M, Carr J, et al. Clinical reasoning strategies in physiotherapy. Phys Ther. 2004;84(4):312–30, discussion 331–5.
15. Wainwright SF, Shepard KF, Harman LB, et al. Novice and experienced physiotherapist clinicians: a comparison of how reflection is used to inform the clinical decision-making process. Phys Ther. 2010;90(1):75–88.
16. Van Genderen FR, de Bie RA, Helders PJM, et al. Reliability research: towards a more clinically relevant approach. Phys Ther Rev. 2003;8(4):169–76.

17. Hayward LM, Black LL, Mostrom E, et al. The first two years of practice: a longitudinal perspective on the learning and professional development of promising novice physiotherapists. Phys Ther. 2013;93(3):369–83.
18. Black LL, Jensen GM, Mostrom E, et al. The first year of practice: an investigation of the professional learning and development of promising novice physiotherapists. Phys Ther. 2010;90(12):1758–73.
19. Jensen GM, Shepard KF, Hack LM. The novice versus the experienced clinician: insights into the work of the physiotherapist. Phys Ther. 1990;70(5):314–23.
20. Jensen GM, Gwyer J, Hack LM, et al. Expertise in physiotherapy practice. 2nd ed. St. Louis, MO: Saunders Elsevier; 2007.
21. Wainwright SF, Shepard KF, Harman LB, et al. Factors that influence the clinical decision making of novice and experienced physiotherapists. Phys Ther. 2011;91(1):87–101.
22. Piper MC, Darrah J. Motor assessment of the developing infant. Vol. 81, 1st ed. Philadelphia: W.B. Saunders; 1994.
23. Bayley N. *Bayley Scales of Infant and Toddler Development, 3rd edition.* San Antonio (TX): Psychological Cooperation; 2006.
24. Henderson SE, Sugden DA, Barnett LA. *Movement Assessment Battery for Children, second edition, manual, Nederlandse bewerking.* London: Pearson Education International; 2010.
25. Shoukri MM, Asyali MH, Donner A. Sample size requirements for the design of reliability study: review and new results. Stat Methods Med Res. 2004;13(4):251–71.
26. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60(1):34–42.
27. Bartlett JW, Frost C. Reliability, repeatability and reproducibility: analysis of measurement errors in continuous variables. Ultrasound Obstet Gynecol. 2008;31(4):466–75.
28. Schouten HJA. Measuring pairwise interobserver agreement when all subjects are judged by the same observers. Stat Neerl. 1982;36(2):45–61.
29. de Vet HCW, Mokkink LB, Terwee CB, et al. Clinicians are right not to like Cohen's κ. BMJ. 2013;346:f2125.
30. Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. J Clin Epidemiol. 1990;43(6):543–9.
31. van Haastert IC, de Vries LS, Helders PJ, et al. Early gross motor development of preterm infants according to the Alberta Infant Motor Scale. J Pediatr. 2006;149(5):617–22.
32. Palisano RJ, Walter SD, Russell DJ, et al. Gross motor function of children with Down syndrome: creation of motor growth curves. Arch Phys Med Rehabil. 2001;82(4):494–500.

Chapter 5

Reliability and responsiveness of the Observable Movement Quality scale for children with mild to moderate motor impairments

Lieke MA Dekkers
Bert JM de Swart
Marianne Jonker
Pauline van Erp
Anneke Wisman
Philip J van der Wees
Maria WG Nijhuis-van der Sanden
Anjo JWM Janssen

Accepted for publication in: Journal of Physical and Occupational Therapy in Pediatrics

Abstract

Aim

The Observable Movement Quality (OMQ) scale measures generic movement quality and is used alongside standardized age-adequate motor performance tests. The scale consists of 15 items, each focusing on a different aspect; together, the entire construct of movement quality is assessed. This study aimed to determine interrater and intrarater reliability, and responsiveness of the OMQ scale.

Methods

A prospective intervention study with pre-post design in paediatric physiotherapy practices. For interrater reliability, 3 physiotherapists observed video-recorded motor assessments of 30 children with mild to moderate motor impairments—aged 4 to 12 years—using the OMQ scale. One therapist scored baseline assessment a second time for intrarater reliability, and to calculate smallest detectable change (SDC). Responsiveness (n = 28) was tested by comparing outcomes before and after intervention.

Results

Interrater reliability was moderate to good ($ICC_{2,1}$: 0.79); intrarater reliability was high ($ICC_{2,1}$: 0.97). Responsiveness results revealed an SDC of 2.4 and a minimal important change of 2.5; indicating sufficient validity in differentiating groups of children showing improved versus unchanged movement quality.

Conclusion

The OMQ scale is reliable and responsive to change when used to assess movement quality in clinical practice for children with mild to moderate motor impairments, aged 4–12 year.

Introduction

The assessment of movement quality is perceived by physiotherapists as relevant for recognizing motor problems, evaluating interventions and predicting recovery and offers insight into the developing child's possibilities for reacting or adapting to changing conditions.[1–5] Movement quality represents the interaction between personal characteristics and learning experiences; the task difficulty; and environmental conditions.[6] Furthermore, movement quality gives an impression of how movements are controlled and coordinated.[7]

During children's development, the mastery of new movements and skills increases, which can be observed through quantitative and qualitative changes.[8] Quantitative changes reflect the acquisition of new and more complex motor skills, whereas changes in quality of movements are demonstrated by more subtle characteristics such as an increase in accuracy, fluency and automated movements.[9] Available and commonly used discriminative motor tests in paediatric physiotherapy specifically assess quantitative aspects by comparison with peers. These motor tests are validated, and norm referenced. For movement quality, however, available and commonly used measurement instruments are designed for particular diagnostic groups (e.g., children with cerebral palsy) for children in a specific age frame or to assess the functioning of extremities.[4,5,10–12]

The Observable movement Quality (OMQ) scale[13] can be used to assess movement quality in children, over time and for all age categories, as a generic evaluative measurement instrument. The OMQ scale is a criterion-based measurement instrument containing 15 items, each measuring one aspect of the whole construct of movement quality. The paediatric physiotherapist completes the OMQ scale directly after the assessment with an age-specific, discriminative or disease-specific motor test, in approximately 5-10 minutes. During the development of the OMQ scale, content validity was established,[13] followed by the determination of the scale's interrater reliability in a group of children from 6 months to 6 years of age.[14] To test the OMQ scale in a broader age group, research is needed with a focus on reliability among older children. Furthermore, to use the OMQ scale as an evaluative instrument, it is necessary to gain insight into its ability to detect change over time as a result of either development or intervention.[15–17] For the latter, the smallest detectable change (SDC)[18] and the minimal important change (MIC)[18–20] are important outcomes to determine the applicability and interpretability of the OMQ scale.[21,22]

The aim of this study is to determine interrater and intrarater reliability as well as responsiveness of the OMQ scale (including SDC and MIC) in daily physiotherapist practice among children from 4 to 12 years of age.

Methods

Design and Setting

This was a prospective intervention study with a pre–post design in paediatric physiotherapy practices. Children were assessed by their treating physiotherapist using the Movement Assessment Battery for Children, 2nd edition, Dutch version (MABC-2-NL),[23,24] at baseline and after an intervention period of 3 months, consisting of at least one physiotherapy session per week. Movement quality was assessed using the OMQ scale during observations of video recordings of the motor performance assessment by examiners. To test the reliability, outcomes were compared among examiners. To test responsiveness, baseline assessment outcomes of the children were compared with outcomes after the intervention period. The medical ethical committee of Radboud University Medical Centre approved the study, which conforms to the principles of the Declaration of Helsinki (registration number 2016-2832).

Participants

Video-recorded Children

Children with mild to moderate motor impairments were recruited from November 2016 to March 2017 through two paediatric physiotherapy practices in the central part of the Netherlands. Inclusion criteria were (1) being 4 to 12 years old and (2) being indicated for treatment by a physiotherapist. To meet the inclusion criteria for the MABC-2-NL, children with neurological disorders (e.g., cerebral palsy), children unable to walk independently and children with severe cognitive impairments were excluded. The MABC-2-NL was video-recorded during a 30- to 45-minute session at baseline and after an intervention period of at least 3 months. All parents and 12-year old children signed informed consent for the recording and use of the video for this study.

Examiners

One paediatric physiotherapist (LD) and two bachelor physiotherapists (PE and AW), who were at that time completing their master's education in paediatric physiotherapy, examined the video recordings using the OMQ scale. Before the start of this study, the master students received a 4-hour training session on scoring the OMQ scale. The paediatric physiotherapist (LD) was experienced in scoring the OMQ. The students had no previous experiences with the OMQ scale. The training outlined the purpose of the scale and explained the definitions of the items. The students were educated in the development and aim of the OMQ scale and observed videos of children showing severe deviant movement quality as a frame of reference. Thereafter they watched, together with two expert paediatric physiotherapists (AJ and LD), two video recordings of a child with motor impairments and completed the OMQ scale individually. Finally, the scores were compared, differences and problems in scoring were discussed and unclear issues were resolved.

Instruments

OMQ scale

The OMQ scale[13] was designed for children from 3 months to 16 years of age. The scale needs to be filled in alongside an age-specific, discriminative or disease-specific motor test—for this study, the MABC-2-NL—to observe and score movement quality relative to what is expected for a child's age. The 15 items are scored on a 5-point Likert scale; thus, total scores range from 15 to 75 (see Table 3 for the 15 scale items). Lower scores indicate lower movement quality.

Global Perceived Effect

As a comparator instrument to measure change in movement quality, a global perceived effect (GPE) rating scale was used.[25] Treating physiotherapists were asked—before the assessment with the MABC-2-NL after the intervention period of 3 months—to answer a single question to indicate how much movement quality had changed since baseline.[26] The question asked to the therapists was: 'To what extent has the quality of movement of the child improved since the start of the paediatric physiotherapy intervention?'. Responses were scored on a 7-point Likert scale ranging from: 1 = 'very much improved' to 2 = 'much improved'; 3 = 'a little improved'; 4 = 'no change'; 5 = 'a little deterioration'; 6 = 'much deterioration'; and 7 = 'very much deterioration'.[27] Test–retest reliability for GPE is high (ICC = 0.997); however, construct validity is moderate.[28]

Procedure

The treating physiotherapist performed the assessment of the MABC-NL as a usual part of the diagnostic procedure in daily practice using the standardized procedures for administration and instructions for calculation of the test scores, as specified in the test manual. All motor test items were recorded using a predesigned video protocol. The video recordings were edited by the master students (PE and AW) to a roughly 20-minute-long video, ensuring that they showed all test items of the MABC-2-NL and that all aspects of the OMQ scale were observable. This was a technical procedure in which the video part with instructions from the therapist to the child was deleted, and multiple files of an assessment part were combined, if necessary. The students did not observe the videos during editing. The video recordings were saved on a password-protected hard disk, only accessible by the main researchers of this study (LD and AJ).

For data sampling, examiners observed the video recordings of the children individually and completed an OMQ scale scoring sheet for each recording. For interrater reliability, videos of the baseline assessment were scored by three examiners (LD, PE and AW). For intrarater reliability, the videos of the baseline assessment were scored a second time, by the expert examiner (LD). In addition, for responsiveness, videos of the assessments after intervention were scored by the expert examiner (see figure 1). The two master student examiners (PE and

AW) scored all video recordings within two weeks after the baseline assessment. The expert examiner (LD), who scored both baseline assessment and assessment after intervention, started with scoring at the end of the project when all motor performance assessments—before and after intervention—were gathered. The video-records were scored in a random order, while the examiner was blinded for measurement time point. This examiner also scored baseline assessment a second time (with at least two weeks in-between) blinded for previous scores. The expert examiner was not involved in the children's paediatric physiotherapy assessment or intervention, which prevented practitioner bias.

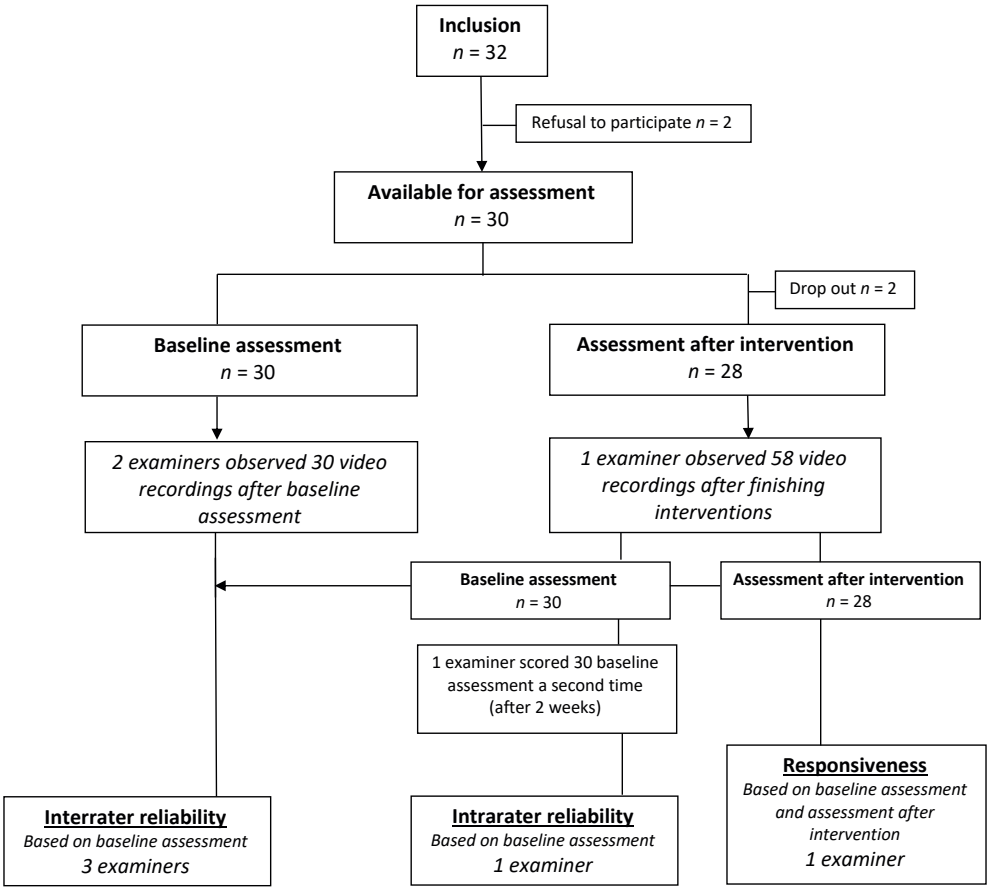


Figure 1. Flow chart

Statistical Methods

Descriptive statistics of the characteristics of the video-recorded children were presented as numbers and percentages for categorical variables and as median and interquartile range (IQR) for ordinal variables. For continuous data, means and standard deviations (SD) were reported. The data were checked for outliers. MABC-2-NL standard scores were used for analyses. The score distribution of the OMQ scale was examined for floor and ceiling effects. Floor and ceiling effects exceeding 20% of the participants were considered substantial.[29,30]

To estimate interrater and intrarater reliability of the OMQ scale, interclass correlation coefficients type 2:1 ($ICC_{2,1}$)—a two-way random-effects single-measures model of absolute agreements—with 95% confident intervals (CI) were determined,[31,32] along with the standard error of measurement (SEM) and the limits of agreement (LoA).[15] The SEM was calculated by the SD of the measurement at baseline using ICC (as SEM = SD x √[1-ICC]),[18,33] and the SEM was used to calculate the LoA (as: $\bar{d}$ ±1.96 x √2 x SEM).[18] SEM ≤ SD/2 was taken as the criterion of acceptable precision.[33]

Bland-Altman plots were used to represent the agreement between measurements graphically.[34] For the Bland-Altman plot representing interrater reliability, OMQ total scores for the 3 examiners were plotted against each other and visualized; for intrarater reliability, the OMQ scale baseline assessment, score twice by one examiner, were plotted against each other. Furthermore, item agreement presented the percentage of observed agreement (P_O). For sample size calculations, we assumed an ICC of 0.8 (i.e., good reliability) and larger than 0.6 (i.e., moderate reliability). To obtain a power of 80% (alpha = 0.05, F-test), we needed 31 videos.[35]

The overall effect of paediatric physiotherapy intervention was defined by the effect size (ES)—a standardized measurement of change calculated by dividing the mean change between baseline measurement and measurement after the intervention period by the SD of the baseline measurement—and the standardized response mean (SRM)—calculated as the mean change in scores between baseline measurement and measurement after the intervention period divided by the SD of that change score.[36] ES and SRM were calculated for both MABC-2-NL and OMQ scale total scores. A positive SRM indicated improvement, whereas a negative SRM indicated deterioration.[36] Outcomes for ES and SRM of 0.20 were considered as small, 0.50 as moderate and 0.80 as large.[37]

To assess the responsiveness of the OMQ scale, the SDC (as 1.96 x √2 x SEM) was calculated. If the change was above the SDC value in individual patients, one could be 95% confident that it was not caused by measurement error.[38] Furthermore, the MIC value for the OMQ scale was calculated to examine the discriminative ability of change scores for the OMQ scale.[20,39] To explore the interpretability of change scores, the SDC was compared to the MIC; to distinguish clinically important change from measurement error, we tested whether the MIC was greater than the SDC.[26]

The perceived improvement of movement quality on the GPE was used as an anchor (gold standard)[26,40] Outcomes for GPE were classified as 'improved' (defined as GPE scores 1–2) and 'unchanged' (defined as GPE scores 3–7). The MIC values for the OMQ scale were calculated by subtracting the mean change score of the children classified as unchanged from the mean change score of those classified as improved. To establish the validity of the anchor, a two-sample t-test was performed to test the difference between the two groups across OMQ scale scores.[30] Receiver operating characteristic (ROC) curves were used to examine various cut-off values for the OMQ scale change scores.[18] In a ROC curve, sensitivity and 1-specificity values from the 'improved' and 'unchanged' groups were plotted on a y- and x-axis. The ROC cut-off point was detected by finding the minimal distance to the upper left corner of the ROC curve, which was assumed to represent the optimal trade-off between sensitivity and specificity for detecting clinical improvement.[26] The area under the ROC curve (AUC) was used as an indicator for responsiveness. For sufficient responsiveness, an AUC over 0.70 is recommended.[26]
All statistical tests were two-sided, and $p < 0.05$ was considered significant. Data were analysed using IBM Statistical Package for the Social Sciences (IBM SPSS Statistics) version 25 (IBM Corporation, Armonk, NY).

Results

In total, 32 children were recruited for participation. The parents of one child refused to sign informed consent, and one child refused to sign for video recordings. Finally, we were able to include 30 children in this study. Table 1 shows the characteristics of the 30 children—19 boys (63%)—aged 4 years to 12 years with a mean age (SD) of 7 years and 5 months (2 years and 6 months). Indication for physiotherapy intervention was diverse; however, the majority of the children were diagnosed with motor developmental delays (63%). Two children (7%) dropped out of the intervention, both due to severe health problems of one of the parents. MABC-2-NL standard scores at baseline had a mean (SD) of 6.17 (3.51) and, after the intervention period, a mean (SD) of 7.64 (4.50). OMQ scale total scores at baseline had a mean (SD) of 67.63 (4.97) and, after the intervention, a mean (SD) of 70.07 (5.19). For possible floor and ceiling effects of the OMQ scale, none (0%) and two (7%) of the children had initially the lowest or highest possible scores, respectively (compared to 0% and 14% after the intervention period). Results on GPE ranged from 2 ('much improved') to 4 ('no change').

Table 1. Characteristics of included children (*n*=30), indication for physiotherapy intervention; outcome on Movement Assessment Battery for Children, 2nd edition, Dutch version (MABC-2-NL), Observable Movement Quality (OMQ) scale, and Global Perceived Effect (GPE) scores

Characteristics	*n*	mean	(SD)	range
Boys	19			
Girls	11			
Age in years	30	7yr5mth	(2yr6mth)	4 – 12yr
Indication for physiotherapeutic intervention				
Motor developmental delay	16			
In combination with DCD	2			
In combination with PDD-nos	1			
Clumsy motor skills	3			
Manual dexterity developmental delay	3			
Hypermobility	2			
Musculoskeletal injury	2			
Scoliosis	1			
Outcome MABC-2-NL (standard scores)				
MABC-2-NL at baseline	30	6.17	(3.51)	
MABC-2-NL after intervention	28	7.64	(4.50)	
Outcome OMQ scale (total scores)				
OMQ scale score at baseline	30	67.63	(4.97)	51 – 75
OMQ scale score after intervention	28	70.07	(5.19)	55 – 75
Outcome GPE (at t_1)		median		range
Treating paediatric physiotherapist	30	3		2 – 4

SD = Standard Deviation; yr = years; mth = months; GA = Gestational Age; DCD = Developmental Coordination Disorder; PDD-nos = Pervasive Developmental Disorder-Not otherwise Specified.

The interrater reliability indicated a moderate to good reliability (ICC2,1: 0.79; 95% CI: 0.62, 0.89), and for intrarater reliability, a high reliability ($ICC_{2,1}$: 0.97; 95% CI: 0.93, 0.98) was shown; see Table 2. The SEM values for both interrater and intrarater reliability met the criteria (SEM ≤ SD/2), suggesting an acceptable measurement precision of the OMQ scale. The LoA for interrater reliability was 33.0, and for intrarater reliability it was 7.03, indicating a better measurement precision for intrarater reliability. Bland-Altman plots for interrater reliability showed a systematic difference in OMQ total scores (–5.16) and an increase in the plots for the expert examiner (examiner 3), indicating higher total scores for this examiner (see Figure 2). Median score for all items of the OMQ scale varied between 4 and 5 for both interrater and intrarater reliability (see Table 3). For interrater and intrarater item

Table 2. Reliability of the Observable Movement Quality (OMQ) scale total scores (*n* = 30).

Characteristics	$ICC_{2,1}$ (95% CI)	SEM	SDC	LoA
Inter-rater reliability (N = 3)	0.790 (0.615 – 0.893)	3.85	10.67	33.0
Intra-rater reliability (N = 1)	0.965 (0.926 – 0.983)	0.86	2.38	7.03

OMQ = Observable Movement Quality; *n* = number of video-recorded children $ICC_{2,1}$ = intraclass correlation coefficient, a two-way random effects single-measures model of absolute agreement; SEM = standard error of mean; SDC = smallest detectable change; LoA = Limits of Agreement; N = number of observers.

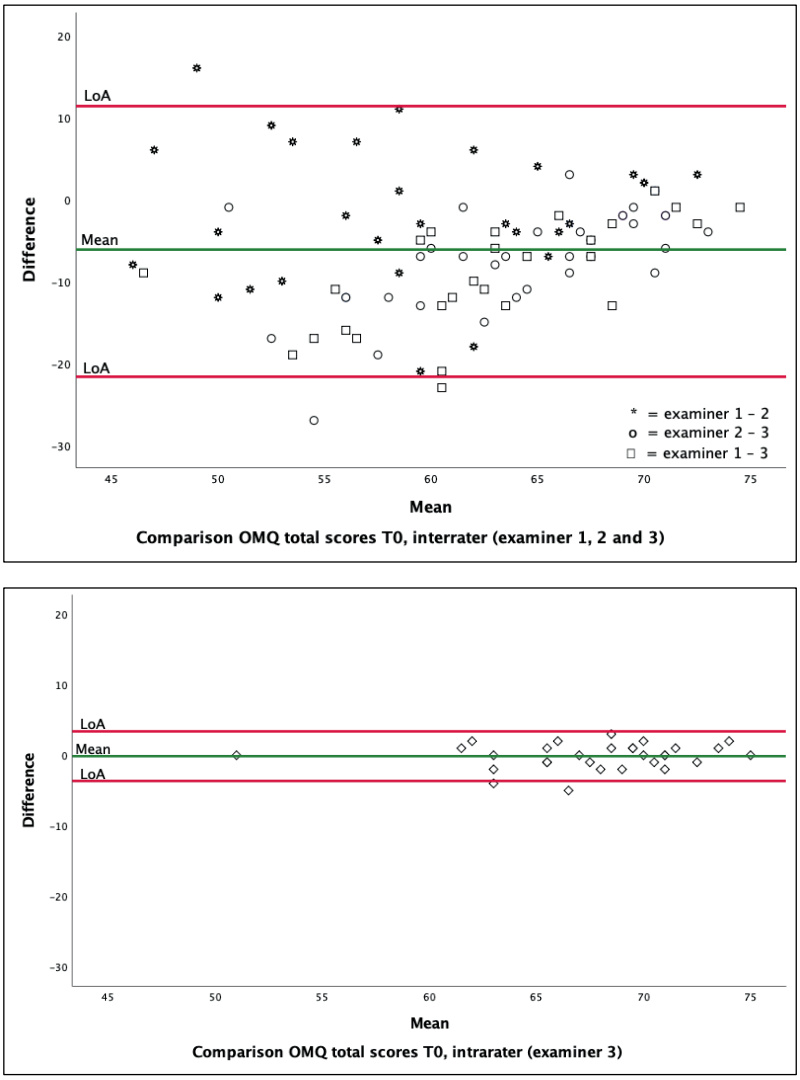


Figure 2. Bland-Altman plots for comparison of Observable Movement Quality scale total scores at baseline assessmen

Table 3. Item Agreement for each item (*n*=15) of the Observable Movement Quality (OMQ) scale scored on a 5-point Likert scale by expert examiner (*n* = 1) for intrarater and by all examiners (*n* = 3) for interrater over video-record children (*n* = 30) for baseline assessments.

		Interrater		Interrater	
	OMQ item	Median (range)	P_O mean (range)	Median (range)	P_O mean (range)
1	Appropriate fine motor movements	4 (2–5)	0.70 (0.33–1.00)	4 (2–5)	0.75 (0.50–1.00)
2	Appropriate gross motor movements	4 (2–5)	0.70 (0.33–1.00)	4 (2–5)	0.90 (0.50–1.00)
3	Fluency of movements	4 (2–5)	0.69 (0.33–1.00)	4 (2–5)	0.78 (0.50–1.00)
4	Reduced muscle tone	5 (2–5)	0.67 (0.33–1.00)	5 (4–5)	0.97 (0.50–1.00)
5	Increased muscle tone	5 (3–5)	0.74 (0.33–1.00)	5 (3–5)	0.97 (0.50–100)
6	Tremors	5 (2–5)	0.82 (0.33–1.00)	5 (2–5)	0.97 (0.50—1.00)
7	Slow and/or delayed movements	5 (2–5)	0.69 (0.33–1.00)	5 (4–5)	1.00 (1.00–1.00)
8	Accelerated and/or abrupt movements	4 (2–5)	0.61 (0.33–1.00)	5 (3–5)	0.88 (0.50–1.00)
9	Asymmetry in movements	5 (2–5)	0.73 (0.33–1.00)	5 (4–5)	0.98 (0.50–1.00)
10	Accuracy (well-aimed)	4 (2–5)	0.62 (0.33–1.00)	5 (3–5)	0.80 (0.50–1.00)
11	Strength regulation	4 (2–5)	0.53 (0.33–0.67)	5 (3–5)	0.88 (0.50–1.00)
12	Variation in movements	4 (2–5)	0.68 (0.33–1.00)	4 (3–5)	0.87 (0.50–1.00)
13	Involuntary movements	4 (1–5)	0.61 (0.33–1.00)	4 (3–5)	0.80 (0.50–1.00)
14	Automated movements	4 (1–5)	0.67 (0.33–1.00)	4 (2–5)	0.88 (0.50–1.00)
15	Stereotype movements	5 (2–5)	0.67 (0.33–1.00)	5 (4–5)	0.98 (0.50–1.00)

[a] Linear weighting, CI = Confidence Interval, Po =proportion of observed agreement

agreement the proportions of observed agreement were moderate to good; 0.53–0.83 and 0.75–1.00, respectively.

The overall effect of paediatric physiotherapy intervention is presented in Table 4. ES and SRM values for OMQ scale total scores between baseline assessment and assessment after intervention reflect small to moderate effects (0.48 and 0.73, respectively). For MABC-2-NL, ES and SRM values between baseline assessment and assessment after intervention reflect small effects (0.44 and 0.43, respectively). Outcomes for responsiveness of the OMQ scale are also presented in Table 4. The SDC at the 95% confidence interval for the OMQ scale was 2.38, implying that a change of 2 points or more is likely to represent true change in movement quality as measured by the OMQ scale.

According to the GPE scores, 46% of the children (*n* = 13) were categorized as improved and 54% (*n* = 15) as unchanged for motor quality. The MIC for the OMQ scale total score was identified as 3.15, implying that a change of 3 points or more is likely to represent a therapist-perceived important change by the OMQ scale (see Table 4). A two-sample t-test, applied to the mean change scores for the OMQ scale between the improved and unimproved groups, revealed significant difference (p = 0.009), with the improved group scoring higher than the unimproved group. The MIC calculated from the ROC curve using the cut-off point nearest the upper left-hand corner of the graph was 2.5 points for OMQ scale total scores (sensitivity 84%, specificity 77%); the AUC for change in OMQ scale total score was 0.77.

Table 4. Responsiveness statistics for Observable Movement Quality (OMQ) scale total score (*n* = 28) and Movement Assessment Battery for Children, 2nd edition, Dutch version (MABC-2-NL) standard scores.

	Mean difference	(SD)	95% CI	Range	sign	SEM	ES	SRM	SDC	MIC	ROC cut-off
OMQ scale											
Total score	2.39	(3.28)	1.12 – 3.67	-2 – 11	0.001	0.86	0.48	0.73	2.38	3.15	2.50
MABC-2-NL											
MABC-2-NL standard score	1.50	(3.46)	0.16 – 2.84	-5 – 11	0.030	NA	0.44	0.43	NA	NA	NA

SD = Standard Deviation; CI = confidence interval; sign = significance; SEM = Standard Error of Measurement; ES = Effect Size; SRM = Standardized Response Mean; SDC = Smallest Detectable Change; MIC = Minimal Important Difference; ROC = Receiver Operating Curve; NA = not applicable

Discussion

In this prospective intervention study, the reliability and responsiveness of the OMQ scale was determined in physiotherapists' daily practice with children from 4 to 12 years of age as participants. This study showed that the OMQ scale is a reliable and valid measurement instrument to assess movement quality in clinical practice and to monitor and evaluate movement quality as a result of the treatment's progress. The OMQ showed a moderate to good interrater reliability and high intrarater reliability, with excellent item agreement. Our study showed, furthermore, a SEM of 0.62 for OMQ scale total scores, SDC of 2.38—both based on the intrarater scores—and MIC of 3.15. The MIC calculated using a ROC curve was 2.5. Because the MIC should be detectable beyond measurement error,[39,40] and above the SDC,[26] our research showed that it is possible for the OMQ scale to detect change in movement quality among children from 4 to 12 years of age.

Responsiveness of the OMQ scale was assessed using a GPE as a comparator instrument to measure change in movement quality. This was chosen because a construct approach—in which a priori hypotheses of expected associations between scores of the OMQ scale and other assessment tools that measure more or less the same construct would be assessed—was not possible.[15,16] The reason for development of the OMQ scale was the lack of a generic measurement instrument to assess movement quality in children.[13] Therefore, no hypothesis for expected correlations between changes in scores on the OMQ scale and those on other similar instruments could be set. Using the GPE as a comparator instrument is the most common external criterion.[41] Furthermore, we used a 7-point transition question, focusing on change in movement quality, as recommended.[27] However, when scoring a GPE, patients are known to have difficulty taking their baseline status into account; as such, GPE ratings are strongly influenced by patients' current health status.[28,40,42–44] Moreover, the MIC depends significantly on the anchor's definition of important change.[40]

In our study, we decided not to ask the parents of the children to rate the perceived change in movement quality, because we anticipated that they would have difficulties estimating changes in movement quality; a professional concept. Above all, we expected the parents to be influenced by the current health status of their child or even to want to please the physiotherapist by saying their child had improved.[40] Therefore, we decided to ask the treating physiotherapists to rate the change on the GPE ratings scale before the start of the assessment after intervention. This allowed them a perspective on what they would consider important improvement or deterioration, although some practitioners' bias could have influenced the rating on the GPE.[45]

In this study, we used ROC curves to examine cut-off values for the OMQ scale change scores.[18] The perceived improvement of movement quality on the GPE was used as an anchor,[26] which required the choice of a sensible cut-off point of important change.[46] There is debate about whether the category 'a little improved' should be considered as change.[19,33,40,47] We concluded that it should not, in accordance with Demoulin et. al.,[47] who

stated that the accuracy to differentiate patients who improved from those who did not will decrease if patients who report little improvement are considered as improved. By not including patients who were 'a little improved' in the group classified as 'improved', we assumed the concept of important improvement was better reflected.

The responsiveness and MIC of measurement instruments are often population- and context-specific and should be taken into account before generalizing to other populations.[40] A limitation of our study for the measurement properties of the OMQ scale was that these have so far only been examined in Dutch paediatric physiotherapy practices. Consequently, it is necessary to investigate whether the measurement properties are consistent with other countries and other populations of children, and therefore we should assess whether they adequately reflect the purpose of the OMQ scale.[16] Future studies for the OMQ scale, to include data of children treated in multiple paediatric physiotherapy contexts, in other countries and within other patient populations (e.g., neuromuscular diseases and syndromes), will provide further evidence of validity for the use of the OMQ scale in clinical practice.

A limitation of our study was that the SDC was derived from intrarater reliability measures and not from test-retest situations. Therefore, the SDC from our study could be an underestimation as it only considers the examiner as a source of variance. Another limitation of our study was the inclusion of 30 children, while sample size calculation indicated 31 video recorded children. For our study, we were able to recruit 32 children; two children refused to participate in second thought, unfortunately. Due to the duration of our study, it was not possible to include the indicated 31 children. However, we expect that this did not affect the results of our study. Also, a limitation of our study was the inclusion of only 1 examiner for intrarater reliability and 3 examiners to establish interrater reliability. This decision was based on the time investment for the physiotherapists to observe the video-recorded children. The inclusion of only 3 examiners could possibly have contributed to the somewhat higher outcomes for interrater reliability and item agreement demonstrated by our present study compared to those in our previous study—in which 28 paediatric physiotherapists also unfamiliar with the OMQ and a short introduction observed video recordings of 9 children from 6 months to 6 years of age.[14] Although, these higher outcomes can be related to differences in the population, as for example the age, as well. Also, a limitation was the inclusion of only children aged 4 to 12 years. In our former study for interrater reliability of the OMQ scale, we included children from 6 months to 6 years of age,[14] and measurement properties for children within the age frame of 12–16 years have not yet been investigated. A future study including a larger group of children—also of older ages—within other patient populations, and with intervention periods over a 6-month period, will benefit the generalizability of the results.

Conclusion

The OMQ scale demonstrates a moderate to good interrater reliability and high intrarater reliability when used by paediatric physiotherapists to assess movement quality in children aged 4 to 12 years with mild to moderate motor impairments. Our findings show, furthermore, that the OMQ scale is responsive to change when used for children treated in daily paediatric physiotherapist practice, although only a small change in motor performance was seen within our study population. Our findings on reliability and responsiveness indicate that the OMQ scale can be used in daily clinical practice. Moreover, our findings show that an assessment with the OMQ scale—which is completed in approximately 5–10 minutes— is complementary to outcomes for motor performances tests. Training in the use of the OMQ scale is recommended to ensure reliable scoring, which will be developed after completing the validation of the OMQ scale.

Funding details
The Dutch Research Council (NWO) supported Lieke Dekkers: grant 023.004.037. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgements
We would like to sincerely thank all the children, their parents, and the paediatric physiotherapists for participating in this study.

References

1. Knudson D V, Morrison CS. *Qualitative analysis of human movement.* 2nd ed. Leeds: Human Kinetics; 2002.
2. Heineman KR, Hadders-Algra M. Evaluation of neuromotor function in infancy-A systematic review of available methods. *J Dev Behav Pediatr.* 2008;29(4):315-323.
3. Harris SR, Heriza CB. Measuring infant movement. Clinical and technological assessment techniques. *Phys Ther.* 1987;67(12):1877-1880.
4. Sorsdahl AB, Moe-Nilssen R, Kaale HK, Rieber J, Strand LI. Change in basic motor abilities, quality of movement and everyday activities following intensive, goal-directed, activity-focused physiotherapy in a group setting for children with cerebral palsy. *BMC Pediatr.* 2010;10:26.
5. Wright FV, Rosenbaum P, Fehlings D, Mesterman R, Breuer U, Kim M. The Quality Function Measure: reliability and discriminant validity of a new measure of quality of gross motor movement in ambulatory children with cerebral palsy. *Dev Med Child Neurol.* 2014;56(8):770-778.
6. Skjaerven LH, Kristoffersen K, Gard G. An eye for movement quality: a phenomenological study of movement quality reflecting a group of physiotherapists' understanding of the phenomenon. *Physiother Theory Pract.* 2008;24(1):13-27.
7. Magill R, Anderson D. Motor Control Theories. In: *Motor learning and control. Concepts and applications.* 10th editi. Singapore: McGraw Hill Education; 2014.
8. Boyce WF, Gowland C, Hardy S, et al. Development of a quality-of-movement measure for children with cerebral palsy. *Phys Ther.* 1991;71(11):820-828; discussion 828-832.
9. Schmidt RA, Wrisberg CA. Getting started. In: *Motor learning and performance, a situation-based learning approach.* 4th editio. Leeds: Human Kinetics; 2008:3-23.
10. Hadders-Algra M. General movements: A window for early identification of children at high risk for developmental disorders. *J Pediatr.* 2004;145(2 SUPPL.):12-18.
11. Hecker E, Baer GD, Stark C, Herkenrath P, Hadders-Algra M. Inter- and intrarater reliability of the Infant Motor Profile in 3- to 18-months-old infants. *Pediatr Phys Ther.* 2016;28(2):217-222.
12. DeMatteao C, Law M, Russell D, Pollock N, Rosenbaum P, Walter S. The relialbilty and validity of the Quality of Upper Extremity Skills Test. *Phys Occup Ther Pediatr.* 1993;13(2):1-18.
13. Janssen AJWM, Diekema ETW, Dolder van R, Kollée LAA, Oostendorp RAB, Nijhuis-van derSanden MWG. Development of a movement quality measurement tool for children. *Phys Ther.* 2012;92:574-594.
14. Dekkers LMA, Nijhuis-Van Der Sanden MWG, Jonker M, De Swart BJM, Janssen AJWM. Interrater reliability of the observable movement quality scale for children. *Physiother Canada.* 2018;70(2):113-119.
15. Mokkink LB., de Vet HCW, Prinsen CAC, et al. COSMIN risk of bias checklist for systematic reviews of Patient-Reported Outcome Measures. *Qual Life Res.* 2018;27(5):1171-1179.

16. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol.* 2010;63(7):737-745.
17. Portney LG, Watkins MP. *Foundations of clinical research, applications to practice.* Third Edit. London: Pearson Education International; 2009.
18. de Vet HC, Terwee CB, Ostelo RW, Beckerman H, Knol DL, Bouter LM. Minimal changes in health status questionnaires: Distinction between minimally detectable change and minimally important change. *Health Qual Life Outcomes.* 2006;4(Mic):3-7.
19. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials.* 1989;10(4):407-415.
20. Revicki DA, Cella D, Hays RD, Sloan JA, Lenderking WR, Aaronson NK. Responsiveness and minimal important differences for patient reported outcomes. *Health Qual Life Outcomes.* 2006;4:1-5.
21. Mokkink LB, Prinsen CAC, Bouter LM, de Vet HCW, Terwee CB. The COnsensus-based Standards for the selection of health measurement INstruments (COSMIN) and how to select an outcome measurement instrument. *Brazilian J Phys Ther.* 2016;20(2):105-113.
22. Rysstad T, Røe Y, Haldorsen B, Svege I, Strand LI. Responsiveness and minimal important change of the Norwegian version of the Disabilities of the Arm, Shoulder and Hand questionnaire (DASH) in patients with subacromial pain syndrome. *BMC Musculoskelet Disord.* 2017;18(1):1-9.
23. Henderson SE, Sugden DA, Barnett LA. *Movement Assessment Battery for Children- second edition (MABC-2).* London: Pearson Education International; 2010.
24. Smits-Engelsman B. Movement Assessment Battery for Children-2. Dutch standardization. Amsterdam: Pearson Education; 2010.
25. Kamper SJ, Maher CG, Mackay G. Global Rating of Change scales: a review of strengths and weaknesses and considerations for design. *J Man Manip Ther.* 2009;17(3):163-170.
26. de Vet HCW, Terwee CB, Mokkink LB, Knol DL. *Measurement in medicine.* Cambridge: Cambridge university press; 2011.
27. Lauridsen HH, Hartvigsen J, Korsholm L, Grunnet-Nilsson N, Manniche C. Choice of external criteria in back pain research: Does it matter? Recommendations based on analysis of responsiveness. *Pain.* 2007;131(1-2):112-120.
28. Kamper SJ, Ostelo RWJG, Knol DL, Maher CG, de Vet HCW, Hancock MJ. Global Perceived Effect scales provided reliable assessments of health transition in people with musculoskeletal disorders, but ratings are strongly influenced by current status. *J Clin Epidemiol.* 2010;63(7).
29. McCarthy ML, Silberstein CE, Atkins EA, Harryman SE, Sponseller PD, Hadley-Miller N a. Comparing reliability and validity of pediatric instruments for measuring health and well-being of children with spastic cerebral palsy. *Dev Med Child Neurol.* 2002;44(7):468-476.
30. Wuang Y, Su J, Su C. Reliability and responsiveness of the Movement Assessment Battery for Children – second edition test in children with developmental coordination disorder. *Dev Med Child Neurol.* 2012;54:160-165. d
31. Koo TK, Li MY. A Guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med.* 2016;15(2):155-163.

32. Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60(1):34-42.
33. Wyrwich KW, Nienaber NA, Tierney WM, Wolinsky FD. Linking clinical relevance and statistical significance in evaluating intra-individual changes in health-related quality of life. *Med Care.* 1999;37(5):469-478.
34. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986;1(8476):307-310.
35. Shoukri M. Sample size for a reliability study. Stat Meth Med Res. 2004;13:251-271.
36. Stratford, Paul W, Binkley, Jill M, Riddle, Daniel L. Health status measures: strategies and analytic methods for assessing change scores. *Phys Ther.* 1996;76(10):1109-1123.
37. Cohen J. *Stastical Power Analysis for the Behavioural Scienses.* New York: Academic Press; 1977.
38. Van Kampen DA, Willems WJ, van Beers LWAH, Castelein RM, Scholtes VAB, Terwee CB. Determination and comparison of the smallest detectable change (SDC) and the minimal important change (MIC) of four-shoulder patient-reported outcome measures (PROMs). *J Orthop Surg Res.* 2013;8(1).
39. Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. *J Clin Epidemiol.* 2003;56(5):395-407.
40. De Vet HCW, Ostelo RWJG, Terwee CB, et al. Minimally important change determined by a visual method integrating an anchor-based and a distribution-based approach. *Qual Life Res.* 2007;16(1):131-142.
41. Ogollah R, Bishop A, Lewis M, Grotle M, Foster N. Responsiveness and minimal important change for pain and disability outcome measures in pregnancy-related low back and pelvic girdle pain. *Physiotherapy.* 2019;99(11):1551-1561.
42. De Vet HCW, Foumani M, Scholten MA, et al. Minimally important change values of a measurement instrument depend more on baseline values than on the type of intervention. *J Clin Epidemiol.* 2015;68(5):518-524.
43. Grøvle L, Haugen AJ, Hasvik E, Natvig B, Brox JI, Grotle M. Patients' ratings of global perceived change during 2 years were strongly influenced by the current health status. *J Clin Epidemiol.* 2014;67(5):508-515.
44. Schuller W, Ostelo RWJG, Janssen R, de Vet HCW. The influence of study population and definition of improvement on the smallest detectable change and the minimal important change of the neck disability index. *Health Qual Life Outcomes.* 2014;12(1):1-7.
45. van Walraven C, Mahon JL, Moher D, Bohm C, Laupacis A. Surveying physicians to determine the Minimal Important Difference. *J Clin Epidemiol.* 1999;52(8):717-723.
46. Stuge B, Krogstad Jenssen H, Grotle M. The Pelvic Girdle Questionnaire: responsiveness and minimal important change in women with pregnacy-related pelvic girdle pain, low back pain, or both. *Phys Ther.* 2017;97(11):1-8.
47. Demoulin C, Ostelo R, Knottnerus JA, Smeets RJEM. What factors influence the measurement properties of the Roland-Morris disability questionnaire? *Eur J Pain.* 2010;14(2):200-206.

Part II

Observations of movement quality

Chapter 6

Educational programs for learning to observe movement quality in physiotherapy; a design-based research approach

Lieke MA Dekkers

Ton Satink

Maria WG Nijhuis-van der Sanden

Bert JM de Swart

Marjo JM Maas

Anjo JWM Janssen

Published ahead of print in: Physiotherapy Theory and Practice,
DOI: 10.1080/09593985.2020.1712754

Abstract

Introduction

Movement observation is a core aspect in physiotherapists' diagnosis to determine which interventions are adequate to improve functional abilities. The aim of this study was to derive design principles for an educational program for the development of observational skills.

Methods

We used a qualitative approach within a design-based research methodology. In four rounds, eight physiotherapy students, 16 teachers and nine practitioners participated in five Nominal Group Technique meetings and six interviews. Meetings and interviews were transcribed verbatim and analysed using thematic analysis.

Results

We identified three themes, each with several design principles: didactics, professional content and conditions for optimal learning. We developed a proto-theory underpinned with underlying educational theories.

Conclusion

To learn observational skills, students, facilitated by an experienced teacher, need to take the lead in their own learning process. This might imply a need for additional training for teachers. A realistic context is a precondition for learning; it might be necessary to increase possibilities for observations in clinical contexts or to invest in training for (simulated) patients as participants in education. Further research is needed to test the applicability of the design principles and a proto-theory for other professionals with a focus on observation and analysis of movements.

Introduction

A core aspect in diagnosis in rehabilitation is the observation of movements as a basis for interventions to improve functional abilities.[1] During assessment, a physiotherapist observes the quantity and quality of movements. The quantity of movements reflects the acquisition or re-acquisition of new or more complex motor skills.[2] The quality of movement gives an impression of how movements are controlled and coordinated.[3] In this way, movement quality represents the interaction between personal characteristics and experiences, the task difficulty, and the environmental conditions; and it gives insight into the possibilities and potential of the person's system for reacting or adapting to changing conditions.[1] For the assessment of movement quality which is important for recognizing motor problems, designing and evaluating interventions, and predicting recovery[1,4] physiotherapists must rely on observational skills. The observation process involves gathering, organizing, and giving meaning to visual, auditory, and sensory information obtained while observing the moving person.[1,4] Earlier studies[1,5] found that how physiotherapists describe their observations is not standardized; it varies among therapists, depending on the theoretical constructs used, which precludes comparability and longitudinal evaluation.

Observation is a fundamental skill for physiotherapists.[1] Curricula for bachelor students in physiotherapy address knowledge and observational skills focused on the measurement and interpretation of the quantity and quality of human movements. However, there is disagreement about the details of what constitutes good clinical observation, its conceptual basis, and how it is learned or developed.[6] General principles for observation were described by Boudreau et. al.[6] as follows: observation has a sensory perceptive and cognitive component, observation is distinct from inference and made concrete through description, and observation is goal oriented, occurs over time, and on different levels. These levels for observation refer to the whole person observed, a body part, the personal or environmental context, behaviours and interactions, and the characteristics of the observer on, e.g., emotional and aesthetic planes.[6,7] It seems evident that the observer should be considered an influencing factor when teaching and evaluating observational skills[8,9] because there is a tendency for perception, interpretation of what is seen, to interfere with observation, and initial observation should be without any judgment.[6,10]

A clear description of how teachers and students in physiotherapy perceive learning observational skills and what didactic principles facilitate learning have not been widely investigated.[11] Teaching observational skills, seems to depend on personal experiences, leading to a variety of information presented and instruction provided, which could cause uncertainty in students. The literature lacks specific educational strategies to support the development of skills for observing movement quality.[8] Existing literature on observational skills learning mainly focuses on learning in the domain of medical education and nursing,[6–11] not on guiding principles for teaching observational skills in physiotherapy. The overall conclusion for the training approaches under study was that bachelor and master students'

observational skill improved in the educational setting, but not in real and complex clinical situations where the incidence of perception failure may increase. In addition to this, currently available research is predominantly limited to learning strategies using static images, as was indicated in the scoping review of Al-Moteri et. al.,[8] while students in physiotherapy need to learn to observe movement. The overall aim of our study was to derive design principles for an educational program to develop observational skills for students in physiotherapy. In addition, we were interested in developing a proto-theory (i.e. a set of theoretical concepts that guides ongoing development and refinement of the educational design) for educational program design for the development of observational skills.

Method

Design-based research

Design-based research is an important methodology for understanding how, when, and why educational innovations work in practice.[12] The design-based methodology is an accepted qualitative research approach in educational sciences, and it triangulates multiple sources of evidence.[12–17] Wang and Hannafin[18] described the definition of design-based research as follows:

> *A systematic, but flexible methodology aimed to improve educational practices through iterative analysis, design, development, and implementation, based on collaboration among researchers and practitioners in a real world setting, and leading to contextually-sensitive design principles and theories. (pp. 6–7)*

The design-based research methodology was applied throughout our study to derive design principles for an educational program to develop observational skills in physiotherapy students. This approach allows for collaboration among students, teachers, practitioners, and researchers, the main stakeholders in our study. Furthermore, the method allows for conducting reflective analyses to iteratively test and refine innovative learning environments as well as to define design principles.[19,20] We followed four sequential steps as stated by Reeves[19] and modified the sequential steps according to our research question, described in figure 1.

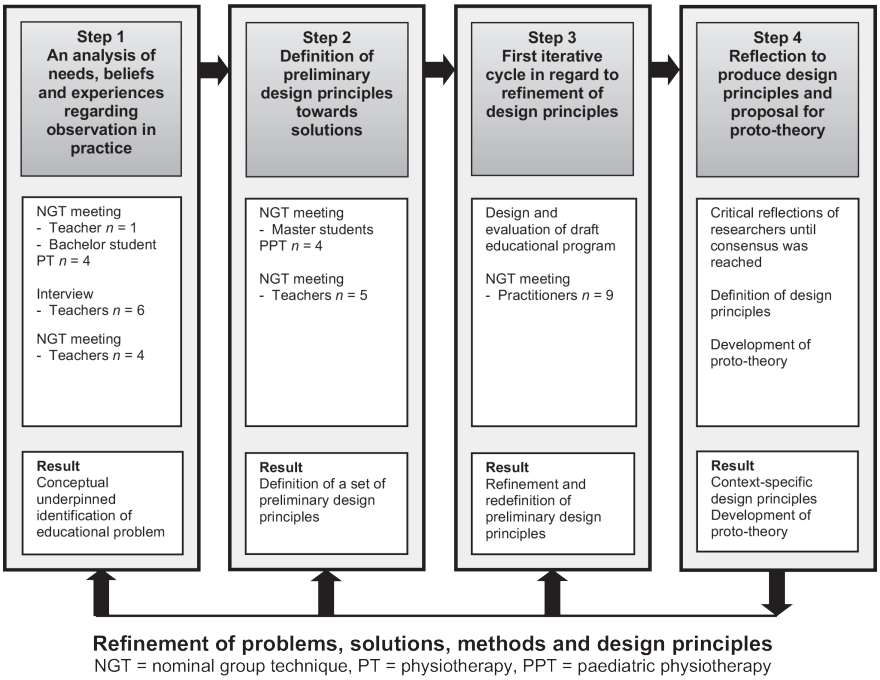


Figure 1. Design-based research, based on Reeves (2006)

Approach and selection of participants

To collect and derive design principles for an educational program, stakeholders from both education and physiotherapeutic practice were asked to participate. We aimed at a maximum of variation in the participants regarding the level of professional development and expertise. Furthermore, we collaborated with various universities of applied sciences in the Netherlands to ensure the integration of different points of view:

(a) Bachelor students in physiotherapy. They must learn to observe movement quality.
(b) Master students in paediatric physiotherapy. They have already gained some experience in observing movement quality. However, to observe the quality of movement in paediatrics, students must understand typical motor development to identify atypical movement quality, which creates an extra challenge in mastering observational skills.
(c) Teachers of physiotherapy who have experience in teaching students to observe movement quality and have knowledge and experience in different teaching strategies.
(d) Practitioners currently working in the field of paediatric physiotherapy. They have experience in developing and mastering observational skills.

Bachelor and master students and teachers were invited by mail to participate in our study. Practitioners were verbally informed about and invited to participate in our study during a post-graduate module to which they were invited. All interested participants received oral and written information about our study, explaining the aim and the total time investment of about 1 to 2 hours. The study was performed in accordance with the Declaration of Helsinki. This study was submitted to the Ethical Advisory Committee of the Faculty of Health at the HAN University of Applied Sciences, Nijmegen, the Netherlands. The Committee reviewed and discussed our research proposal including the consent form for participants. They approved our study, deeming it exempt from further review (registration: EACO17.12/90). All participants volunteered to participate, and anonymity and confidentiality were assured. The participants signed informed consent documents.

Data collection

In design-based research, different research methods can be used.[15,16,21] We used semi-structured interviews and focus group meetings, using the Nominal Group Technique (NGT), which encouraged contributions from everyone.[22,23] For each interview or focus group meeting, new participants were invited, to ensure maximum variety in perspectives, using purposive sampling. The researchers: LD, paediatric physiotherapist and educational scientist, and AJ, paediatric physiotherapist developed an interview guide; research assistants who were trained in qualitative research conducted the interviews and took notes. Subsequently, the researchers prepared five NGT meetings; the first three meetings were conducted by research assistants and supervised by the principal researcher (LD). The principal researcher (LD) conducted the final NGT meeting; a second researcher (AJ) took notes. Both the interviews and NGT meetings took place in university settings. In our study, data collection and analysis was an ongoing iterative process guided by the steps of design-based research.[13,19] The aim, methods, and participants for each step are described below

Step 1: An analysis of needs, beliefs, and experiences regarding observation in practice

The aim was to investigate and analyse the educational needs, beliefs, and experiences, and possible challenges, in teaching and learning methods identified and prioritized by participants. A global search for existing literature on learning observational skills was performed prior to designing an NGT meeting. The NGT meeting was pre-structured and lasted for a maximum of 2 hours. The NGT meeting (n = 5) included one teacher specialized in neurologic physiotherapy, and four bachelor students in physiotherapy all in their fourth and final year of training. During this meeting, information was generated by the participants responding to this question: '*What are the best ways to learn to observe movement quality for bachelor students in physiotherapy?*' The group meeting was followed by six semi-structured interviews (n = 6) with teachers of physiotherapy from different universities of applied sciences from the middle and south regions of the Netherlands to further investigate the generated needs, beliefs, and experiences in teaching and learning

methods for observational skills. The teachers were experts in different domains of clinical practice: musculoskeletal physiotherapy (*n* = 3), neurological physiotherapy (*n* = 2), and sports physiotherapy (*n* = 1). Following the interviews, another NGT meeting was conducted with four teachers of physiotherapy (*n* = 4) from a university in the southeast region of the Netherlands, with the following professional expertise: musculoskeletal physiotherapy (*n* = 3), and neurologic physiotherapy (*n* = 1). Participants responded to the following question: *'What do you think are critical elements for bachelor students in physiotherapy for learning to observe movement quality in the classroom context?'* Through the triangulation of data resources, existing literature, and data from semi-structured interviews and focus groups, a conceptual underpinning and identification of the educational problem was defined at the end of the first step in our study.

Step 2: Definition of preliminary design principles towards solutions

The aim of step 2 was to define a preliminary set of design principles to guide the design of the educational program for learning observational skills as a solution to the identified problem. To define a preliminary set of design principles, two separate NGT meetings were conducted together with the different stakeholders. Those NGT meetings enhanced in-depth exploration of both students' and teachers' needs, beliefs, and experiences in teaching and learning methods for observational skills. Unclear details about needs, beliefs, and experiences, and differences in perspectives on teaching and learning methods among the participants were explained and discussed in the group during the NGT meeting and combined into categories, without the elimination of ideas. We conducted one NGT meeting with four master students of paediatric physiotherapy (*n* = 4) and another meeting with five teachers at a university of applied sciences in physiotherapy (*n* = 5), which professional expertise in musculoskeletal physiotherapy (*n* = 2), neurologic physiotherapy (*n* = 2), and sports physiotherapy (*n* = 1). The question asked was this: '*What is the best method to learn to observe quality of movement, and what supporting materials are needed?'* All master students were educated at a university of applied sciences in the middle of the Netherlands, whereas their former bachelor education in physiotherapy was conducted at different universities of applied sciences based in the north, middle, and southeast regions of the Netherlands. The university teachers worked at the same university of applied sciences in the southeast region of the Netherlands. A set of preliminary design principles was developed at the end of step 2.

Step 3: First iterative cycle in regard to refinement of design principles

A draft educational program was designed based on the preliminary set of design principles and evaluated in an authentic setting in clinical practice. Nine paediatric physiotherapists (*n* = 9) were trained in observing human movement using the draft educational program in a postgraduate program at a university for paediatric physiotherapy. At the end of the program, an NGT meeting was conducted to explore participants' perceptions of the new

educational program. The participants were asked to respond to the following question: *'What are important elements for you in an educational program to learn how to observe movement quality?'* The draft versions of the manual and the supplemental materials were discussed within the study group and revised to meet the needs of all the different participants at the end of step 3. Subsequently, the preliminary set of design principles was refined and redefined.

Step 4: Reflection to produce design principles and proposal for a proto-theory
As the Design-based Research Collective[12] noted, it is important that design-based research does not end with designing and testing particular interventions; rather, it should lead to a shareable 'proto-theory'. This proto-theory aims to support researchers and practitioners in the ongoing development of the educational design.[12] To attain the final goal of our study, we reflected on our overall research procedure, on the results leading to the development of the educational materials, and on the content and use of the materials. During our reflections, we focused on theoretical understanding, which resulted in suggestions for context-specific theoretical design principles. Finally, step 4 was concluded with the design of a proto-theory.

Data analysis
Interviews and NGT meetings were video-recorded and transcribed verbatim. The transcripts were analysed by LD and TS using ATLAS.ti version 8 (ATLAS.ti Scientific Software Development GmbH, Berlin). For each phase of the analysis process, the six steps of thematic analyses, as described by Braun and Clarke[24] were followed (Table 1). The transcripts were coded by LD and TS; codes were discussed until consensus was reached. To enhance credibility, IG, who also specializes in qualitative research methods, reviewed the analysis process. The initial codes were grouped and resulted in a data matrix. Potential themes and preliminary design principles emerged from the data through constant comparison of codes and themes. No contradictory codes or themes needed to be accommodated. The names and definitions of themes and their design principles were discussed with the research team until consensus was reached. During the development of the proto-theory, the construction of a model that depicted the interrelation between the design principles helped us in the further discussion about underlying educational theories. Finally, a proto-theory was developed that reflects the interrelations between the identified design principles and the existing theories on learning observational skills.

Table 1. Phases of thematic analysis, based on Braun and Clarke (2006)

Phase of thematic analysis	Aim	Result
1. Familiarization with the data	Identifying meaningful units of text relevant to learning and teaching observational skills	Initial ideas were noted down
2. Generating initial codes	Grouping together units of text dealing with the same issue	Provisional definitions were given
3. Searching for themes	Different codes were sorted into potential themes regarding learning and teaching principles in learning the observational skill	Collating all the relevant coded data extracts within the identified themes
4. Reviewing themes (based on potential themes in previous steps in the process of design-based research)	Checking if the themes work in relation to the coded extracts and the entire data-set	Thematic map of the analyses resulting in preliminary design principles for our conceptual model
5. Defining and naming themes	Refining the specifics of each theme and sub-theme	Generating clear definitions and names for each theme for the design principles
6. Producing final recommendations for the proto-theory and report	Final analysis; relating back of analysis to the aim of our project and to literature	A proto-theory with explanation and a justification in the report for our project

Results

The overall aim of our design-based research was to derive design principles for an educational program to develop observational skills for physiotherapy students. Our design-based research resulted in the identification of three themes, each of which comprised a cluster of design principles, which are described below and supported by quotes (labelled by participant's role [S=Student, MS=Master Student, T=Teacher, P=Practitioner]). To facilitate students' learning of observational skills, the design of the educational program includes the following:

(1) Didactics: How to learn observational skills, referring to different aspects of student learning, including theoretically, and underpinning perspectives on student learning
(2) Professional content: What needs to be mastered, referring to the understanding of human movement and observation
(3) Conditions for optimal learning: Referring to recommendations for physiotherapists that enable good qualitative observations.

Figure 2 shows the interrelation between the themes and the design principles in the development of an educational program for observational skills. Furthermore, a proto-theory (Table 2) was developed that displays the didactic principles with their underlying educational theories. We will start with the description of the themes and design principles and end with the description of the proto-theory.

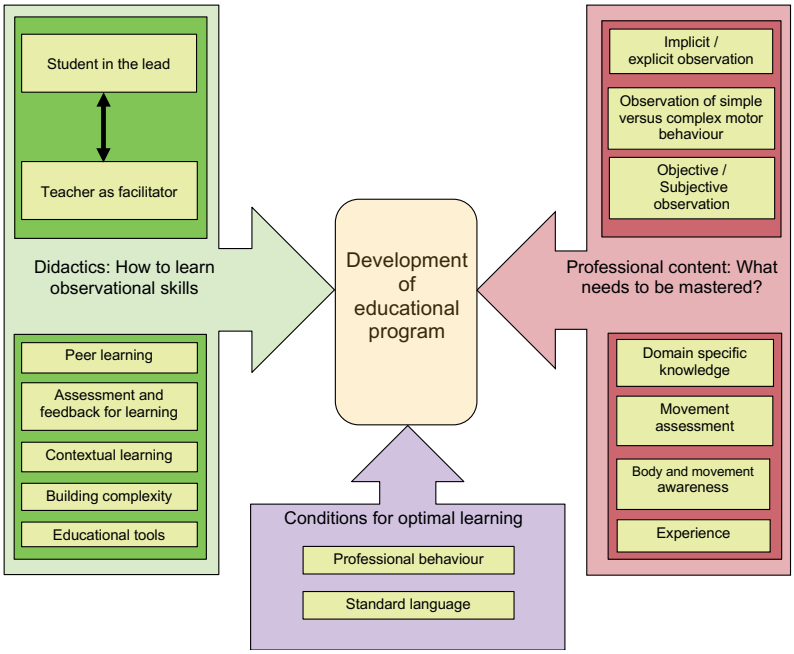


Figure 2. Interrelation between design principles and themes for development of an educational program

Theme A: Didactics: How to learn observational skills

Students, teachers, and practitioners described different ways to facilitate the learning of observational skills. Comments by students and practitioners focused on the process of gathering knowledge, skills, and attitudes from their individual points of view, whereas the lecturers' focus reflected a broader perspective. This perspective included facilitation by teachers, peer students, and students' self-regulation. Seven interrelated design principles were defined for this.

Students in the lead

Students reported that learning to observe human movement requires a sense of urgency in terms of closing gaps in their knowledge and skills. This perceived urgency gives meaning to learning and promotes involvement and an intrinsic motivation to learn. This can be conceptualized as 'taking the lead' or self-regulation.

> *When I observe a patient with a complex health problem, I experience my shortcomings in knowledge and skills. That motivates me to find out what I actually need. (MS-4)*

Students and teachers explained that self-regulation also involves reflection in action, reflection on action, and setting personal learning goals.

> *You just have to practice! At the start I just thought . . . what do I actually see? And when a teacher said what he saw, I thought: 'No way! I didn't see that at all'. But now I do, because I practiced'. (MS-2)*

Teacher as facilitator

Teachers described how they facilitate the learning processes of their students. They emphasized the importance of interacting with students to monitor their learning process and to identify their learning needs.

> *Working with students requires interaction. This provides you with the necessary insights on their current performance level, whether they actually grasp it, or if you have to repeat some of the steps before a step forwards can be made. I have learned you cannot standardize teaching strategies. (T-3)*

Another teacher commented:

> *It has to make sense for the student; only then does it become important. If you can connect the problems they encounter in their learning and what is important for them to learn, that's when learning occurs. (T-6)*

In addition, teachers explained how they facilitate students' taking responsibility for their own learning process by supporting conscious goal setting and action planning.

> *As a teacher, you make sure a student sets personal learning goals related to what they already have learned. You give them their own responsibility to build their body of knowledge and skills. (T-8)*

Teachers reported that students experience observing human movement as difficult. They can watch, but they cannot automatically give meaning to what they see. They need help. Guided instruction, deep questioning, and performance feedback on the observed are the strategies teachers apply to give meaning to the observations. Scaffolding is used as a strategy to facilitate students in reflecting critically on what they have observed, independent of teacher involvement.

> *Learning to observe is very difficult. You have to offer students a framework to guide them. You start to instruct what to observe. Next you ask, 'What did you observe, and what did you notice?' Subsequently you give feedback so that*

> *they become triggered to observe independently, being able to notice without being pointed at. (T-7)*

Thus, the necessity for variation and flexibility in both learning environments and strategies were mentioned as important for encouraging learning.

> *Variation in the educational process: that is what I aspire to. Sometimes I start with performing an observation in a classroom setting; the next time, I let students start with an active observation of a patient. The same with information: sometimes I give necessary information beforehand; other times, I let them search for necessary information and help them to find this later. (T-2)*

Peer learning

Learning with and from each other as peers was described as an important aspect of learning. This was perceived as more fun than individual learning. However, explaining ideas to peers, talking about and questioning each other's views, was perceived as the most important benefit for learning.

> *I learned a lot from working with peers because you have to explain what you think and incorporate the theoretical background in your explanation. (S-1)*

An additional benefit for peer learning was mentioned by lecturers: students learn to understand existing differences of opinion between students and between professionals.

> *The goal is not to state what is correct and what is wrong, but just to recognize and acknowledge the differences. The students have to experience that there are differences and try to figure out what the differences are and why they exist. (T-7)*

In the following quote, a master student describes how his observational skills benefit from explaining to peers what is observed, describing the transfer of domain-specific knowledge to the context of observing human movement.

> *For me, learning skills includes explanations to others. By explaining to others, I have to truly understand the theoretical backgrounds; only by understanding can I explain. (MS-4)*

Assessment and feedback for learning

All participants mentioned the benefits of assessments, both formative and summative, to enhance observational skills learning. Teachers described the valuable impact of assessments, such as performance assessments, peer assessments, and case-based tests, on clinical reasoning. Assessment triggers discussion and critical reflection.

> *For a script concordance test [a written test on clinical reasoning using video recordings of real patients], a panel of experts provide the answer key. The discussion with students about differences of opinion between students and experts is interesting. You need to discuss and explain the test items and results with the students, helping them to set their personal goals. (T-6)*

A student stated:

> *Hearing from a teacher what was correct and what was wrong in my observations of the patient helps me to understand whether I am on the right track or not. (S-5)*

Moreover, the process of providing peer feedback, receiving peer feedback, and discussing feedback encourages students to express their thoughts and to develop a critical attitude towards their own and their peers' performance.

> *Discussing the feedback gives me insights on my thoughts, my own capabilities, or shortcomings. (S-3)*

Contextual learning

Teachers reported the importance of contextual learning in two ways. First, skills learned in the classroom setting should be transferred to the context of their future professional clinical practice. Observing human movement in a learning context, which simulates clinical practice as much as possible, was advocated. Second, teachers should facilitate the process of transferring knowledge and skills. Skills learned on one problem should be transferred and applied in other situations. Third, teachers emphasized that learning a single skill should be embedded in the body of knowledge of the profession and related to the whole process of diagnosis and intervention.

> *In the classroom, you have to refer constantly to related theoretic backgrounds . . . . Physiotherapy is an applied science; it needs the incorporation of knowledge of a physiotherapeutic problem of the patient, a health problem. As a teacher, you have to make sure to constantly refer to and explain the connections. (T-11)*

Building complexity

Teachers mentioned that it is important to consider the degree of complexity of movement observations in the construction of education. The students' achieved level of observational skills led to the decision to start with global or specific observations, with simple or complex cases, or with simulated or real patients.

> *First, students notice large differences in performances, they evolve in their observation, [and] they start noticing smaller, more defined differences. Observing healthy subjects is more complicated than [it is with] patients with real health problems. (T-8)*

> *Basically, the complexity rises the moment the student gets less time to observe and interpret the motor skills performance. (T-6)*

Educational tools

Teachers and students mentioned educational tools as beneficial for learning processes. Examples were the use of video recordings, Internet sources, real patients or volunteers with minor health problems in classroom settings, and patient demonstrations in clinical settings.

> *Tools you always can use are the technical instruments in our motion lab, but also [a device] as simple as a Smartphone or tablet. You can record . . . replay and watch together with the students [and] use it for learning. (T-4)*

Theme B: Professional content: What needs to be mastered

Participants described observational skills as the ability to observe carefully and attentively—to notice, perceive, and register what has been observed as significant. To learn observational skills, students must be aware of three different ways to observe human movement, referring to their respectively different approaches and goals for observation. Furthermore, participants described it as necessary to master domain-specific knowledge, gaining body awareness and building a frame of reference and expertise to learn observational skills. Six interrelated design principles were defined for this theme.

Implicit and explicit observations

Differences between implicit and explicit observations were often mentioned, emphasizing the importance of the ability to explicitly describe findings. Both students and teachers described how students basically start to 'watch' as human beings (implicit), providing them with rich information, though it is not yet related to the physiotherapeutic domain. When they learn to name what they see, they learn to 'observe' (explicit).

> *I think one always has a certain frame of reference . . . . sometimes you have a gut feeling while observing; you think movements are deviant, but you can't explicitly name what you actually see. (MS-3)*

> *As a teacher, you can give structure for how and a language for what they observe; this enables them to explain what they observe. (T-5)*

Physiotherapeutic observations were related to knowing what to expect and recognizing what one sees based on domain-specific knowledge on human movement.

> *Whenever I describe what I see, this should result in recognition. I can only recognize something when I see what I know; otherwise, there is no recognition. Recognition means you did observe and not only watch. (T-1)*

The observation of simple versus complex motor behaviour
Participants reported differences between simple movements; a discrete local movement (e.g., the flexion of an elbow or the action of a knee while walking), and complex motor behaviour (e.g., the behaviour of any motor skill that is influenced by characteristics of the skill itself, the person performing the skill, and the environment in which the skill is performed). The observation of complex motor behaviour was emphasized as significant for physiotherapists.

> *A physiotherapist incorporates the context in the movement performed [and] searches for the intention of the movement; what is the goal to achieve, and is this possible in the circumstances as they are performed in? . . . They are not just observing different parts of a movement. (P-10)*

Objective versus subjective observations
All participants emphasized the importance of both objective observations free of prejudice, and objective interpretation of the observations. They mentioned starting observations subjectively as a pitfall, resulting in early interpretations from predetermined assumptions about the quality of movement. In every patient encounter, a practitioner must have an open mind.

> *By interpreting in an early stage, you have pre-assumptions, causing a focus on certain parts of the skill or body. With that mind-set, you start filtering your observation right away; however, you possibly fail to observe certain other important aspects because they don't fit in your predetermined picture. (T-9)*

A barrier to objective observation might be a limited frame of reference based on expertise in other contexts and knowledge domains.

> *The students who already have observed in a different context [e.g., sports] possibly think they know what observing movement means. But also, they have to observe in the specific context of health problems, physiotherapy, to start building a new frame of reference. Essential in this is the right state of mind, which enables this. (T-6)*

Mastering domain-specific knowledge

Participants stated that domain-specific knowledge regarding human movement is necessary for physiotherapeutic observations. They mentioned specific theoretical contents (e.g., anatomical, kinesiological, physiological, and biomechanical principles). However, they also cited specific theoretical knowledge of motor control, motor learning, and motor development through the life span.

> *Most important for me was to start with theoretical knowledge and backgrounds. I really think it all starts there; you have to know what to look for. Knowledge about what is normal, what is deviant, . . . what is normal for certain ages. Also, theoretical backgrounds such as, e.g., Fitts and Posner are essential while observing; what stages in motor performance are shown? (MS-1)*

Furthermore, the participants differentiated between qualitative and quantitative aspect of movement performances. The need for explicit descriptions on how to appraise either the quality or quantity of movement was clear. Participants gave extended explanations about how to judge the quality and quantity of movements.

> *Whenever you are going to observe and judge quality of movement, you have to know what it is. Ideally, there should be some kind of list with items and descriptions for qualitative movement. (S-4)*

Movement assessment

The availability and use of standardized assessment tools, motor tests, norm and reference values, and checklists for the observation of movement quality were discussed on several occasions. The necessity of using them while working with patients was evident.

> *Interpretations of observations occur based on reference values; a student has to know these. He has to recognize variances of performance regarding*

the norm. This enables the student to recognize deviant movement or when a performance is more or less, according to the norm, a typical movement. (T-1)

It is important that a student has a structure for observations. They have to learn reference values for typical movement. You could use measurement tools in the classroom for this goal. (T-5)

Body and movement awareness

Body awareness and an adequate sense of movement with the observed human movement were perceived as essential to recognizing and understanding problems in functioning.

Before observing, it is beneficial to perform certain activities yourself [student] and reflect on how you perform just to get a picture of the movement. (T-10)

Experience

Teachers explained that mastering observational skills and developing expertise requires deliberate practice and experience in a variety of clinical situations. They stated that experiences are critical for building a personal frame of reference and for mastering observational skills and interpreting the quality of human movement.

Knowing and recognizing is very important; the more you know, the more you recognize, and that is called expertise. (T-1)

Theme C: Conditions for optimal learning

The observation of human movement requires optimal conditions, not only for learning as a student but also for coaching students as a teacher or observing human movement as a physiotherapist. These conditions were described as follows: (1) accurate professional behaviour, including creating mental space and tranquillity; and (2) establishing a standard language. Two design principles were defined.

Professional behaviour

A professional attitude was indicated as necessary to observe human movement. Willingness and ability to show certain behaviour, determined by norms and values, motivation, and personal incentives, were mentioned as essential. Examples were given about learning readiness, perseverance, decisiveness, and willingness to collaborate.

As a physiotherapist . . . you have to take time for your patient, and as a teacher, most importantly, you have to create the opportunity and structure for the student to develop this professional behaviour. (T-8)

Standard language

Participants argued for the use of standard language when observing and interpreting human movement. They mentioned the current diversity in definitions and descriptions of movement quality. Establishing a standard language guided by therapeutic constructs was mentioned as a way to improve communication between physiotherapists.

> *There is a need for clear definitions and descriptions. If definitions and descriptions are clear and mutually shared, then there will be less confusion in communication between health professionals. (P-7)*

Proto-theory for educational program design

The interrelations between our three identified themes 1) didactics; 2) professional content; and 3) conditions for optimal learning, and their design principles are shown in Figure 2. By discussing and reflecting on our analysis and results and by relating our results to existing literature, we developed a proto-theory for an educational program designed for learning observational skills (Table 2). Although the design principles for the didactic, professional content, and conditions for optimal learning themes are interrelated, our proto-theory focuses on the more generic didactic principles that facilitate observational skills learning. We considered those principles as generalizable to other professional content also, whereas conditions for optimal learning are prerequisites for learning. Our proto-theory shows the link between the identified design principles for the didactic theme, associated theories on learning and behavioural change, and recommendations for observational skills learning. Based on the design principles for the didactic themes, we identified six underlying educational theories for our proto-theory. These six theories, with a short description, are as follows:

(1) *Self-regulated learning theory:*[25–28] Students are self-regulated to the degree that they are metacognitively, motivationally, and behaviourally active participants in their learning. Learning goals are set by the student based on both past experiences and current learning experiences. The four areas of self-regulation are cognition, motivation, behaviour, and context.

(2) *Self-determination theory:*[29] Motivation for learning initiates activities and enhances the self to initiate behaviour. Three basic needs for intrinsic motivation are (1) competence or efficacy; (2) relatedness (i.e., the need to feel a sense of belonging and connectedness with others [e.g., peers/ teacher]); and (3) autonomy (i.e., the ability to perform activities on one's own initiative, without experiencing external control or influences).

3) *Social constructivist learning theory:*[30] Knowledge construction enhances attention to, storage of, and retrieval of knowledge from memory using contextual cues to facilitate the transfer of learning from the learning

context to the application context. Teaching occurs in the zone of proximal development (as described by Vygotskiĭ[31]) using scaffolding.[32]

(4) *Social cognitive learning theory:*[33,34] Active participation and knowledge construction occur in collaboration with peers, enhancing the development of self-efficacy beliefs by performing the new behaviour and experiencing the consequences of that behaviour (mastery experience).

(5) *Variation theory:*[35,36] Individuals understand phenomena in the world differently because experience is always partial. Learning takes place when difference occurs against a background of sameness. Conditional for learning are the four patterns of variation: contrast, generalization, separation, and fusion.

(6) *Feedback intervention theory:*[37,38] Feedback changes the locus of attention among three general and hierarchically organized levels of control: task learning, task motivation, and meta-tasks processes (including metacognitive aspects of task learning).

Table 2. Proto-theory for educational program design to learn observational skills.

Design principle	Recommendations for didactic approach	Educational theory
Student in the lead	Building on a sense of urgency for learning knowledge and skills. Designing an improvement plan with personal learning goals, based on feedback.	Self-regulated learning theory [25–28]
	Students take responsibility for their own learning process.	Self-determination theory [29]
Teacher as facilitator	Stimulating critical reflection on performances; reflection-in-action and reflection-on-action.	Self-regulated learning theory [25–28]
	Fading guidance and control. Stimulating deep questioning among each other. Behaving as a coach, not as a lecturer (e.g. empowering students to take responsibility).	Self-determination theory [29]
	Designing a meaningful learning environment with meaningful learning tasks that apply to professional clinical practice. Tailoring instruction or coaching to the actual level of performance. Simulating students to transfer learning experiences to the context of clinical practice.	Social constructivist learning theory [30]
		Social cognitive learning theory [33,34]
	Designing a learning environment that enables students to identify similarities and differences in various patterns of human movement.	Variation theory [35,36]
Peer learning	Working together, explaining ideas to peers, questioning each other's views.	Social constructivist learning theory [30]
Assessment and feedback for learning	Developing a critical attitude towards own and others' performances. Providing performance feedback; emphasizing strengths, challenges, and next steps. Discussing the improvement plan with others.	Self-regulated learning theory [25–28]
		Social cognitive learning theory [33,34]
		Feedback intervention theory [37,38]
Contextual learning	Performing the observational skill individually, by reasoning aloud, and applying in professional context.	Social constructivist learning theory [30]
Building complexity	Helping students to make the transfer of observed differences and similarities to a new context and with higher complexity. Presenting a variety of problems that reflect clinical practice.	Variation theory [35,36]
Educational tools	Enabling the observations of human movement in (real life, or video recorded) patients to support the learning process.	Social constructivist learning theory [30]

Discussion

The overall aim of our study was to derive design principles for an educational program to develop observational skills for students in physiotherapy. After our exploration of the needs, beliefs, and experiences of students, teachers, and practitioners regarding observational skills learning, we used the analysed data to develop design principles and a proto-theory (Table 2). We based the design principles for the three identified themes; (1) didactics; (2) professional content; and (3) conditions for optimal learning, on content-related data. However, the developed proto-theory focuses on more generic didactic principles that facilitate learning and is itself independent of the content.

The decision to focus the proto-theory specifically on the design principles belonging to the didactic theme was partly based on the poor description in the literature of specific educational strategies to support the development of observational skills learning.8 Another consideration was that design principles for the didactic theme and its associated principles could possibly be identified as generic didactic design principles. In contrast to professional content, which depends on the topics, skills, or goals students have to learn or achieve,[39–42] and conditions for optimal learning which are prerequisites for learning and professional behavior.[43] Including the design principles from all three identified themes in a proto-theory would have been possible; however, this would have led to the development of a more specific proto-theory for observational skills learning. While analysing our data, we realized we had the opportunity to develop a generic proto-theory for learning by only including the design principles belonging to the didactic theme. The developed proto-theory shows the link between the design principles for the didactic theme and the associated theoretical concepts on learning and behavioural change, which could be considered generalizable and thus could be combined with other learning content in diverse professions. This proto-theory could also guide the ongoing development and refinement of educational design for skills learning.

One of the results of our study is the proto-theory that comprises two leading design principles: 'student in the lead' and 'teacher as facilitator'. The other principles (i.e., peer learning, assessment and feedback for learning, contextual learning, building complexity, and educational tools) can be linked with these two principles. The first design principle is that students need to take the lead in their own learning to master observational skills. The importance of self-regulating learning processes based on personal learning goals is explained by the self-regulated learning theory;[25–28] and self-determination theory.[29] Supporting and allowing students to take the lead in their own learning and the responsibility for their own learning, depends on both active involvement in learning and the students' motivation to learn.[26–29] Furthermore, the level of motivation is significantly related to the students' self-efficacy beliefs.[33] For the students, this implies deliberate practice, critical reflection on the quality of the demonstrated observational skills, asking for feedback, and designing an improvement plan with personal learning goals.

The second leading principle is the teacher as facilitator. To facilitate student learning, the teacher helps to identify the personal learning needs of students, monitors their learning process, checks understanding, and adjusts instruction or coaching tailored to the actual level of students' performance, as supported by the social constructivist learning theory.[30,31] Furthermore, the teacher enhances collaborative learning and critical performance appraisal among peers.[30] A dialogue with peers is important to build adequate self-perceptions as a tool to organize and give meaning and understanding to thoughts. Interaction with peers can help students to develop a critical attitude towards themselves and towards peers and to create opportunities to help each other in organizing thoughts by stimulating reasoning aloud.[44,45] In this collaborative learning process, the teacher coaches the students to develop adequate skills to support each other by providing constructive feedback. Active engagement in group work, and critical self-appraisal, will help students to optimize the distance between "what is known" and "what is to be learned".[44,46] Additionally, the teacher, as facilitator, is responsible for the choice in complexity of cases and the choice of educational tools to facilitate gradual student development according to learning in Vygotskĭi's zone of proximal development.[31]

Another design principle is the choice of educational tools. More specifically, the choice between enabling students to observe real patients, simulated or standardised patients, or students acting as standardized patients must be considered. Observing real patients adds the most to the learning process of students; however, the possibilities to observe real patients are often limited. Therefore, as an alternative, peer students are often used as standardized patients in the classroom setting. However, the use of same-year students acting as standardized patients may be perceived as less realistic than the use of trained standardized patients.[47] Training senior students as standardized patients for junior students could overcome these concerns, as the study of Mandrusiak et al.[48] demonstrated. Senior students acting as standardized patients resulted in positive experiences for both junior and senior physiotherapy students, with significant improvements in reported self-efficacy and satisfaction.[48]

Although our developed proto-theory and the figure with interrelated design principles could be useful for teachers, it does not necessarily mean that all teachers can apply them in their daily practice of teaching. This implies a need to focus on training for teachers as well. Teachers must monitor and adjust the learning process and identify the learning needs of students. Doing so requires a sense of efficacy for teachers; they must have or develop positive beliefs about their ability to effectively teach, monitor, and assess all students well.49 How to reflect critically on students' performance and provide tailored feedback, possibly based on an assessment, must furthermore be addressed in teacher training. Not all feedback leads to performance improvement, as indicated by the feedback intervention theory.[37,38] The type of feedback and the circumstances in which feedback is provided can affect students' motivation and self-efficacy beliefs.[50,51] Creating opportunities for teachers to observe other teachers performing successfully in class (serving as role models), as well

as experiencing their own success in classroom contexts, might enhance teachers' sense of efficacy and help them to facilitate student learning.[34,49]

Strengths and limitations

A strength of our study was the methodology of design-based research, which enabled us to triangulate multiple perspectives and sources of data collection in developing an educational program to teach observational skills.[12–15] Using design-based research gave us the opportunity not only to invite students, lecturers, and practitioners but also for several individual rounds of data gathering using NGT meetings and interviews. We believe that expertise in learning strategies should be derived not only from professionals but from everyone who is involved in learning.[13,14] Moreover, a research group with a variety of backgrounds and expertise, qualitative as well as design research, education, and physiotherapy, strengthens the review process and discussion of the themes and design principles. We ensured the inclusion of teacher participants with various theoretical and practical backgrounds (e.g., musculoskeletal physiotherapy, sports physiotherapy, neurologic physiotherapy), to reduce the influence that the researchers' experiences and backgrounds (i.e., paediatric physiotherapy) could have had on the outcome.

A challenge of design-based research is the difficulty of deciding when, if ever, a study is completed; because every cycle provides new information, it will be difficult to reach saturation.[13,17] For our study, we modified the four sequential steps developed by Reeves19 according to our research aim: to derive design principles for the development of educational programs to teach observational skills for physiotherapy. A possibility was to conduct at least one more round of testing an education program, which would have improved the research design and its efficiency. However, we chose to conduct only one iterative cycle in step 3 and to continue the iterative process in step 4. By conducting NGT meetings and interviews and evaluating the draft educational program, we had the opportunity to go back to the participants in different rounds and listen to their opinions. After the last round, no new data emerged, so we decided to focus on the theoretical understanding of the derived design principles through reflection in step 4.

Although we did not identify any differences in design principles and approaches for observational skill learning between bachelor and master students in either our study or the literature,[8] we did not explore this explicitly; it would be interesting to study this in more depth in the future. A limitation was the decision in step 3 of our study to test the draft educational program during a postgraduate program for which learning outcomes were set. In the NGT meeting following this training, the participants possibly not only responded to the question about their perceived needs, beliefs, and experiences regarding observational skills learning, but also to the draft version of the manual and supplemental materials in relation to the learning outcomes, which may have contributed to the data and analysis. Another limitation is that although it is unlikely that observation of movement quality differs between countries, language differences and differences in didactic approaches

may influence the international transferability of the results. A recommendation for further research is to continue the iterative process of the development of educational materials with students, teachers, and practitioners to complete the implementation of materials in practice.

Conclusion

We conclude that, to learn observational skills, students must take the lead in their own learning process, facilitated by an experienced teacher. This might imply a need for additional teacher training to strengthen their didactic skills. Another precondition for learning observational skills is a realistic context. This might imply the consideration to enable students to observe real or standardised patients to increase possibilities for observation in a clinical context. Although our proto-theory was developed for physiotherapy students to learn observational skills, it might be interesting to study whether this proto-theory is applicable to other professions and in programs with different content related to learning observational skills, for example, interviewing patients. Regarding the research methodology used for designing educational programs, we can conclude that a design-based approach suited the project well. Its advantage is that all stakeholders fully participated in the design process, enhancing the applicability of the design in the context of its end users.

Acknowledgements
We thank Irene van der Glind for her support during the analyses process.

Funding
This work was supported by the Dutch Research Council (NWO) [023.004.037].

Declaration of Interest Statement
The authors report no conflict of interests.

References

1. Skjaerven LH, Kristoffersen K, Gard G. An eye for movement quality: a phenomenological study of movement quality reflecting a group of physiotherapists' understanding of the phenomenon. *Physiother Theory Pract*. 2008;24(1):13-27.
2. Schmidt RA, Wrisberg CA. Getting started. In: *Motor Learning and Performance, a Situation-Based Learning Approach*. 4th editio. Leeds: Human Kinetics; 2008:3-23.
3. Magill R, Anderson D. Motor Control Theories. In: *Motor Learning and Control. Concepts and Applications*. 10th editi. Singapore: McGraw Hill Education; 2014.
4. Knudson D V, Morrison CS. Interdisciplinary nature of qualitative analysis. In: *Qualitative Analysis of Human Movement*. 2nd editio. Leeds: Human Kinetics; 2002.
5. Janssen AJWM, Diekema ETW, Dolder van R, Kollée LAA, Oostendorp RAB, Nijhuis-van derSanden MWG. Development of a movement quality measurement tool for children. *Phys Ther*. 2012;92:574-594.
6. Boudreau JD, Cassell EJ, Fuks A. Preparing medical students to become skilled at clinical observation. *Med Teach*. 2008;30(9-10):857-862.
7. Berger L. A hierachy of clinical observations. *Pediatrics*. 1980;65(2):357-358.
8. Al-Moteri MO, Plummer V, Cooper S, Symmons M. Training paradigms to enhance clinical observational skills in clinical practice : A scoping review. *J Nurs Educ Pract*. 2015;5(12):96-103.
9. Jarodzka H, Scheiter K, Gerjets P, van Gog T. In the eyes of the beholder: How experts and novices interpret dynamic stimuli. *Learn Instr*. 2010;20(2):146-154.
10. Shapiro J, Rucker L, Beck J. Training the clinical eye and mind: using the arts to develop medical students' observational and pattern recognition skills. *Med Educ*. 2006;40(3):263-268.
11. Jarodzka H, Balslev T, Holmqvist K, et al. Conveying clinical reasoning based on visual observation via eye-movement modelling examples. *Instr Sci*. 2012;40(5):813-827.
12. The Design-Based Research Collective. Design-based research: an emerging paradigm for educational inquiry. *Educ Res*. 2003;32(1):5-8.
13. Anderson T, Shattuck J. Design-based research: A decade of progress in education research? *Educ Res*. 2012;41(1):16-25.
14. Barab S, Squire K. Design-based research: putting a stake in the ground. *J Learn Sci*. 2009;13(1):1-14.
15. Brown AL. Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *J Learn Sci*. 1992;2(2):141-178.
16. McKenney S, Reeves TC. Systematic review of design-based research progress: Is a little knowledge a dangerous thing? *Educ Res*. 2013;42(2):97-100.
17. Vanderhoven E, Schellens T, Vanderlinde R, Valcke M. Developing educational materials about risks on social network sites: a design based research approach. *Educ Technol Res Dev*. 2016;64(3):459-480.
18. Wang F, Hannafin MJ. Technology-enhanced learning environments. *Educ Technol Res Dev*. 2005;53(4):5-23.

19. Reeves TC. Design research from a technology perspective. In: DenAkker J V, Gravemeijer K, McKenny S, Nieveen N, eds. *Educational Design Research*. London: Routledge; 2006:52-66.
20. Kennedy-Clark S. Research by Design: Design-Based Research and the Higher Degree Research student. *J Learn Des*. 2013;6(2):26-32.
21. Herrington J, McKenney S, Reeves T, Oliver R. Design-based research and doctoral students: Guidelines for preparing a dissertation proposal. 2007;2007(2007):4089-4097.
22. Potter M, Gordon S, Hamer P. The Nominal Group Technique: A useful consensus methodolgy in physiothearpy research. *New Zeal J Physiother*. 2004;32(2):126-130.
23. van de Ven A, Delbecq A. The Nominal Group as a Research Instrument for Exploratory Health Studies. *Am J Public Health*. 1972;March:337-342.
24. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol*. 2006;3(2):77-101.
25. Zimmerman BJ. Self-regulated learning and academic achievement: an overview. *Educ Psychol*. 1990;25(1):3-17.
26. Zimmerman BJ. Becoming a self-regulated learner: an overview. *Theory Pract*. 2002;41(2):64-70.
27. Schön D. The reflective Practitioner. In: *How Professionals Think in Action*. San Fransisco: Jossey-Bass Inc.; 1983.
28. Greene JA, Azevedo R. A Theoretical Review of Winne and Hadwin's Model of Self-Regulated Learning: New Perspectives and Directions. *Rev Educ Res*. 2007;77(3):334-372.
29. Ryan R, Deci E. Self-determination theory and the facilitation of intrinsic motivation. *Am Psychol*. 2000;55(1):68-78.
30. Norman G, Bordage G, Page G, Keane D. How specific is case specificity? *Med Educ*. 2006;40(7):618-623.
31. Omrod J. *Human Learning*. 4th ed. Upper Saddle River: Pearson Prentice Hall; 2004.
32. Rasmussen J. The importance of communication in teaching: A systems-theory approach to the scaffolding metaphor. *J Curric Stud*. 2001;33(5):569-582.
33. Bandura A, Locke EA. Negative self-efficacy and goal effects revisited. *J Appl Psychol*. 2003;88(1):87-99.
34. Bandura A. The Explanatory and Predictive Scope of Self-Efficacy Theory. *J Soc Clin Psychol*. 1986;4(3):359-373.
35. Marton F, Säljö R. On qualitative differences in learning: Outcome and process. *Br J Educ Psychol*. 1976;46(1):4-11.
36. Marton F, Trigwell K. Variatio Est Mater Studiorum. *High Educ Res Dev*. 2000;19(3):381-395.
37. Kluger AN, DeNisi A. Effects of feedback intervention on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychol Bull*. 1996;119(2):254-284.
38. Kluger AN, van Dijk D. Feedback, the various tasks of the doctor, and the feedforward alternative. *Med Educ*. 2010;44(12):1166-1174.
39. van de Wiel M, Boshuizen H, Schmidt H. Knowledge restructuring in expertise development: Evidence from pathophysiological representations of clinical cases by students and physicians. *Eur J Cogn Psychol*. 2000;12(3):323-356.

40. Lubarsky S, Dory V, Audétat M-C, Custers E, Charlin B. Using script theory to cultivate illness script formation and clinical reasoning in health professions education. *Can Med Educ J*. 2015;6(2):e61-70.
41. Norman G, Eva K. Diagnostic error and clinical reasoning. *Med Educ*. 2010;44(1):94-100.
42. Monteiro SM, Norman G. Diagnostic Reasoning: Where We've Been, Where We're Going. *Teach Learn Med*. 2013;25(SUPPL.1):37-41.
43. Grace S, Trede F. Developing professionalism in physiotherapy and dietetics students in professional entry courses. *Stud High Educ*. 2013;38(6):793-806.
44. Chi MTH, de Leeuw N, Chiu M, LaVancher C. Eliciting Self-Explanation Improves Understanding. *Cogn Sci*. 1994;18:439-477.
45. Sluijsmans DMA, Brand-Gruwel S, van Merriënboer JJG, Bastiaens TJ. The Training of Peer assessment Skills to Promote the Development of Reflections Skills in Teacher Education. *Stud Educ Eval*. 2003;29(03):23-42.
46. Seenan C, Shanmugam S, Stewart J. Group peer teaching: a strategy for building confidence in communication and team working skills in physiotherapy students. *Physiotherapy*. 2016;102(3):S282.
47. Mounsey AL, Bovbjerg V, White L, Gazewood J. Do students develop better motivational interviewing skills through role-play with standardised patients or with student colleagues? *Med Educ*. 2006;40(8):775-780.
48. Mandrusiak AM, Isles R, Chang AT, et al. Senior physiotherapy students as standardised patients for junior students enhances self-efficacy and satisfaction in both junior and senior students. *BMC Med Educ*. 2014;14(1):1-7.
49. Delale-O'Connor LA, Alvarez AJ, Murray IE, Milner HR. Self-Efficacy Beliefs, Classroom Management, and the Cradle-to-Prison Pipeline. *Theory Pract*. 2017;56(3):178-186.
50. Gielen S, Peeters E, Dochy F, Onghena P, Struyven K. Improving the effectiveness of peer feedback for learning. *Learn Instr*. 2010;20(4):304-315.
51. Govaerts M. Workplace-based assessment and assessment for learning: threats to validity. *J Grad Med Educ*. 2015;7(2):265-267.

Chapter 7

General Discussion

The aim of this thesis was to gain insights into the measurement properties of the Observable Movement Quality (OMQ) scale through the determination of its reliability, validity and responsiveness. The second aim was to investigate what students in physiotherapy need in their educational program to develop observational skills and which didactic principles facilitate this learning. The findings are discussed in this chapter, followed by reflection on and discussion of the measurement properties of the OMQ scale, the challenges to the observation of movement quality and the integration of observational skills in a competency-based physiotherapy profile. Subsequently, the methodological considerations and recommendations for future research are addressed, followed by recommendations for implementation in clinical practice and education.

Measurement properties of the OMQ scale

The Consensus-based Standards for the Development of Measurement properties (COSMIN)[1,2] guidelines for measurement properties were followed in our studies; reliability, validity and responsiveness of the OMQ scale were established. In an exploratory validation study for construct validity, described in **Chapter 2**, 6 out of 7 hypotheses were confirmed for the OMQ scale, supporting the content validity of the scale to measure movement quality. **Chapter 3** describes a prospective longitudinal cohort study for individual neurodevelopmental trajectories over 5 years. A moderate to good correlation was shown on outcomes for movement quality assessments, at 3 months of age, using General Movements (GMs) assessment and the OMQ scale; correlations between motor developmental outcomes at 5 years of age and GMs were high and significant, as were correlations between motor developmental outcomes and OMQ scale scores. These results confirmed the construct validity of the OMQ scale, and, furthermore, they indicated a tendency towards good predictive and concurrent validity. Moreover, in a cross-sectional study – described in **Chapter 4** – interrater reliability for the OMQ scale was shown to be moderate, with a good item agreement among participants, when using the scale in the age category ranging from 6 months to 6 years of age. And last, in **Chapter 5**, a prospective intervention study showed a moderate to good interrater reliability of the OMQ scale, with good item agreement, while the intrarater reliability was shown to be high. Furthermore, a responsiveness to change was seen when used to assess movement quality after physiotherapy intervention in children aged 4 to 12 years.

Altogether, the findings of our studies on measurement properties indicate that the OMQ scale is valid for measuring movement quality in clinical practice in addition to motor performance tests. However, measurement properties for cross-cultural validity – the degree to which the performance of the items on a translated or culturally adapted instrument are an adequate reflection of the performance of the items of the original version[1] – have yet to be established for the OMQ scale; we investigated the OMQ scale in Dutch clinical practice alone. Moreover, the OMQ scale was not yet tested in the older age group of children aged 12 to 16 years.

Observation of movement quality

The observation process involves gathering, organising and giving meaning to visual, auditory and sensory information obtained while observing a moving person.[3,4] Observation has sensory, perceptive and cognitive components; is distinct from inference or judgment; and is made concrete through description.[5] Furthermore, observations are goal-oriented, occur over time, carry ethical obligations and occur on different levels: the whole person observed, a body part, the personal or environmental context and behaviours and interactions.[5,6] The aim for the development of the OMQ scale was to realise a measurement instrument to observe movement quality independently from a specific age, motor task and predetermined theoretical construct to allow comparable observations for physiotherapists and longitudinal evaluations.[7] For observations of movement quality using the OMQ scale, our cross-sectional study with 28 participating paediatric physiotherapists unfamiliar with the scale – detailed in **Chapter 4** – showed moderate interrater reliability, with no differences based on either work experience or work setting. Comparable outcomes for interrater reliability of the OMQ scale were found in our prospective intervention study – outlined in **Chapter 5** – in which two master's students in paediatric physiotherapy and one paediatric physiotherapist were involved. Those results did not support our hypothesis for the expected differences in introspective judgment of movement quality on the basis of clinical expertise. However, the outcomes and OMQ total scores of the master's students and the experienced physiotherapist were compared using Bland-Altman plots, showing a systematic difference in OMQ total scores and an increase in the plots for the experienced therapist, indicating higher total scores for this examiner. The systematic difference in scores, and the higher OMQ total scores for the experienced therapist, could have been a result of the difference in the amount of clinical experience. Experience with the observation and assessment of typical and atypical development, as well as specific diagnoses or diseases, enables a therapist to recognise deviations in movement quality more adequately. Although results for both studies do support the assertion that the OMQ scale focuses on observable aspects of movement and can be scored by physiotherapists independent from work experiences or work setting, results for reliability could indicate that clinical experience is important for the interpretation of outcomes of movement observations.

Observation is a fundamental skill for physiotherapists,[3] informing the process of clinical reasoning fundamental to the design of interventions aimed at improving functional abilities. For observational skill learning, as described in **Chapter 6**, students must take the lead in their own learning process, facilitated by an experienced teacher. This teacher needs to help the students prepare for clinical care, for which the integration of physiotherapeutic knowledge and skills needs to be addressed.[8] For the development of the observational skills in a classroom setting, the students need to understand that scoring demands an introspective judgment of movement quality based on systematic observations and internal reflection, incorporating the (future) therapist's knowledge and specific experiences with the

target group.[3,4] An interpretation of the outcomes of movement observations is necessary to design interventions for patients to improve functional abilities; for this, students need to develop clinical reasoning skills.[8] A description for clinical reasoning in physiotherapy was conceptualised by Huhn et. al.[9] as integrating cognitive, psychomotor and affective skills; it is contextual in nature and involves both therapist and client perspectives. It is adaptive, iterative and collaborative; the intended outcome is a biopsychosocial approach to patient care.9 Through practical experience, observational and clinical reasoning skills gradually increase, influenced by a variety of patients.[8,10–12] Repetitive clinical exposure to a variety of patients enables students to compare and reflect on differences and similarities in approaches observed in students and professional therapists, enhancing the development of clinical reasoning skills.[8] This might imply not only a need for additional training for teachers to strengthen their didactic skills to facilitate observable skill learning but also a need for additional training to strengthen didactic skills to facilitate the development of clinical reasoning strategies. Furthermore, it might be necessary to evaluate curricula for students in physiotherapy and to increase possibilities for the students to learn to observe in a realistic context.

The results from our reliability studies show that the OMQ scale can be validly scored by physiotherapists independent from work experiences or work setting. However, results for reliability could also indicate that clinical experience increases observational and reasoning skills,[5,7–9] supporting our recommendation to involve experienced teachers in helping students prepare for clinical care.

Observational skills in a competency-based physiotherapy profile

Competencies for professionals involve the ability – based on knowledge, skills, attitude and role conceptions – to act adequately in a complex professional situation and to be able to account for and reflect on choices and decisions made. In the Netherlands, a competency framework for the physiotherapy profession using the CanMEDS-model[13] was designed, resulting in the description of seven metacompetencies.[14] For each of the seven metacompetencies, four key competencies were described.[14] To be able to perform professional practice following the competency profile, a body of knowledge, body of skills and body of attitudes is required, as explained in *National Transcript Physiotherapy*.[12] The curricula for physiotherapy in the Netherlands are competency-based, and these competencies guide outcome expectations for students at the completion of their training together with the explanations in National Transcript Physiotherapy. However, detailed descriptions of what a specific body of knowledge, body of skill or body of attitudes entails appear in neither the competency framework nor the National Transcript Physiotherapy. Therefore, the competency framework and the National Transcript Physiotherapy provide generalized descriptions to guide learners, their teachers, supervisors at practical placements and educational institutions in teaching and assessment.[16] The expected outcomes across programs vary because there are no uniform or consistent guidelines for setting those

expected outcomes, nor have widely accepted outcomes been adopted in the professional education community.[17] Accordingly, a challenge to competency-based physiotherapy curricula is finding ways to structure and assess the essential integration of knowledge, skills and attitudes that shape an entry-level physiotherapist who is self aware and can be trusted to practice without direct supervision.[18,19]

Facilitating the translation from a competency-based profile to education

The entrustable professional activity (EPA) concept was conceived to facilitate the translation from the competency framework to the world of health care, addressing the concern that competency frameworks would otherwise be too theoretical to be useful for training and assessment in education and daily practice.[16,18] An EPA represents a unit of essential professional activity that an individual can be trusted to perform in a health-care context, once sufficient competence has been demonstrated.[18] Generally, an EPA requires the integration of competencies from two or more competency domains – including the key dimensions critical to the profession, such as knowledge, communication and clinical skills (see Table 1).[19] The behavioural descriptors that indicate the level of performance for a given competency are represented by milestones and identify knowledge, skills and attitudes, organised in a developmental framework from less to more advanced, applying critical thinking to patient care.[19,20] In short, EPAs are tasks that must be accomplished, whereas milestones are stages in the development of specific competencies, providing insights into the abilities of the individual.[16,19] During training, the teacher must evaluate the student's abilities, which can be accomplished through a small number of assessments that enclose multiple milestones integrated into an EPA. The teacher can observe students' progress on an EPA multiple times, initially for coaching purposes and later to record performance. Furthermore, EPAs and milestones will provide both teacher and student with clear expectations of the skills and abilities they need at each stage of training, which will help in planning learning, teaching and coaching opportunities.

Table 1. Concepts for designing an assessment framework for physiotherapy curricula[a].

	Definition	Example for physiotherapy assessment and interpretation
Milestones	Behavioural descriptor indicating a level of performance for a competency in the development of a student in physiotherapy	Milestones describes the knowledge, skills and attitude needed. Are organised from less to more advanced to enhance development. • Beginner: recognises needed knowledge, and starts to interpret findings from observations • Advanced: summarises information obtained from multiple sourced, and develops an assessment plan • Competent: Takes expected and unexpected outcomes into account in the development of a physiotherapeutic assessment and intervention plan
Competencies	Ability to integrate knowledge, skills, and attitude. Can be measured and assessed.	The student collects essential and correct information about the patient using anamnesis and physiotherapy assessment Demonstrates the ability to adapt to the physical, cognitive, and cultural needs of the patient
Domains of competence	Key dimension or area critical to the profession of physiotherapy.	Clinical skills Observational skills • **Observation of simple versus complex motor behaviour*** • **Implicit versus explicit observation of movements*** • **Objective versus subjective observation of movements*** • **Movement assessment*** **Professional behaviour*** Knowledge for practice Communication

Entrustable professional activity
(EPA)
	Represents a unit of essential physiotherapy activity
	Represents the core elements of physiotherapy.
	Units of physiotherapy practice that all graduates can perform unsupervised.
	Require generally an integration of competencies from two or more domains of competences.
		Conducts a physiotherapeutic assessment of a patient to investigate the cause or nature of the patients' condition or problem.
		• Description of performance of a complete physiotherapy assessment relevant to the patients' needs, including:
		  • Collecting quantitative assessment data
		  • Collection data from observation of movements
		  • Collecting data from observation of functional abilities

[a] Adapted from Chesbro et al. (2017);
*identified design principles for an educational program to develop observational skills (Chapter 6)

The competency profile for physiotherapy describes four competencies within the meta-competency 'physiotherapy activities': screening, diagnosis, intervention and evaluation.[14] Observation of movement quality is integrated into all four of those competencies, given that these observations inform the process of clinical reasoning, fundamental to diagnosing, designing and evaluating interventions. Thereby, the observational skill could be regarded as a domain of competence – or a key dimension – of the profession that physiotherapy students have to master. This key dimension of observational skills is represented in Table 1 and described based on the in Chapter 6 identified design principles from the theme 'professional content'. Students can master this skill by taking the lead in their own learning process, facilitated by an experienced teacher. However, student and teacher would benefit from described EPAs, domains of competency and milestones, leading to clearer expectations for learning skills and mastering abilities. The development and implementation of an EPA-based undergraduate clinical curriculum in the Netherlands was described by Ten Cate et al.;[21] for physiotherapy education no descriptions of curricula with EPAs are yet available. Chesbro et al.[19] explained in their article that constructing EPAs for physiotherapy education, which assesses essential expectations for physiotherapist development across the learning spectrum, would be a good start. Therefore, the construction and implementation of EPAs should be investigated in education research, keeping a focus on improving the quality of care and ensuring we are meeting the needs of society.[19] This implies that our proto-theory can be informative for designing an educational program to develop observational skills. However, to assess the development of such skills in students, they should be considered as a domain of competence, as part of an EPA – the representation of an integrated unit of essential professional activity – and not as a single skill. The concepts for a framework to develop EPAs appear in Table 1 and include an example of the development and integration of observational skills that could be used in the implementation of teaching and assessment of observational skills in curricula for physiotherapy.

Methodological considerations

Within this thesis, we used data from a variety of sources and stakeholders (literature, patients, physiotherapists, students and teachers). Furthermore, we combined both quantitative and qualitative designs; the studies to validate the OMQ scale used quantitative approaches, while the investigation of which didactical principles facilitated observable skills learning used a qualitative design. We chose these designs carefully to answer the research questions, to collect necessary data and to generate knowledge matching the current state of research within the field of paediatric physiotherapy and education. In general, this broad approach was a strength of our studies; data from different sources and different levels of specificity and evidence were compiled and compared.

Involvement of stakeholders

The stakeholders for learning the observational skills are, in our opinion, students, teachers and practitioners. However, patients do benefit from good observational skills because the observation of movements informs the process of clinical reasoning and is fundamental to the design of interventions that aim to improve functional abilities. Only the patient can tell whether an intervention was successful. Based on those considerations, it could be interesting to involve patients in possible future studies as stakeholders as well. On the other hand, involving patients in learning observational skills might interfere with their natural tendency toward perception and interpretation of what they see. Initial observation should be without any judgment.[22,23] However, the observer is an influencing factor when teaching and evaluating observational skills.[24,25] Involving patients in designing training programs could thus focus the observer more on perception than on judgment-free observations. Although it would be interesting to investigate and discuss this, we have decided not to do so in our study because we wanted to focus on student learning and which didactic principles facilitated this learning.

During the investigation of the adequateness of the OMQ scale, only physiotherapists were considered stakeholders. However, we decided to include master students, novices and expert physiotherapists in the studies. We based this decision on the consideration that the observation of movement is based on the introspective judgment of movement quality, influenced by knowledge, reasoning and personal experiences with the target group.[3,4] To estimate interrater and intrarater reliability of the OMQ scale, only physiotherapists could have been invited to participate in our studies. However, it would have been interesting to also include bachelor's students in physiotherapy. Comparison of outcomes for bachelor's students in physiotherapy and experienced therapists could have told us whether or not clinical experience increases observational and reasoning skills; our choice to not include these students is a limitation for our studies.

For the estimation of responsiveness of the OMQ scale, (parents of) the children could also have been considered as valuable stakeholders. Only they could tell whether or not the physiotherapist had designed an intervention to improve the abilities and whether or not the intervention was successful. However, when questioned about the effects of an intervention, patients have difficulty taking their baseline status into account because such ratings are strongly influenced by patients' current health status.[26–30] In the study on responsiveness of the OMQ scales, we decided not to ask the children to rate their perceived change in movement quality because they were too young to answer this question. Moreover, we anticipated that parents would have difficulties estimating changes in movement quality because descriptions were based on professional concepts. However, above all, we expected the parents to be influenced by the current health status of their child or even to want to please the physiotherapist by saying their child had improved,[30] and we decided to ask the treating physiotherapists to rate the change in movement quality, which we believe is a strength of our study.

Sample size

The number of participants included in a research study, or the sample size, is an important consideration. Decisions on sample sizes for quantitative studies are usually based on statistical sample size estimations, which calculate the statistical power – the ability of a study to enable detection of a statically significant difference when one truly exists.[31] Including too few participants will result in statistically inconclusive outcomes; too many participants, however, will either expose a higher number of participants to the procedure or will be less feasible as a result of the time consumed and its costs. In our cross-sectional reliability study (described in **Chapter 3**), we were able to include 28 paediatric physiotherapists instead of the usual 2 or 3 raters. We based sample size estimations for our study on the number of therapists included, reasoning that including a large number of therapists would increase the generalisability of the results to clinical practice.[32] One could argue that only a relatively small group of children was included. However, our decision on the number of children to be included was based on how much time we could reasonably expect the physiotherapist to give to the study; we estimated a total time investment of 6 hours to observe and score children as reasonable. In 6 hours, we estimated therapists to be able to observe and score 9 video-recorded children; we took those numbers into account during our estimations of sample size. Furthermore, including children with a variety of diagnoses representative of daily paediatric physiotherapeutic practice contributed to the generalisability of the results. However, a limitation for our study was the inclusion of two children with Down syndrome. Conversations with therapists revealed that they found it challenging to score these children using the OMQ scale. This was possibly because the paediatric physiotherapists involved were trained to use developmental trajectories for children with Down syndrome as reference values while observing them. In our study, the therapists had to change perspective and compare their observations with typical development. As the scoring differences on the OMQ scale demonstrated, this change in perspective proved difficult. In the development of the training for the OMQ scale, we will take these perceived difficulties into account and expand the focus on observation regardless of expected motor performance for certain diagnoses or syndromes.

In our prospective intervention study (detailed in **Chapter 5**), our aim was to estimate interrater and intrarater reliability and responsiveness of the OMQ scale. We based sample size calculations on estimations of reliability.33 Although a larger group of children was observed in this study – increasing generalisability of the result to clinical practice – no statistical estimations of sample size required for detecting important change were performed.34 We followed the COSMIN2 guidelines, in which it was explained that there are no standards for sample size responsiveness in single studies in which it is possible to pool the results. Although COSMIN gave no standards, as a rule of thumb, a minimum of 50 patients is often mentioned, which was not feasible for our study due to the time investment for the participating therapists, and could thereby be seen as a limitation. Had we been able

to include more children in our study, this could have increased the determination of the usefulness of the outcomes.

Recommendations

Recommendations for future research and clinical practice

Our findings on validity, reliability and responsiveness indicate that the OMQ scale can be used in clinical practice. So far, however, measurement properties for the OMQ scale have only been examined in Dutch paediatric physiotherapy practices. Measurement instruments are often population- and context-specific, and this should be taken into account before generalising to other populations.[30] Consequently, it is necessary to investigate whether the measurement properties are consistent with other countries and other populations of children, and therefore, a future study should assess whether they adequately reflect the purpose of the OMQ scale.[35] Future studies of the OMQ scale that include data for children treated in multiple paediatric physiotherapy contexts, in other countries and within other patient populations (e.g., neuromuscular diseases and syndromes) will provide further evidence of validity for the use of the OMQ scale in clinical practice. Moreover, although in our studies we have investigated measurement properties in children from 6 months to 12 years of age, measurement properties for the age frame of 12–16 years have not yet been investigated. The collection of data in clinical practice from a large group of children – also of older ages – within multiple patient populations and with intervention over a 6-month period will benefit the generalisability of the results. Furthermore, the interpretability – the degree to which one can assign a clinical meaning to the scores or changes in scores[1] – of the OMQ scale has yet to be determined. A future study should investigate norm references for typical and deviant movement quality and reference values for changes in scores; this will expand the usefulness for the OMQ scale in clinical practice.

Recommendations for education

From our qualitative study, we concluded that, to learn observational skills, students must take the lead in their own learning process. To stimulate students to take responsibility for their own learning, it is recommended to enable them to exercise observations of human movement in a realistic context to encourage them to learn together, which must be facilitated by an experienced teacher. Our proto-theory, displaying the didactic principles with their underlying educational theories, can inform the design of an educational program to teach observational skills during bachelor's and master's education.

The recommendation to involve an experienced teacher to stimulate students to take the lead in their own learning process might imply a need for additional training for teachers to strengthen their skills. This training should focus on how teachers can tailor instruction or coaching to the actual level of performance needed. For this, the teacher needs to develop skills for behaving as a coach, not as a lecturer. Other important didactic skills to address

in this training are how to stimulate critical reflection in students on performances – i.e., reflection in action and reflection on action – and how to stimulate a critical attitude towards one's own and others' performances. By designing a meaningful learning environment with tasks applicable to professional clinical practice, in which students are encouraged to work and learn together, the student can learn to take responsibility for his or her own education. Another recommendation for the design of an educational program is to keep in mind that a precondition for learning observational skills is a realistic context. This might imply enabling students to observe real – or standardised – patients to increase possibilities for observation in a clinical context. Observing real patients will allow the students to actively engage in developing observational and clinical reasoning skills[8,10–12] and facilitate competence building to diagnose patients and design interventions to improve functional abilities. In addition, possibilities for working together with peers, explaining ideas and questioning each other's views will help the students develop a critical attitude toward their own and others' performances. Thus, students will take responsibility for their own learning process once they are enabled to exercise observations of human movement and are encouraged to learn together (facilitated by an experienced teacher), which should be the basis for the design of an educational program.

For further education evaluation, it is best to: continue the iterative process of the development of educational materials for the development of observational skills with students, teachers and practitioners to complete the implementation of materials in educational contexts; and construct EPAs – including descriptions of domains of competence and milestones – for both bachelor's and master's education to guide and assess essential expectations for learning skills and mastering abilities in a physiotherapist's professional development.

References

1. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res*. 2010;19(4):539-549.
2. Mokkink LB., de Vet HCW, Prinsen CAC, et al. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. *Qual Life Res*. 2018;27(5):1171-1179.
3. Skjaerven LH, Kristoffersen K, Gard G. An eye for movement quality: a phenomenological study of movement quality reflecting a group of physiotherapists' understanding of the phenomenon. *Physiother Theory Pract*. 2008;24(1):13-27.
4. Knudson D V, Morrison CS. Interdisciplinary nature of qualitative analysis. In: *Qualitative Analysis of Human Movement*. 2nd editio. Leeds: Human Kinetics; 2002.
5. Boudreau JD, Cassell EJ, Fuks A. Preparing medical students to become skilled at clinical observation. *Med Teach*. 2008;30(9-10):857-862.
6. Berger L. A hierachy of clinical observations. *Pediatrics*. 1980;65(2):357-358.
7. Janssen AJWM, Diekema ETW, Dolder van R, Kollée LAA, Oostendorp RAB, Nijhuis-van derSanden MWG. Development of a movement quality measurement tool for children. *Phys Ther*. 2012;92:574-594.
8. Wijbenga MH, Bovend'Eerdt TJH, Driessen EW. Physiotherapy students' experiences with clinical reasoning during clinical placements: A qualitative study. *Heal Prof Educ*. 2019;5(2):126-135.
9. Huhn K, Gilliland SJ, Black LL, Wainwright SF, Christensen N. Clinical reasoning in physiotherapy: A concept analysis. *Phys Ther*. 2019;99(4):440-456.
10. Black LL, Jensen GM, Mostrom E, et al. The first year of practice: an investigation of the professinal learning and development of promising novice physiotherapists. *Phys Ther*. 2010;90:1758-1773.
11. Hayward LM, Black LL, Mostrom E, Jensen GM, Ritzline PD, Perkins J. The first two years of practice: a longitudinal perspective on the learning and professional development of promising novice physiotherapists. *Phys Ther*. 2013;93(3):369-383.
12. Jensen GM, Shepard KF, Hack LM. The novice versus the experienced clinician: insights into the work of the physiotherapist. *Phys Ther*. 1990;70:314-323.
13. Frank J, Fréchette D, Marks M, Valk N, Bourgeois G. *Report of the CanMEDS Phase IV Working Groups*. Ottawa; 2005.
14. de Vries C, Hagenaars L, Kiers H, Schmitt M. *Beroepsprofiel Fysiotherapeut*; 2014.
15. Peppen R. *Landelijk Opleidingsprofiel Bachelor Fysiotherapy. SROF*.; 2017.
16. ten Cate O. Nuts and bolts of Entrustable Professional Activities. *J Gen Intern Med*. 2013;5(1):157-158.
17. Jensen GM, Nordstrom T, Segal RL, McCallum C, Graham C, Greenfield B. Education research in physiotherapy: Visions of the possible. *Phys Ther*. 2016;96(12):1874-1884.
18. ten Cate O. Entrustability of professional activities and competency-based training. *Med Educ*. 2005;39(12):1176-1177.

19. Chesbro SB, Jensen GM, Boissonnault WG. Entrustable professional activities as a framework for continued professional competence: Is now the time? *Phys Ther*. 2018;98(1):3-7.
20. Furze JA, Tichenor CJ, Fisher BE, Jensen GM, Rapport MJ. Physiotherapy residency and fellowship education: Reflections on the past, present, and future. *Phys Ther*. 2016;96(7):949-960.
21. ten Cate O, Graafmans L, Posthumus I, Welink L, van Dijk M. The EPA-based Utrecht undergraduate clinical curriculum: Development and implementation. *Med Teach*. 2018;40(5):506-513.
22. Boudreau JD, Cassell EJ, Fuks A. Preparing medical students to become skilled at clinical observation. *Med Teach*. 2008;30(9-10):857-862.
23. Shapiro J, Rucker L, Beck J. Training the clinical eye and mind: Using the arts to develop medical students' observational and pattern recognition skills. *Med Educ*. 2006;40(3):263-268.
24. Al-Moteri MO, Plummer V, Cooper S, Symmons M. Training paradigms to enhance clinical observational skills in clinical practice : A scoping review. *J Nurs Educ Pract*. 2015;5(12):96-103.
25. Jarodzka H, Scheiter K, Gerjets P, van Gog T. In the eyes of the beholder: How experts and novices interpret dynamic stimuli. *Learn Instr*. 2010;20(2):146-154.
26. Schuller W, Ostelo RWJG, Janssen R, de Vet HCW. The influence of study population and definition of improvement on the smallest detectable change and the minimal important change of the neck disability index. *Health Qual Life Outcomes*. 2014;12(1):1-7.
27. De Vet HCW, Foumani M, Scholten MA, et al. Minimally important change values of a measurement instrument depend more on baseline values than on the type of intervention. *J Clin Epidemiol*. 2015;68(5):518-524.
28. Kamper SJ, Ostelo RWJG, Knol DL, Maher CG, de Vet HCW, Hancock MJ. Global Perceived Effect scales provided reliable assessments of health transition in people with musculoskeletal disorders, but ratings are strongly influenced by current status. *J Clin Epidemiol*. 2010;63(7):761-766.
29. Grøvle L, Haugen AJ, Hasvik E, Natvig B, Brox JI, Grotle M. Patients' ratings of global perceived change during 2 years were strongly influenced by the current health status. *J Clin Epidemiol*. 2014;67(5):508-515.
30. De Vet HCW, Ostelo RWJG, Terwee CB, et al. Minimally important change determined by a visual method integrating an anchor-based and a distribution-based approach. *Qual Life Res*. 2007;16(1):131-142.
31. Eng J. Sample size estimation: How many individuals should be studied? *Radiology*. 2003;227(2):309-313.
32. van Genderen FR, de Bie RA, Helders PJM, Meeteren NLU. Reliability research: towards a more clinically relevant approach. *Phys Ther Rev*. 2003;8:169-176.
33. Shoukri MM, Asyali MH, Donner A. Sample size requirements for the design of reliability study: review and new results. *Stat Methods Med Res*. 2004;13(4):251-271.
34. Terwee CB, Roorda LD, Knol DL, De Boer MR, De Vet HCW. Linking measurement error to minimal important change of patient-reported outcomes. *J Clin Epidemiol*. 2009;62(10):1062-1067.
35. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*. 2010;63(7):737-745.

Chapter 8

Summary

Summary

While growing up, children learn to use countless motor skills. This development starts with learning basic motor skills – such as sitting, walking, reaching and grasping, chewing and talking – followed by more specific skills based on the needs and demands of the environment – such as riding a bike, playing a hockey match or writing a letter. During the development of motor skills, both quantitative and qualitative changes in performance can be observed. Quantitative changes give information on the motor skills a child is able to perform, while qualitative changes give information on how movements are executed. For a physiotherapist, the observation of movements is both essential for diagnoses and a basis for intervention to improve functional abilities. For the assessment of movement quantity, valid motor tests are available. However, no generic test is available to assess movement quality in children over time and for all age categories.

Chapter 1 introduces background information and the aims and outline of the thesis. The chapter focuses on the definition and terminology of movement quality, observation, education and validation of measurement instruments. We describe the development and clinical use of the Observable Movement Quality (OMQ) scale as well as what is known about the observational skills of students in physiotherapy.

This thesis is divided into 2 parts. The aim of **PART 1** (Chapters 2–5) is to determine the reliability validity and responsiveness of the OMQ scale. In **PART 2** (Chapter 6), the aim is to investigate what students in physiotherapy need in their educational program to develop observational skills and which didactic principles facilitate this learning.

PART I: Measurement properties of the OMQ scale

The first aim of this thesis was to provide insights on the measurement properties of the OMQ scale by the determination of the reliability, validity and responsiveness of the scale. All studies followed the Consensus based standards for the Selection of Health Measurement Instruments (COSMIN), and we studied them in the context of physiotherapeutic care for children.

In **Chapter 2** we describe the determination of the construct validity of the OMQ scale using 7 hypotheses. For the 7 hypotheses, we defined direction, magnitude and rationale, concerning the relationships between OMQ scores and the severity of motor disabilities, or the outcomes of motor tests (n = 2), the probability of low scores on specific OMQ items in children diagnosed with spasticity, psychomotor retardation, mitochondrial diseases or ataxia (n = 4), and the difference in the level of OMQ scores between diagnosis subgroups (n = 1). Data collection primarily took place as part of a multidisciplinary assessment during

diagnostic trajectories for children suspected of mitochondrial dysfunction or disease. We chose this trajectory because these children show either a wide range of motor problems – with additional signs and symptoms – or an almost normal development. To ensure even sample sizes per age group, for gender and for a diversity in diagnosis, we added data from cases of outpatient multidisciplinary evaluations from other trajectories (e.g., children born preterm or diagnosed with ataxia telangiectasia). A paediatric physiotherapist assessed motor performance in children using an age-specific motor test and the OMQ scale. We conformed 6 out of 7 hypotheses, indicating sufficient construct validity. We found a significant positive relationship between OMQ total scores and severity of motor disabilities ($r = 0.72$) and z-scores on motor tests ($r = 0.60$). We confirmed probabilities for low scores on OMQ items for children diagnosed with spasticity, psychomotor retardation, mitochondrial diseases and ataxia – exceeding chi-square's critical value – except probabilities for low scores on strength regulation for children with ataxia. The OMQ total scores for children who were non-ambulant due to neurological conditions were significantly different from the scores for children who were non-ambulant due to fatigue ($r = 0.66$). However, our sample of children was based on theoretical assumptions about relevant variations in clinical representations; based on our results, it appears that children with low strength regulation were underrepresented. The confirmation of nearly all hypotheses supports the validity of the OMQ scale for measurement of movement quality in clinical practice in addition to standardised age-adequate motor performance tests.

For **Chapter 3** we studied the concurrent and predictive validity for movement quality for the OMQ scale, comparing outcomes measured using the OMQ scale and General Movements assessment (GMs) at three months of age. We based our correlations on data collected in a prospective longitudinal cohort study for individual neurodevelopmental trajectories over 5 years in children treated with hypothermia for perinatal asphyxia. In this longitudinal cohort study, we assessed 18 children at 3 (t_1), 6 (t_2), 12 (t_3) and 24 (t_4) months, and at the age of 5 (t_5) years, with standardised norm-referenced tests. Of these 18 children, 6 showed abnormal movement quality assessed with GMs (t_1), and all showed severe neurodevelopmental disabilities at t_5. Correlations for GMs and the assessment of movement quality, scored on the OMQ scale, was moderate to good at 3 months ($r_s = 0.65$). The correlation between assessment of motor development, using Movement Assessment Battery for Children (MABC) and GMs ($r_s = 0.84$), was high and significant, as was the correlation between the MABC and OMQ total scores ($r_s = 0.75$). These outcomes indicate that the OMQ scale can be considered as an alternative for GMs in children at risk for developmental delays; however, more research is needed with larger groups of children and with different diagnoses before final conclusions can be drawn.

For **Chapter 4** we investigated the interrater reliability – including the Standard Error of Measurement, and the limits of agreement (LoA) – of the OMQ scale in a cross-sectional

study in with a stratified sample of paediatric physiotherapists (*n* = 28) with a variety of clinical expertise based on work setting and work experiences. They observed video-recorded assessments of age-related motor tests in children (*n* = 9) aged 6 months to 6 years and filled in the OMQ scale. All therapists were unfamiliar with the OMQ scale and received a 2-hour training to gain insight into the aim of the scale and movement quality. For analyses, we used linear mixed models without fixed effects, and the results showed moderate interrater reliability ($ICC_{2,1}$: 0.67, 95% CI [0.47, 0.88]) with no influence from work setting or work experience. The standard deviation of the random measurement error was 5.7, and the LoA, 31.5. Item agreement was good (P_O total: 0.82–0.99). The OMQ scale seems a promising tool to test movement quality; however, feedback from participating physiotherapists suggested a need for a more comprehensive training for use of the scale in clinical practice.

For **Chapter 5**, the reliability and responsivity of the OMQ scale were investigated in a prospective intervention study with a pre–post design, conducted in centres for paediatric physiotherapy practice. For this study, 3 paediatric physiotherapists observed 30 video-recorded assessments of children – aged 4–12 years – using the MABC and the OMQ scale. To determine intrarater reliability, 1 physiotherapist scored baseline assessments for a second time. We tested responsiveness by comparing outcomes before and after intervention. Interrater reliability was moderate to good (intra-class correlation coefficient [ICC2,1]: 0.79; 95% CI: 0.62, 0.89); intrarater reliability was high (ICC2,1: 0.97; 95% CI: 0.93, 0.98). Responsiveness results revealed a Smallest Detectable Change of 2.38 for OMQ total scores and a Minimal Important Change (MIC) of 3.15. Based on a receiver operating curve, a MIC of 2.5 (sensitivity 84%, specificity 77%) was shown (area under the curve of 0.77). The results of this study demonstrate a moderate to good interrater reliability for the OMQ scale, a high intrarater reliability and a scale that is responsive to change when used to assess movement quality in children aged 4–12 years. Altogether, the findings indicate that the OMQ scale is valid for measuring movement quality in clinical practice in addition to motor performance tests.

PART II: Observations of movement quality

The second aim of this thesis was to define the needs of students in physiotherapy in their educational program to develop observational skills as well as to investigate which didactic principles facilitate this learning. We studied this in diverse educational contexts.

Chapter 6 describes the development of a proto-theory for an educational program for physiotherapy students to learn observational skills. To develop this proto-theory, we derived design principles from students, teachers, practitioners and researchers using a qualitative approach within a design-based methodology. In 4 rounds, 8 physiotherapy

students, 16 teachers and 9 practitioners participated in 5 Nominal Group Technique meetings and 6 interviews. We transcribed meetings and interviews verbatim and analysed them using thematic analysis. Three themes were identified, each with several design principles: didactics, professional content and conditions for optimal learning. We developed a proto-theory with underlying educational theories. We explained that students must take the lead in their own learning process, facilitated by an experienced teacher to learn observational skills. This might imply a need for additional training for teachers to strengthen their didactic skills. Another precondition for learning is a realistic context; it might be necessary to increase possibilities for observations in clinical contexts or to invest in training for (simulated) patients as participants in education. Further research is needed to test the applicability of the design principles and the proto-theory for other professionals with a focus on observation and analysis of movements.

Finally, **Chapter 7** contains a general discussion of the main findings of the studies from Chapters 2–6, which are described and discussed within 2 different themes: 'Measurement properties of the OMQ scale' and 'Observation of movement quality'. A description of the development of a competency framework for the physiotherapy profession in the Netherlands follows, including ways in which the quality of movement observation fits into a competency-based physiotherapy profile. A challenge in competency-based physiotherapy curricula is finding ways to structure and assess the essential integration of knowledge, skills and attitudes. We discuss how to facilitate the translation from a competency-based profile to an educational setting. To facilitate this translation, we adopted and explained the entrustable professional activity (EPA) concept. We present a concept for a framework to develop EPAs for physiotherapy curricula. This framework includes examples of the development and integration of observational skills in an EPA. This can be used and further developed in the implementation of teaching and assessment of observational skills in physiotherapy curricula.
Subsequently, we describe issues to consider related to methodological choices. In a recommendation paragraph, we offer suggestions for future research in clinical practice, as well as recommendations for education. Finally, we provide an overview of possible next steps within the field of education. We argue that the necessary next step will be the construction of EPAs for both bachelor's and master's education, which can guide and assess essential expectations for learning skills and mastering abilities in physiotherapists' professional development.

Nederlandse samenvatting

Tijdens het opgroeien leren kinderen ontelbare motorische vaardigheden. De ontwikkeling van deze motorische vaardigheden start met het leren van basisvaardigheden zoals zitten, lopen, reiken en pakken, kauwen en zelf eten. Hierna ontwikkelt een kind meer specifieke motorische vaardigheden, waarbij gedacht kan worden aan fietsen, schrijven of vaardigheden om mee te kunnen doen aan een sportwedstrijd. Deze specifieke vaardigheden zijn gebaseerd op behoeften en eisen vanuit de omgeving en geven het kind de mogelijkheid om mee te doen aan de samenleving. Gedurende de ontwikkeling van motorische vaardigheden vallen zowel kwantitatieve als kwalitatieve veranderingen in de uitvoering op. Kwantitatieve veranderingen geven informatie over *welke* vaardigheden een kind op dat moment beheerst, terwijl de kwalitatieve veranderingen een indruk geven over *hoe* deze vaardigheden worden uitgevoerd. Voor een fysiotherapeut is het observeren van bewegen essentieel. Deze observaties zijn noodzakelijk om diagnoses te kunnen stellen waarop behandelingen worden gebaseerd die gericht zijn op het vergroten van het functioneel bewegen van het kind. Voor het meten van de kwantiteit van motorische vaardigheden bij kinderen zijn betrouwbare meetinstrumenten beschikbaar. Echter, voor het meten van de kwaliteit van motorische vaardigheden was er geen meetinstrument beschikbaar dat gebruikt kon worden voor kinderen van alle leeftijden en diagnosegroepen en dat in staat was de veranderingen in de kwaliteit van bewegen over de tijd vast te leggen.

In **Hoofdstuk 1** vindt u algemene achtergrondinformatie, de doelen en de opzet van dit proefschrift. Er wordt uitgelegd wat verstaan kan worden onder kwaliteit van bewegen, hoe deze te observeren, wat er nodig is om dit aan te leren en aan welke eisen meetinstrumenten moeten voldoen. Verder wordt de ontwikkeling en het gebruik van de Observable Movement Quality (OMQ) schaal beschreven, het meetinstrument waar dit proefschrift op voortbouwt. Ook wordt ingegaan op wat er al bekend is over de vaardigheid 'observeren' voor studenten fysiotherapie.

Dit proefschrift is onderverdeeld in twee gedeelten. Het doel van **Deel 1** (hoofdstuk 2–5) is om de betrouwbaarheid, validiteit en responsiviteit van de OMQ-schaal te bepalen. In **Deel 2** (hoofdstuk 6) is het doel om te onderzoeken wat studenten fysiotherapie nodig hebben in hun opleiding om de vaardigheid observeren te leren en welke didactische principes dit leren kunnen faciliteren.

Deel 1: Psychometrische eigenschappen van de OMQ-schaal
Tijdens de ontwikkeling van meetinstrumenten worden de psychometrische eigenschappen van het betreffende meetinstrument bepaald. Deze psychometrische eigenschappen geven inzicht in de mate waarin het instrument het construct meet waarvoor het instrument bedoeld is, of dit betrouwbaar gebeurt en hoe groot de meetfout is waar rekening mee

gehouden moet worden. Kennis over de meeteigenschappen en de focus van een instrument helpt de clinicus om het meest geschikte meetinstrument te kiezen. Het eerste doel van dit proefschrift is inzicht geven in de betrouwbaarheid, validiteit en responsiviteit van de OMQ-schaal. Bij onze studies is de Consensus based standards for the Selection of Health Measurement Instruments (COSMIN) gevolgd als leidraad. Daarnaast zijn de onderzoeken zoveel mogelijk uitgevoerd in de dagelijkse praktijk.

Hoofdstuk 2 beschrijft de bepaling van de constructvaliditeit, ook wel begripsvaliditeit genoemd, van de OMQ-schaal. Met de bepaling van de constructvaliditeit wordt vastgesteld of de meetresultaten van de OMQ-schaal, die voorspeld worden op basis van theoretische kennis, overeenkomen met wat er in de klinische praktijk daadwerkelijk gemeten wordt. De constructvaliditeit van de OMQ-schaal werd bepaald met behulp van zeven vooraf opgestelde hypotheses. Voor deze zeven theoretisch onderbouwde hypotheses werden zowel de richting als de sterkte van het verband voorspeld. De hypotheses hadden betrekking op: de relatie tussen de OMQ-totaalscore en de ernst van de motorische beperkingen en de uitkomsten op motorische testen (n = 2), de waarschijnlijkheid van lage scores op specifieke OMQ-items voor kinderen gediagnostiseerd met spasticiteit, verstandelijke beperking, mitochondriële ziekten (stofwisselingsziekten) of ataxie (n = 4) en verschillen in de hoogte van de OMQ-scores tussen twee subgroepen binnen één diagnose (n = 1). De dataverzameling vond voornamelijk plaats als onderdeel van multidisciplinaire diagnostiek bij kinderen met verdenking op een mitochondriële aandoening of ziekte. Voor deze groep werd gekozen omdat bekend is dat bij deze kinderen milde, matige of zeer ernstige motorische problemen kunnen voorkomen. Om ervoor te zorgen dat we gelijke groepsgrootten hadden voor geslacht, leeftijd en diversiteit aan diagnoses, werd data van poliklinische multidisciplinaire evaluaties van andere diagnosegroepen toegevoegd (bijvoorbeeld: kinderen die prematuur geboren werden of kinderen met de diagnose ataxie). Eén kinderfysiotherapeut onderzocht de motorische vaardigheden van kinderen waarbij de OMQ-schaal werd gebruikt om de kwaliteit van bewegen vast te leggen naast een leeftijdsgerelateerde motorische test om de kwantiteit van bewegen vast te leggen. Zes van de zeven hypotheses werden bevestigd, waarmee kon worden vastgesteld dat de constructvaliditeit toereikend is. Resultaten lieten een significant positieve relatie zien tussen de OMQ-totaalscore en ernst van de motorische beperkingen (r = 0.72) en de uitkomsten op motorische testen (r = 0.60). De waarschijnlijkheid voor lage scores op OMQ-items werd bevestigd voor kinderen gediagnosticeerd met spasticiteit, verstandelijke beperking, mitochondriële aandoeningen en ataxie. De verwachte waarschijnlijkheid voor lage item scores voor krachtregulatie bij kinderen met ataxie werd echter niet bevestigd. Uit de resultaten bleek verder dat de OMQ-totaalscores voor kinderen die niet ambulant zijn door een neurologische aandoening, significant verschillen van kinderen die niet ambulant zijn door vermoeidheid (r = 0.66). Hoewel de inclusie van de kinderen aan dit onderzoek gebeurde op basis van theoretische aannames over relevante variaties in motorische functioneren, blijkt op basis van onze

resultaten dat kinderen met verminderde krachtregulatie ondervertegenwoordigd waren. Met de bevestiging van vrijwel alle hypotheses wordt aangetoond dat de OMQ-schaal een valide instrument is om de kwaliteit van bewegen te meten in de praktijk. Deze metingen van de kwaliteit van bewegen kunnen worden uitgevoerd als aanvulling op gestandaardiseerde leeftijdsgerelateerde motorische testen die de kwantiteit van bewegen vastleggen.

In **Hoofdstuk 3** werden de concurrente en predictieve validiteit van de OMQ-schaal bestudeerd. De concurrente validiteit beschrijft in hoeverre de resultaten van de OMQ-schaal overeenkomen met resultaten die worden gemeten met een ander meetinstrument bedoeld om de kwaliteit van bewegen vast te leggen. Predictieve validiteit heeft betrekking op de vraag in hoeverre de resultaten van de OMQ-schaal kunnen voorspellen hoe de motorische vaardigheden van een kind zich verder zullen ontwikkelen. Om de concurrente en predictieve validiteit van de OMQ-schaal te onderzoeken, werden meetresultaten van de OMQ-schaal vergeleken met die van de General Movements assessments (GM) bij kinderen van 3 maanden oud. Deze data werden verzameld gedurende een prospectieve longitudinale cohortstudie voor kinderen die zuurstofgebrek opliepen rond de geboorte en hiervoor in het ziekenhuis behandeld werden met hypothermie. Tijdens deze cohortstudie werden individuele ontwikkelingstrajecten van de betrokken kinderen gedurende vijf jaar gevolgd. In totaal werden 18 kinderen onderzocht op een leeftijd van 3 (t_1), 6 (t_2), 12 (t_3) en 24 (t_4) maanden en op 5 (t_5) jaar, waarbij gestandaardiseerde en genormeerde testen werden gebruikt. Van deze 18 kinderen vertoonden zes kinderen een abnormale kwaliteit van bewegen gemeten met de GM op 3 maanden (t_1); zij vertoonden allen ernstige neurologische aandoeningen op vijfjarige leeftijd. De correlatie tussen de uitkomsten van de GM en de uitkomsten van de OMQ-schaal op de leeftijd van drie maanden was matig tot goed (r_s = 0.65). De correlatie tussen de uitkomsten op een motorische test (Movement Assessment Battery for Children [MABC]) op 5 jaar en de GM op 3 maanden was hoog en significant (r_s = 0.84) evenals de correlatie tussen totaalscores van de MABC en de OMQ-schaal (r_s = 0.75). De uitkomsten geven aan dat de OMQ-schaal kan worden overwogen als een alternatief voor de GM bij kinderen die risico lopen op een vertraagde of gestoorde ontwikkeling. Er is echter meer onderzoek nodig in een grotere groep kinderen en met verschillende diagnoses om de gevonden trend te bevestigen.

In **Hoofdstuk 4** werd onderzocht in hoeverre metingen van de OMQ-schaal vrij zijn van meetfouten. Hiervoor werd de interbeoordelaarsbetrouwbaarheid bepaald, inclusief de standaard meetfout (Standard Error of Measurement [SEM]) en de spreiding in overeenstemming (Limits of Agreement [LOA]). In een cross-sectioneel onderzoek met een gestratificeerde steekproef van kinderfysiotherapeuten (n = 28) vanuit diverse werksettingen en met een uiteenlopende mate van werkervaring werd met behulp van video's de betrouwbaarheid onderzocht. De deelnemende kinderfysiotherapeuten observeerden video's van kinderen die getest werden met een leeftijdsgerelateerde motorische test (n

= 9). De kinderen varieerden in leeftijd van 6 maanden tot 6 jaar en hadden verschillende diagnoses. Aan de hand van de video-opnamen werd een OMQ-schaal scoreformulier ingevuld. Alle therapeuten waren vooraf getraind in het afnemen van de schaal in een twee uur durende training om inzicht te krijgen in wat de doelstellingen van de OMQ-schaal zijn en om te leren hoe de kwaliteit van bewegen gescoord kan worden gebruikmakend van de verschillende items. Voor statistische analyses werden linear mixed models gebruikt. De resultaten lieten een matige interbeoordelaarsbetrouwbaarheid zien (intraclass correlation coefficient [$ICC_{2,1}$]: 0.67, 95% CI [0.47, 0.88]). De werksetting of werkervaring bleken geen invloed op de scores te hebben. De standaarddeviatie van de random meetfout was 5.7 en de LoA 31.5. De itemovereenkomst was goed (P_o totaal: 0.82 tot 0.99). De OMQ-schaal lijkt een veelbelovend meetinstrument om de kwaliteit van bewegen te meten, maar op basis van de feedback van deelnemende fysiotherapeuten kon geconcludeerd worden dat een uitgebreidere training nodig is om de OMQ-schaal beter te kunnen gebruiken.

Hoofdstuk 5 geeft een beschrijving van het onderzoek naar de betrouwbaarheid en responsiviteit van de OMQ-schaal. Dit werd onderzocht in een prospectieve interventiestudie met een pre post design, in eerstelijnspraktijken voor kinderfysiotherapie. Voor deze studie observeerden drie kinderfysiotherapeuten 30 video's van kinderen (in de leeftijd van 4 tot 12 jaar) waarbij met behulp van de MABC motorische vaardigheden werden getest. Tijdens de observatie van de video's werden alle items van de OMQ-schaal gescoord. Om de intrabeoordelaarsbetrouwbaarheid te bepalen werden de baseline metingen voor een tweede keer gescoord door één van de drie kinderfysiotherapeuten. Responsiviteit werd bepaald door de uitkomsten voor en na de interventie te vergelijken. De interbeoordelaarsbetrouwbaarheid was matig tot goed ($ICC_{2,1}$: 0.79; 95% CI: 0.62, 0.89) en de intrabeoordelaarsbetrouwbaarheid was hoog ($ICC_{2,1}$: 0.97; 95% CI: 0.93, 0.98). De resultaten voor responsiviteit lieten een smallest detectable change van 2.38 zien voor de OMQ-totaalscores en een minimal important change (MIC) van 3.15. Een receiver operating curve liet een MIC van 2.5 zien (sensitiviteit 84% en specificiteit 77%) met een area under the curve van 0.77. De resultaten demonstreren een matige tot goede interbeoordelaarsbetrouwbaarheid voor de OMQ-schaal en een hoge intrabeoordelaarsbetrouwbaarheid. De resultaten laten zien dat de OMQ-schaal responsief is voor verandering wanneer deze gebruikt wordt om de kwaliteit van bewegen te meten bij kinderen in de leeftijd van 4 tot 12 jaar. De resultaten geven daarmee aan dat de OMQ-schaal valide gebruikt kan worden in de praktijk om de kwaliteit van bewegen te meten bij deze doelgroep van kinderen.

Deel II: Het observeren van de kwaliteit van bewegen

Het tweede doel van deze thesis was enerzijds om te onderzoeken wat studenten fysiotherapie nodig hebben tijdens hun opleiding om observatievaardigheden te ontwikkelen en anderzijds om te bepalen welke didactische principes dit leren kunnen faciliteren.

Hoofdstuk 6 beschrijft de ontwikkeling van een proto-theorie voor een onderwijskundig programma voor fysiotherapie studenten om te leren observeren met behulp van kwalitatief onderzoek met een design-based benadering. In vier rondes namen acht studenten fysiotherapie, 16 docenten fysiotherapie en 9 fysiotherapeuten deel aan vijf Nominale Group Technique bijeenkomsten en zes interviews. De bijeenkomsten en interviews werden woord voor woord uitgeschreven, waarna deze thematisch werden geanalyseerd. Er werden drie thema's geïdentificeerd, elk met meerdere design principes: didactiek, professionele content en condities voor optimaal leren. Vervolgens werd er een proto-theorie ontwikkeld welke werd onderbouwd vanuit onderwijstheorieën. In dit onderzoek kwam naar voren dat studenten de leiding moeten nemen in het eigen leerproces, gefaciliteerd door een ervaren docent, om de vaardigheid observeren te leren. Dit kan betekenen dat er extra aandacht moet komen om docenten te trainen zodat hun didactische vaardigheden worden vergroot. Een andere voorwaarde voor leren is een realistische context: het is wellicht noodzakelijk om de mogelijkheden voor studenten te vergroten om observaties in een klinische context te laten plaatsvinden, of om te investeren in training voor (simulatie) patiënten om te participeren in het onderwijs. Verder onderzoek is nodig om de bruikbaarheid van de design principes en proto-theorie te testen voor andere professionals die eveneens een focus hebben op observeren en analyseren van beweging.

Het laatste hoofdstuk, **Hoofdstuk 7**, bediscussieert de bevindingen van dit proefschrift uit hoofdstuk 2–6. Deze bevindingen zijn beschreven en worden bediscussieerd in twee verschillende thema's: 'Psychometrische eigenschappen van de OMQ-schaal' en 'het observeren van de kwaliteit van bewegen'. Hierna volgt een beschrijving van het competentieprofiel van de fysiotherapeut en de wijze waarop observatie van bewegen in een op competenties gebaseerd opleidingsprofiel voor fysiotherapeuten past. Een uitdaging in een op competenties gebaseerde curriculum voor fysiotherapie, is het vinden van manieren om de integratie van kennis, vaardigheden en attitudes te structureren en te beoordelen. Bediscussieerd wordt hoe de vertaling van een competentieprofiel naar een onderwijskundige setting gemaakt kan worden. Om deze vertaling te kunnen maken is het Entrustable Professional Activity (EPA) concept verduidelijkt. Vervolgens wordt een concept voor een kader om EPA's te ontwikkelen voor fysiotherapie curricula gepresenteerd. Dit kader bevat voorbeelden van de ontwikkeling en integratie van de vaardigheid observeren in een EPA. Deze kan gebruikt en verder ontwikkeld worden bij het implementeren van onderwijs en de beoordeling van de vaardigheid observeren in curricula voor fysiotherapie. Ook worden methodologische keuzes en overwegingen voor de onderzoeken beschreven. In een paragraaf met aanbevelingen worden suggesties voor toekomstig onderzoek en voor het onderwijs gedaan. Afsluitend wordt een overzicht gegeven over mogelijke vervolgstappen binnen het onderwijs. Beargumenteerd wordt dat de noodzakelijke vervolgstap de constructie van EPA's zal zijn, voor zowel bachelor als masteronderwijs. Deze EPA's kunnen verwachtingen managen met betrekking tot het leren van de vaardigheid observeren en

het opdoen van ervaring tijdens de hierop volgende professionele ontwikkeling van de fysiotherapeut.

Appendices

Data Management
PhD portfolio
Curriculum Vitae
Dankwoord

DATA MANAGEMENT

General information about the data collection

Each study protocol in this PhD trajectory was submitted to the local Medical Ethics Committee CMO Arnhem-Nijmegen (Chapter 2–5) or the Ethical Advisory committee at the Faculty of Health of the HAN University of Applied Sciences (Chapter 6). All studies were officially declared exempt from ethical approval for human subjects' research. Participants volunteered to participate, and anonymity and confidentiality were assured. We obtained informed consent for collecting data from the participants and/or from their parents (or legal representatives) for those younger than 18 years. All procedures performed in the reported studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee, and with the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards.

FAIR PRINCIPLES

Findable

Raw and processed data and accompanying files (descriptive files, syntaxes etc.) for the projects in this thesis are stored in a folder on the server of the Department of Rehabilitation, accessible only by the main researchers on this project. After publication, data are stored on the secured Rehabilitation archive server in folders called 'OMQ-validation' and 'observation-MQ'.

Accessible

Only the main researchers can access the data. Paper records, including paper assessment forms, are stored in the Rehabilitation department's archive. DVDs containing video-recordings of motor assessments of children were returned to the main researchers after observation by the participating therapists and are stored in a key-locked cupboard at the Department of Rehabilitation.

It is not possible to make the data available in a public repository because participants only gave informed consent to use their data for purposes as explained on the signed informed consent form. However, requests for data can be made by contacting the staff secretary of the Department of Rehabilitation of the Radboud University Medical Center (secretariaatstaf.reval@radboudumc.nl). A suitable way to share the data will then be sought.

Interpretable

The data are stored in SPSS, SAS, or Atlas-Ti format. No existing data standards such as vocabularies, ontologies or thesauri have been used.

The involved health-care provider collected health record data, and no identifying patient information was shared with the researcher. Data collected during group meetings

and interviews were recoded. We deleted audio-recordings after finalising the coded transcriptions. We stored identifying information for participating physiotherapists separately from the data, in a secured folder on the Department of Rehabilitation server, to which only the main researchers had access. We deleted the identifying information after finishing the study.

Reusable

The data will be stored for at least 10 years after publication and can therefore also be reused within this period by the main researchers, for purposes as explained in the informed consent form signed by the participants.

PhD Portfolio

Institute for Health Sciences
Radboudumc

Name	Lieke Dekkers (v)
PhD period	September 2014 – December 2019
Department	Rehabilitation
Graduate school	Radboud Institute for Health Sciences
Promotors	Em. Prof. Dr. MWG Nijhuis-van der Sanden
	Prof. Dr. PJ van der Wees
Co-promotors	Dr. AJWM Janssen
	Dr. BJM de Swart

Activities	Year	ECTS
Training activities and courses		
RIHS PhD introduction course	2015	0.0
Effective Writing Strategies	2015	3.0
Klinimetrie: het ontwikkelen en evalueren van meetinstrumenten (V40) (afdeling epidemiologie & Biostatistiek, VU medisch centrum, Soesterberg)	2015	1.25
Academic Writing	2015	3.0
Qualitative research methods in Health Care – introduction (IQ Health Care)	2016	0.8
Presenteren Eigen Onderzoek (for PhD's)	2016	1.5
Scientific Integrity for PhD candidates	2016	0.0
BROK (EMWO, Nijmegen)	2017	1.5
Perfecting your Academic Writing Skills	2017	1.5
Basiskwalificatie Examinering (HAN)	2018	3.0
Teaching		
Bachelor Physiotherapy HAN University of Applied Sciences	2015-2019	37
Master Neurorehabilitation and Innovation HAN University of Applied Sciences	2019	3.4
Symposia, Congresses and Conferences		
WCPT Congress 2015 (Singapore) - oral presentation	2015	1.25
NVFK Jaarcongres 2015 (Utrecht)	2015	0.25
4th European Congress of ER-WCPT on Physiotherapy Education, Practice, Research (Liverpool, UK)	2016	0.5
Eindsymposium Godiva studie 2016 (Hogeschool van Utrecht)	2016	0.25
NVFK Jaarcongres 2016 (Utrecht)	2016	0.25
WCPT Congress 2017 (Cape Town, SA) – poster presentation	2017	1.25
Cohehre conference 'Integrated Care: New trends in higher education and research' 2019 (Vic, Spain)	2019	0.75
European Paediatric Physiotherapy Congress 2019 (HU) - oral presentation	2019	1.0

Curriculum vitae

Lieke Dekkers was born in Zeeland (NB), the Netherlands, on February 11th 1972. She graduated in 1990 from the higher secondary school (HAVO) at 'Kruisheren Kollege', Uden. In 1995 Lieke completed her physiotherapy education at the now called HAN University of Applied Sciences in Nijmegen. Her working life as a physiotherapist started in hospital settings at Østfold Sentralsykehus, Fredrikstad, Norway. After returning to the Netherlands, Lieke worked in private practice within paediatrics. She studied paediatric physiotherapy with the Transfer group Rotterdam, the Netherlands, from 1998 to 2001. Since 2006, Lieke was appointed a teachers' position at the Bachelor Physiotherapy of HAN University of Applied Sciences in Nijmegen, which gave her the opportunity to start her study for a Master of Science degree in Physiotherapy and Education at the University of Brighton, UK, in 2009, which she completed in 2011. Since October 2012, Lieke combines her work as teacher with a paediatric physiotherapists position at Radboudumc, Nijmegen. The Netherlands Organization for Scientific Research rewarded Lieke with a Doctoral Grant in September 2014, which enabled her to start with a PhD trajectory on the observable quality of movement at Radboud University, Nijmegen. In addition to her work for the bachelor Physiotherapy, Lieke started in 2019 as a teacher at the master Neurorehabilitation and Innovation, HAN University of Applied Sciences, Nijmegen.

Lieke Dekkers is married to Joost van Wijchen since 1998. Together they have three children: Mees, Siem and Veerle.

Dankwoord

Wat een bijzonder moment is dit. Het schrijven van het dankwoord. Dat betekent dat mijn boekje nu echt af is en dat ik antwoord kan geven aan iedereen die me in de afgelopen tijd de vraag stelde: “Wanneer ben je eigenlijk klaar met het schrijven van je proefschrift?” Trots kan ik zeggen: “Dat is nu!!” Het traject om het proefschrift te kunnen schrijven heb ik gelukkig niet alleen hoeven afleggen, maar met vele fijne mensen om me heen. Op verschillende manieren ben ik geïnspireerd, gemotiveerd, geholpen en gesteund. Mede door deze mensen heb ik de afgelopen jaren met plezier aan dit onderzoek gewerkt. Graag wil ik iedereen hiervoor heel erg hartelijk bedanken. Uit de grond van mijn hart kan ik schrijven dat zonder hen dit proefschrift nooit tot stand gekomen zou zijn. Heel erg veel dank!

Allereest dank aan alle kinderen, hun ouders en de kinderfysiotherapeuten voor hun bijdrage aan de onderzoeken. Zonder jullie zou er geen data zijn geweest en had het valideren van de OMQ-schaal nooit plaats kunnen vinden.
Daarnaast dank voor alle studenten en docenten fysiotherapie voor het deelnemen aan interviews en focusgroepen: jullie meningen, inzichten en ideeën hebben me erg geïnspireerd.

Mijn speciale dank gaat uit naar mijn promotor Ria Nijhuis-van der Sanden, co-promotoren Anjo Janssen en Bert de Swart, en tweede promotor Philip van der Wees.

Beste Ria, wat heb ik veel van jou geleerd! Je sterke analytisch vermogen, je snelle denken en je heldere kritische kijk gaven me veel inzichten. Wat fijn om een promotor te hebben die niet alleen veel ervaring heeft in onderzoekstechnieken en methoden maar die ook inhoudelijk als kinderfysiotherapeut een expert is. Dank voor alle publicaties die je doorstuurde, voor de ontelbare tips die je me gaf. Dank daarnaast voor alle nauwgezette feedback die ik ontving op alle verschillende versies van de artikelen die de hoofdstukken vormen van mijn proefschrift. Wat ontzettend fijn dat je telkens zo snel en uitgebreid reageerde zodat ik weer verder kon met mijn schrijfwerk.

Beste Anjo, wat ben ik blij dat jou als co-promotor. Ik durf met honderd procent zekerheid te zeggen dat dit proefschrift zonder jouw inbreng echt nooit tot stand gekomen zou zijn. Ik kon voor al mijn vragen altijd bij jou terecht. Hoe druk je het ook had, hoe hectisch het op dat moment op de afdeling kinderfysiotherapie ook was; of ik je nu vroeg om mee te denken bij de opzet van dataverzameling, om hulp bij de analyses, of bij het schrijven van

een artikel; je nam altijd alle tijd voor me en stond voor me klaar. Mede hierdoor heb ik een ontzettend fijne en leerzame tijd gehad. Heel veel dank daarvoor!

Beste Bert, door jou is het hele traject van promoveren tot ontwikkeling gekomen. Vanuit je rol als lector neurorevalidatie op de HAN heb je me gestimuleerd om na het afronden van mijn masteropleiding verder te gaan. Hierdoor ben ik als kinderfysiotherapeut en onderzoeker op het Radboudumc gaan werken naast mijn werk als docent op de HAN. De interesse in wetenschappelijk onderzoek en de ambitie om te promoveren werd daarmee niet alleen gevoed, er werd ook een mogelijkheid geschapen om echt aan de slag te gaan. Dank voor dit duwtje in m'n rug, voor alle hulp bij het zoeken naar oplossingen voor de (praktische) problemen die ik onderweg tegenkwam en natuurlijk voor je eeuwige optimisme.

Beste Philip, je bent vanaf het moment dat je als hoogleraar werd aangesteld als tweede promotor betrokken bij mijn promotietraject. Op dat moment was ik met de laatste loodjes bezig: het schrijven van het laatste artikel en het opzetten van dit proefschrift. Tijdens het schrijven hiervan zorgde je ervoor dat ik scherp bleef door je frisse blik en je kritische vragen. Door je opbouwende feedback heb je me steeds weer gemotiveerd om er nog eens goed voor te gaan zitten. Dank hiervoor!

Onmisbare was de ondersteuning bij de ingewikkelde en voor mij soms lastige statistische berekeningen. Marianne Jonker, Reinier Akkermans en Rogier Donders dank voor jullie vakkundige begeleiding en het geduld dat jullie opbrachten bij alle vragen die ik jullie stelde. Hierbij wil ik ook mijn co-auteurs bedanken voor hun bijdrage aan het tot stand komen van mijn proefschrift.

Graag wil ik de leden van de manuscript commissie bestaande uit Prof. Dr. B. Steenbergen, Prof. Dr. M. Willemsen en Prof. Dr. M. Jongmans bedanken voor het beoordelen van dit manuscript.

Veel dank ben ik verschuldigd aan het NWO. Doordat er een promotiebeurs aan mij werd toegekend kreeg ik tijd en ruimte om niet alleen mijzelf verder te ontwikkelen maar ook om bijdrage te leveren aan de ontwikkeling van het vak (kinder)fysiotherapie.
Daarnaast is het geweldige dat ik de kans gekregen van mijn leidinggevenden op de HAN om een NWO-promotiebeurs aan te vragen. Ik dank Theo Joosten, Menno Pistorius, Saskia van der Lyke en natuurlijk mijn direct leidinggevende Herman Berndt, voor het in mij gestelde vertrouwen op een goede afloop van het promotietraject.

Graag wil ik mijn collega's van de HAN bedanken. Als eerste Ton en Marjo. Dank voor jullie hulp en bijdrage aan het analyseren van de kwalitatieve data én voor de hulp om mijn gedachtenspinsels op leesbare manier op papier te krijgen. Daarnaast veel dank aan Wim,

Rob, Elvira, Hans, Niki, Els, Wiebke, Eefje, Esther, Annemarie, Margot en Aukje. Ook al was ik tijdens mijn promotietraject wat minder op de HAN, ik kon altijd bij jullie terecht met om vragen te stellen, om samen te sparren of gewoon om even te kletsen en een kop thee te drinken. Verder wil ik ook mijn andere collega's van de opleiding fysiotherapie en de collega's van zowel het lectoraat als de master neurorevalidatie bedanken voor de gezellige tijd en de prettige samenwerking. Ik vind het erg fijn dat ik vanaf dit jaar (weer) een grotere bijdrage ben gaan leveren aan het werk binnen de teams.

Natuurlijk wil ik ook mijn (oud)collega's van het Radboudumc van de afdeling kinderfysiotherapie en het secretariaat heel erg bedanken. Marlou, Perijn, Anke, Ineke, Maaike, Leo, Merel, Maaike, Daniëlle, Marieke en Christine. Ik heb me altijd erg welkom gevoeld in jullie team, dank voor de samenwerking, het meedenken en de interesse naar mijn promotietraject. Ook al ben ik de laatste tijd wat minder zichtbaar geweest op de afdeling, ik heb onze samenwerking zeer gewaardeerd!
Speciale dank voor de ondersteuning vanuit het secretariaat van IQ Health Care door Annick; wat fijn dat ik bij jou altijd terecht kon met mijn vragen.

Heel veel dank ben ik verschuldigd aan Lard, Maxime, Rick, Grietje, Manon, Mariska, Annuska, Erin, Folkert, Moniek, Eefje, Saskia, Claudia, Anneke en Pauline. Tijdens de afstudeeropdrachten voor jullie bachelor- of masteropleiding zijn er focusgroepen gehouden en interviews afgenomen, zijn er kinderen motorisch onderzocht, video's opgenomen en onderzoeken uitgezet. Wat fijn dat jullie je allemaal zo hebben ingezet. Ik ben jullie erg dankbaar voor alle data die verzameld is!

Naast mijn collega's wil ik natuurlijk ook mijn vrienden bedanken. Dorith en Kirsten wat fijn dat jullie mijn paranimfen zijn vandaag! Jullie betrokkenheid is hartverwarmend. Ik wil jullie graag bedanken voor jullie steun en aanwezigheid tijdens mijn promotie.
Inger, Marlies, Dion, Raph, Torbjørn, Inge, Patricia, Paul en Hildegarde, ik wil jullie niet alleen bedanken voor de fijne vriendschap die we hebben, maar ook voor jullie luisterend oor tijdens (hard)loop rondjes, etentjes, bij een borrel, tijdens een gezamenlijk uitstapje of vakantie. Ook al zag ik niet ieder van jullie zo vaak als ik zou willen, ik heb onze gesprekken zeer gewaardeerd.
Verder aan alle andere vrienden, vriendinnen en familieleden: de gezellige etentjes en avonden, de uitjes en het samen sporten, deze gaven me een welkome afleiding tijdens het onderzoek, dank jullie wel!

Papa en mama, ik wil jullie bedanken voor jullie onvoorwaardelijke steun. Ik heb van jullie geleerd om het beste uit mezelf te halen, om door te zetten en niet te snel op te geven. Maar ook om te van het leven te genieten en te doen wat ik leuk vind. Daar heb ik heel veel aan gehad; hierdoor is het me gelukt om de balans te houden en dit traject tot een goed

einde te brengen. Trees en Ben, ook jullie wil ik bedanken voor alle steun, jullie gezelligheid en alle hulp. Jullie zijn fijne schoonouders.
Dank ook aan mijn zus Tanja en mijn zwager Koen, door de jaren heen hebben we heel wat etentjes gehad, musea bezocht en (reis)plannen besproken. Jullie gezelligheid heeft me gesteund op momenten dat het veel en druk was tijdens dit traject. Daarnaast wil ik ook graag mijn zwager Daan en schoonzus Nicole bedanken voor alle interesse die jullie in mij hebben getoond.

Als laatste een woord van dank voor de vier allerbelangrijkste personen in mijn leven: Joost, Mees, Siem en Veerle.
Lieve Mees, lieve Siem en lieve Veerle, jullie betekenen zo ontzettend veel voor me. Alle drie hebben jullie een fijne, unieke persoonlijkheid. Jullie interesses lopen soms uiteen; toch zijn jullie erg aan elkaar gehecht en hebben jullie een oprechte belangstelling voor de ander. Samen hebben we altijd genoeg om over te praten, te discussiëren, te lachen of om samen te doen. Lieve Joost, je hebt me altijd gesteund en was altijd vol vertrouwen dat ik dit traject tot een goed eind zou brengen. Het was niet altijd makkelijk om naast ons werk voldoende tijd voor elkaar te vinden. Gelukkig kunnen we beide ook de kleine dingen in het leven waarderen. Samen de natuur in, een museum bezoeken of gewoon lekker een film kijken: het maakt niet zoveel uit wat we doen, als het maar samen is en we van elkaars gezelschap kunnen genieten. Heel fijn vind ik dat we ondanks alle drukte in de afgelopen jaren toch steeds de tijd genomen om met z'n vijven mooie reizen te maken. Met ons gezin hebben we veel gezien, ondernomen, maar vooral genoten. Ik hoop dan ook dat er nog veel van dit soort momenten zullen volgen.

Institute for Health Sciences
Radboudumc

NWO