OCCUPATION-BASED EVALUATION AND INTERVENTION

Validity of the Assessment of Motor and Process Skills When Used with Persons with Mental Retardation

by

Anders Kottorp

Umeå 2003
To Pär Gyllander and
to Marianne Rogalin, in memoriam
CONTENTS

ABSTRACT

SVENSK SAMMANFATTNING

ABBREVIATIONS

ORIGINAL PAPERS

INTRODUCTION 1

Occupation and Occupational Performance 3

Persons with Mental Retardation 6

Mental Retardation and Related Conditions 8

Mental Retardation and Adaptive Behavior 8

Client-centered Practice 10

Occupational Therapy, ADL, and Persons with Mental Retardation 11

Mental Retardation and ADL Evaluation 13

The Assessment of Motor and Process Skills 15

Validity 20

Classical measurement statistics limitations 21

Rasch Measurement Models 21

Rasch Measurement Statistics 23

Person ability ($B_i$) and item difficulty ($D_i$) 23

Standard error ($SE$) 24

Mean square ($MnSq$) and standardized ($z$)
goodness-of-fit statistics 24

Rasch-based Evidence of Validity 25

Rasch Measurement Model Limitations 25

RATIONALE 28

AIMS 30
METHODS 31

Participants 31
Study I 31
Study II and III 32
Study IV 33

Assessments, Procedures, and Interventions 34
Assessment of Motor and Process Skills 34
Assessment of level of mental retardation 34
Assessment of Awareness of Disability 35
Occupational therapy intervention program 36

Data Analyses and Designs 37
Rasch measurement statistics 37
Statistical analyses 38
Specific study designs 39

Ethical Considerations 39

RESULTS 40

Evidence of Validity of the AMPS Ability Measures Based on Relationships to Levels of Mental Retardation 40

Evidence of Validity Based on Response Processes and Internal Structure 43
Fit among tasks and items 43
Fit among participants 43

Further Evidence of Validity Based on Internal Structure 45

Group Profiles in ADL Performance Skills 46

The Usefulness of the AMPS Ability Measures for Detecting Change 47

DISCUSSION 48

Methodological Considerations 48
Sample selection, representation, and size 48
Choice of specific Rasch statistics and criteria 51
Single case research design limitations 52
ABSTRACT

The ability to perform everyday life occupations is a critical component in both evaluation and intervention for persons with mental retardation (MR). While the ability to perform personal and instrumental activities of daily living (ADL) has always been important for occupational therapy (OT) practice, there is an absence in OT literature and research with a focus on ADL and persons with MR. The overall aim of this thesis was to evaluate the validity of the Assessment of Motor and Process Skills (AMPS) for evaluation and intervention of ADL ability for persons with MR.

In order to evaluate the evidence of validity of the AMPS ability measures based on relation to level of MR, two groups of participants with MR were evaluated with the AMPS (n=22; n=39). The results indicated expected moderate relationships between ADL motor and ADL process ability measures and level of MR, despite different methods used for evaluating level of MR. The results also indicated that the results of the AMPS evaluation could be used to directly describe and measure the consequences in performance of ADL tasks for persons with different levels of MR.

The evidence of validity of the AMPS was further examined in a study including participants with different types of developmental disabilities (e.g., MR, cerebral palsy, spina bifida) (n=1724). An application of many-faceted Rasch analysis was used to examine goodness-of-fit of the responses for the tasks, skill items, and participants included in the study. All tasks and all items except one demonstrated acceptable goodness-of-fit to the model on the ADL motor and ADL process scales. An expected proportion of participants demonstrated acceptable goodness-of-fit on the ADL motor scale. On the ADL process scale, a slightly lower proportion of participants than expected demonstrated acceptable goodness-of-fit. The results indicated further that persons with more severe levels of MR and persons with more limited ADL process abilities demonstrated different response patterns across tasks and possibly items.

The evidence of validity of the internal structure of the AMPS scales was also evaluated between persons with mild and moderate MR (n=178; n=170). Group specific ADL motor and ADL process skill item hierarchies were generated using many-faceted Rasch analyses and compared. The hierarchies of ADL motor and ADL process skill items remained stable across groups, indicating evidence
of validity of the AMPS scales when used to evaluate persons with MR. The results also indicated that although participants with moderate MR demonstrated overall lower mean ADL motor and ADL process ability, they did perform some specific ADL motor and ADL process skills at a similar level as persons with mild MR.

Finally, the utility of the AMPS ability measures for detecting change were examined in an intervention study including three female participants with moderate MR. The study was based on a single case design and evaluated the effectiveness of a structured occupational therapy intervention program. Improvements were found for the participants in relation to the implementation of the program, but the pattern of changes were different between the participants and across the dependent variables. ADL process ability was the only variable that improved across all participants. The results supported the ADL process abilities as sensitive measures for detecting changes in ADL ability of persons with MR.

In conclusion, the results of these studies contribute to the evidence of validity of the AMPS ability measures and scales, specifically in relation to the evaluation of persons with MR. The finding that an OT program resulted in improved ADL process ability also suggest that the results of the AMPS can be used to plan as well as evaluate outcomes of OT practice. Further research is also suggested in order to improve validity evidence and utility of the AMPS when used with persons with MR.

Keywords: Mental retardation, intellectual disability, developmental disabilities, occupational therapy, activities of daily living, ADL assessment, performance skills, occupational performance, Many-faceted Rasch measurement, single case design.
SVENSK SAMMANFATTNING


Syftet med Studie I var att undersöka den samtidiga validiteten mellan AMPS förmågemått i relation till graden av utvecklingsstörning. Två grupper av personer med olika grad av utvecklingsstörning bedömdes med AMPS (n=22; n=39). Resultaten indikerade moderata till starka samband mellan motorisk ADL-förmåga och process-ADL-förmåga i relation till grad av utvecklingsstörning, oavsett vilken metod som användes för att fastställa graden av utvecklingsstörning i de olika grupperna. Resultaten indikerade också att AMPS-bedömningen på ett mer detaljerat sätt kunde beskriva och mäta konsekvenserna i utförandet av vardagsaktiviteter för personer med olika grad av utvecklingsstörning.

I studie II undersöckes vidare validiteten av AMPS när bedömningen användes med personer med olika typer av medfödda funktionshinder (t. ex. utvecklingsstörning, cerebral pares, ryggmärgsbråck) (n=1724). En tillämpning av fler-facetterad Rasch-analys användes för att undersöka s.k. goodness-of-fit från responserna av de i AMPS ingående ADL-uppgifterna, motoriska ADL- och process-ADL-färdigheterna samt de ingående personerna i studien. Resultaten visade att alla ADL-uppgifter och alla ADL-färdigheter utom en uppräknade acceptabla goodness-of-fit-värden i AMPS motoriska ADL-skala och process-ADL-skala. En förväntad andel personer


Resultaten av studierna bidrar sammantaget till att styrka validiteten av AMPS när bedömningen används med personer med utvecklingsstörning. Resultaten stödjer också att AMPS kan användas för att såväl bedöma som planera interventioner samt utvärdera resultaten av arbetsterapi-insatser för dessa personer.
ABBREVIATIONS

AAD        Assessment of Awareness of Disability
AAMR       American Association on Mental Retardation
ADL        Activities of daily living
AERA       American Educational Research Association
AMPS       Assessment of Motor and Process Skills
AOTA       American Occupational Therapy Association
APA        American Psychological Association
Bn         Person ability measure
CAOT       Canadian Association of Occupational Therapists
d         Effect size
Di         Item difficulty calibration
DSM-IV     Diagnostic and Statistical Manual of Mental Disorders, 4th edition
IADL       Instrumental, or domestic, activities of daily living
ICD-10     International Statistical Classification of Diseases and Related Health Problems, 10th edition
ICF        International Classification of Functioning, Disability, and Health
ID         Intellectual disability
IQ         Intelligence quotient
IRT        Item response theory
M          Mean value
MFR model  Many-faceted Rasch model
MnSq       Mean square
MR         Mental retardation
NCME       National Council on Measurement in Education
OT         Occupational therapy
OTIPM      Occupational Therapy Intervention Process Model
PADL       Personal activities of daily living
PEDI       Pediatric Evaluation of Disability Inventory
SD         Standard deviation
SE         Standard error
SIB-R      Scales of Independent Behavior – Revised
WHO        World Health Organization
The present thesis is based on the following papers, which will be referred to by their Roman numerals:


The original articles have been reprinted with the kind permission of the publishers.
INTRODUCTION

I think I am good at almost everything. It is a few things that can sometimes get messy.

[Jag tycker att jag är bra på nästan allt. Det är bara vissa saker som ibland kan köra ihop sig.]

This quote was from a man in his mid-thirties, living in a small community in the county of Västerbotten, when he was interviewed about his experiences of everyday life activities. He had mild mental retardation, lived in a two-room apartment in the community, and received minimal assistance from caregivers. He was participating in a study with a focus on how persons with mild mental retardation, living independently in the community, perceived and described their performance of everyday life tasks.

As members of multidisciplinary habilitation teams working with clients with mental retardation, it is critical that all professions contribute their unique expertise and perspective on functioning in evaluation and intervention. Each profession also has the responsibility to use evaluation tools and administration procedures that have shown evidence of validity for the specific target group, and also contribute with information useful for the intervention planning process. Working as a clinical occupational therapist in the late 1980s and 1990s, I found it difficult to plan, implement, and evaluate the outcomes of my interventions in order to improve the functioning in everyday life activities for persons with mental retardation, due to the absence of valid and reliable occupational therapy evaluations. The evaluations available were not specifically focused on the actual doing of everyday life tasks, but rather were designed to evaluate individual and/or environmental characteristics.

Therefore, the overall focus of my thesis is evaluation of occupational therapy methods to describe, assess, and intervene in order to improve the doing of everyday life tasks for persons with mental retardation. It is critical that we, as clinical occupational therapists, use valid assessments of everyday life tasks, or personal and domestic (instrumental) activities of daily living (ADL), that also support the continuing intervention planning process. We must also evaluate and document if and how the implemented interventions are effective or
Introduction

not in improving the performance of those tasks. My personal frames of reference and professional experiences led to the following foci of this thesis:

*Persons with mental retardation* — a heterogeneous group of persons that share some characteristics, related to an incomplete development of the mind, that impact on the overall level of intelligence and adaptive behavior. These people will be the primary target group studied.

*Client-centered practice* — where the professional provider of services starts the intervention process from the client’s perspective and personal experience, and where there is a continuous collaboration between the professional and the client throughout the evaluation and intervention process. The evaluation assessments and intervention methods studied will all have a client-centered focus.

*The doing of everyday life occupations* — the performance of actions, which results in the completion of a specified everyday life task, that is meaningful and purposeful for the individual. The skillfulness of the actions supports the quality of the doing, resulting in an efficient, safe, effortless, and independent performance of ADL. The quality of the doing of everyday life tasks will be evaluated using the Assessment of Motor and Process Skills (AMPS).

*Methods to describe and measure performance* — the challenge of generating measures from the performance of everyday tasks. Traditional statistics have been shown to have limited use in generating valid and reliable measures from ordinal data. In this thesis, an alternative approach to traditional psychometric statistics will therefore be used, the family of Rasch measurement models.

*Occupational therapy* — an academic discipline and a profession, offering services to persons with mental retardation; the view of the themes described above will be through the lens of occupational therapy theory and practice.
Introduction

Occupation and Occupational Performance

Occupational therapy has used the word *occupation* to frame and describe the area of concern for its theory development, research, and practice. Although different words may be used internationally to describe our profession (e.g., arbetsterapeut, ergoterapeut), the original word chosen was *occupational* therapy (A. G. Fisher, 1995, 2001a). The word *occupation* has its origin in the Latin words *occupare/occupatio*, and means seizing, taking possession, or occupying space and time (*The Oxford English Dictionary*, 1989). The word *occupation* also means the activities or actions a person performs in his or her profession, as part of his or her daily life, or for pleasure\(^1\) (Ahlberg, Lundqvist, & Sörbom, 1982). Since the word relates to the actual *doing* of something (the verb, e.g., *playing* with the children, *working* in the office), the focus for occupational therapy should therefore emphasize the engagement of performances that comprise the process of doing (A. G. Fisher, 1994b).

Occupations should also be meaningful and purposeful for the person who engages in them. The meaning of the occupation relates to the source of motivation for doing for the specific person. The purpose relates to the goal or the aim of the occupation, and supports the organization of the doing (A. G. Fisher, 1994b, 1998; Trombly, 1995a). Occupations are also performed in a context, that includes personal, physical, social, temporal, and cultural aspects, that impact and form the doing of occupations (American Occupational Therapy Association [AOTA], 2002; A. G. Fisher, 1998, 2002; Kielhofner, 2002).

Even though the word occupation, in itself, refers to the performance of activities, the concept *occupational performance* is also commonly used in occupational therapy theories and research to highlight the actual doing of occupations. A selection of proposed definitions from occupational therapy literature follows:

> Occupational performance [refers] to the day-to-day engagement in occupations that organize our lives and meet our needs to maintain

\(^1\) As an interesting notation for the Swedish reader, “occupatio” was also translated to the Swedish words “sysselsättning/sysselsatt”, but not the word “aktivitet”.

---

3
ourselves, to be productive, and to derive enjoyment and satisfaction within our environments.

(Christiansen & Baum, 1991, p.27)

We view occupational performance as a meaningful sequence of actions in which the person enacts and completes a specified task that is relevant to his or her culture and daily life roles.

(Fisher, 2001a, p. 24)

Occupational performance refers to doing an occupational form (Nelson, 1988). For example, when persons do such tasks as shining shoes, baking a cake, mowing the lawn, and painting a room, they are performing those occupational forms.

(Kielhofner, 2002, p.116)

Occupational performance refers to the ability to choose, organize, and satisfactorily perform meaningful occupations that are culturally defined and age appropriate for looking after one's self, enjoying life, and contributing to the social and economic fabric of a community.

(Townsend, 1997, p.30)

Occupational performance, or the ability to carry out activities during daily life, depends on the individual's culture and gender, and the roles that he or she wishes to undertake, and the environment in which he or she lives. Thus, occupational performance is an individual concept.

(Trombly, 1995b, p.43)

There are common themes in these proposed definitions, although there are some differences. When considered together, these definitions suggest that occupational performance can be defined as the doing of occupations; the ongoing performance of actions that results in the accomplishment of a task within a given role. The emphasis, however, can be viewed from different aspects, or levels, of occupational performance:

Roles define the nature of occupational performance at various points in one's lifespan. This level of occupational performance refers to what is expected of performance in order to fulfil different social roles (Christiansen & Baum, 1991; Townsend, 1997; Trombly, 1995b). Kielhofner (2002) refers to this dimension of doing as occupational participation; to engage in tasks that are part of one's
sociocultural context and desired and/or necessary for one’s well-being. Examples of roles can be related to work and employment (e.g., teacher), domestic life (e.g., household maintainer), and relationships (e.g., parent). This level of occupational performance is captured in the Participation dimension in the International Classification of Functioning, Disability, and Health (ICF), from the World Health Organization [WHO] (WHO, 2001).

Activities are often viewed as the basic unit of occupational performance. Activities consist of specific goal-oriented behavior directed toward the performance of a task (Christiansen & Baum, 1991; Kielhofner, 2002). In relation to the examples described above, tasks a person might perform in his or her role as household maintainer may be preparing breakfast, folding laundry, and vacuuming the living room. This level of occupational performance is captured as tasks in the Activity dimension, as defined in the ICF (WHO, 2001), and as occupational performance by Kielhofner (2002).

Finally, occupational performance can also be viewed from the perspective of the discrete purposeful actions that, when sequentially performed, result in the doing of a specified task. These smallest units of occupational performance are referred to as occupational skills (A. G. Fisher, 2001a; Kielhofner, 2002). Skills are practiced, goal-directed actions that show deftness, dexterity, and confidence in performance (Connolly & Dalgleish, 1989). This level of occupational performance is captured as actions in the Activity dimension in ICF (WHO, 2001).

All of these levels, and specifically the level of skills, are to be differentiated from underlying body functions or personal factors in ICF (WHO, 2001). In contrast to body functions, which refer to the underlying capacities of the individual (e.g., mental functions, movement related functions), skill refers to the concrete actions that are performed in the midst of undertaking a task performance, where the person’s characteristics (including performance capacity) interact with the environment (AOTA, 2002; A. G. Fisher, 2001a; Kielhofner, 2002). A skilled performance is always jointly determined by the person, the task, and the precise environment in which the actions take place (Connolly & Dalgleish, 1989).
Introduction

The different levels of occupational performance are critical to differentiate between. They can impact on the design and choice of data gathering processes used, the interpretation of the information gathered, and the focus of interventions in occupational therapy practice.

Persons with Mental Retardation

Persons with mental retardation are not a homogeneous diagnostic group, but a heterogenous population in relation to aetiology, capacities, and abilities (Accardo & Capute, 1998; Greenspan, 1999; Hogg & Raynes, 1987; Kluver, Clark, & Hoffman, 1998). Moreover, they are different persons related to their earlier life experiences, current environmental surroundings, cultural backgrounds, and personal interests and values. All of these aspects impact on their choices, organization, and actual performance of occupations.

The definition of mental retardation in the International Statistical Classification of Diseases and Related Health Problems (ICD-10) states that mental retardation is a condition of arrested or incomplete development of the mind, especially characterized by impairments of skills manifested during the developmental period, that will contribute to the overall level of intelligence (i.e., cognitive, language, motor, and social abilities) (WHO, 1992). There are two dimensions evaluated in the diagnosis of mental retardation in the ICD-10: limited intelligence (IQ<70) and limitations in adaptive behavior. The diagnosis is usually estimated by the use of standardized intelligence tests. As a supplement to these, scales assessing adaptive behavior in different environments are also used, but are not specified further. These two dimensions together contribute to the approximate indication of the degree, or severity, of mental retardation. The degree of mental retardation is defined as mild, moderate, severe, and profound. IQ scores define the degree of mental retardation, but no criteria for defining and clarifying limitations in adaptive behavior in relation to those are specified.

The Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV) from the American Psychiatric Association (1994) provides a similar definition and degree of mental retardation as proposed in ICD-10 (WHO, 1992). In contrast to the suggested IQ cutoff of two
standard deviations below the mean for intelligence as a criterion acceptable for the diagnosis of mental retardation, there is no mention of cutoff scores or criteria for adaptive behavior limitations, even though the authors suggest that one or more instruments be used to assess different domains of adaptive behavior.

The American Association of Mental Retardation [AAMR] (1992) has proposed a somewhat different definition of mental retardation as compared to the ICD-10 and DSM-IV definitions (American Psychiatric Association, 1994; WHO, 1992). Although the AAMR definition shares the same concepts of limited intelligence and adaptive behavior, the definition differs from the other two in two distinct ways. First, the AAMR definition operationally defines limitations in adaptive behavior as limitations in 2 or more out of 10 adaptive skill areas. Second, the AAMR definition does not categorize persons with mental retardation into levels of severity. Instead, they are categorized based on the amount of support they need. The operational definition of adaptive behavior and the changes of levels of severity into levels of support have been criticized and discussed from conceptual, psychometric, and clinical perspectives (Llewellyn, 1994; MacMillan, Gresham, & Siperstein, 1993; Reiss, 1994; Simeonsson, Granlund, & Bjorck-Akesson, in press).

While the term mental retardation is consistently used in the ICD-10 and DSM-IV, and by the AAMR (c.f., AAMR, 1992; American Psychiatric Association, 1994; WHO, 1992), the term used to define this population varies between countries. In Australia, the term intellectual disability is used. In Europe, the terms intellectual disability, mental handicap, learning disability, and learning difficulty, are commonly used. The choice of using the concept mental retardation in this thesis was primarily based on the fact that the term used in the ICD-10 and DSM-IV was mental retardation. As the reader will note, however, the equivalent term intellectual disability was used in Studies II and III, since they were submitted to British and Australian journals, respectively.

Finally, the term developmental disability is used to refer to people with a variety of congenital or early childhood onset mental and/or physical diagnoses and impairments that result in delayed development, and that include, but are not limited to, people with the diagnosis of mental retardation. The term developmental disability was also used in Study II in this thesis, where a sample of persons with different
disabilities associated with developmental delays, including persons with mental retardation, was evaluated.

**Mental Retardation and Related Conditions**

The diagnosis of mental retardation is also frequently associated with other specific conditions (e.g., autism, cerebral palsy, visual impairment, seizure disorders). Sometimes, however, these secondary conditions are not diagnosed as they are overshadowed by the mental retardation. Calculations based on the national Health Interview Survey (NHIS) data from 1994/95 (Larson et al., 2001) revealed that 38% of the noninstitutionalized persons with mental retardation also had one or more of the identified conditions autism, spina bifida or cerebral palsy. Persons with mental retardation are also, to a higher extent, affected by other health problems than people without mental retardation. High prevalences of visual and hearing impairments has been reported (van Schrojenstein Lantman de Valk et al., 1997). Similarly, psychiatric disorders are more common among persons with mental retardation (Day & Jancar, 1994). Research has also indicated that persons with Down syndrome are more likely to develop dementia of the Alzheimer type, but not persons with mental retardation of other forms (Zigman, Schupf, Sersen, & Silverman, 1995). The issue of diagnostic overshadowing in persons with mental retardation has been confirmed in research in relation to mental health problems (White et al., 1995). It is unclear if other disorders (e.g., motor, social, cognitive) also may be less likely to be detected in persons with mental retardation due to diagnostic overshadowing. It is also important to consider the fact that since persons with mental retardation have difficulties in adapting to problems encountered in everyday life in general, the difficulties will be even more pronounced when attributable disorders are interacting and impacting on the doing of everyday life tasks.

**Mental Retardation and Adaptive Behavior**

The diagnosis of mental retardation always includes a component of adaptive behavior, despite the different definitions described earlier. The concept of adaptive behavior has been a formal part of the diagnosis of mental retardation since the late 1950s (Reschly, Myers, & Hartel, 2002). The concept of adaptive behavior is
multidimensional in nature, confirmed both by definitions but also by factor analytic studies (Thompson, McGrew, & Bruininks, 1999). Assessments of adaptive behavior always include items related to self-care, home-living, ability to work, and function in the community. That is, there is always one distinct domain found in adaptive behavior scales: independent living skills or daily living skills (Widaman, Borthwick-Duffy, & Little, 1991; Widaman & McGrew, 1996). In the early precursors to today's adaptive behavior assessments, the performance of personal and instrumental ADL was the primary measure of adaptive behavior (Reschly et al., 2002).

According to Reschly et al., (2002), the development of different scales of adaptive behavior was initiated during the 1960s. During the following decades, these scales were updated and refined psychometrically. There are now different methods used for evaluation of adaptive behavior. Most adaptive behavior scales use structured or semi-structured interviews as data gathering methods, sometimes complemented by checklists filled in by caregivers, parents, or teachers. Since adaptive behavior is defined as applying to a broad range of different life situations, assessments based on direct observation are usually considered as time-consuming and impractical. Today there are at least 200 published scales of adaptive behavior that can be used for diagnosis setting, individualized intervention plan development, program evaluation, and administration. Although several instruments have tried to include both the breadth required for appropriate diagnosing, and the detailed depth required for individual use, most fall short of accomplishing both.

When using evaluations for individual intervention planning with persons with mental retardation, there are other important aspects to consider in the evaluation process: How can the methods used for data gathering incorporate aspects about what is meaningful and important for the specific person in his or her life? How can professionals (e.g., occupational therapists) working with persons with mental retardation incorporate evaluations into client-centered practice?
Client-centered Practice

The term client-centered practice was introduced by Carl Rogers in the 1930s, based on his clinical work with maladjusted children in the United States of America [USA] (C. R. Rogers, 1939). The concept was further formulated and described in his later work (C. R. Rogers, 1951, 1957, 1961). The ideas behind client-centered practice was based on Rogers’ experience in psychotherapy with which clinical conditions were effective in initiating a constructive change in the personality and behavior of the person seeking help (C. R. Rogers, 1961). The provision of these conditions became the basis for an approach to psychotherapy that was described as client-centered or non-directive therapy. Rogers also suggested that these conditions could serve as criteria to other programs that aim towards constructive changes in the personality and/or behavior of the individual (C. R. Rogers, 1957).

More specifically, client-centered therapy is not a specific method, but more an approach, based on a close relationship between the therapist and the client. Rogers chose the term *client* because it implied that the person is someone who is active and voluntarily seeks support or help to solve a problem, and also takes responsibility for the situation (C. R. Rogers, 1951). Client-centered therapy highlights the necessity of a continuous relationship between the client and the therapist, where the therapist is a congruent, genuine, and integrated person. If the therapist is presenting a façade, knowingly or unknowingly, he or she is not being freely and deeply him- or herself, which is contrary to client-centered therapy (C. R. Rogers, 1957). The therapist should also experience an unconditional positive acceptance of each aspect of the client. This means a caring for the client as an individual and separate person, with permission to have his or her own feelings, experiences, values, and priorities. The therapist should experience an accurate, empathic understanding of the client’s awareness of his or her own experience. Finally, the client should also perceive, to some extent, the acceptance and empathy the therapist experiences for him or her. These defined conditions must be present in order for a constructive change to occur (C. R. Rogers, 1951, 1957). If they are, the therapist can facilitate problem-solving by stimulating the individual’s desire and ability to understand problems unique to his or her life situation, and propose appropriate solutions. However, the magnitude of the change is not related to the experience and professional knowledge of
Introduction

the therapist. Rather, the greater degree to which the conditions of the relationship between the client and therapist exist, the more marked the change will become (C. R. Rogers, 1957).

Occupational therapy has incorporated the ideas of client-centered practice into different frameworks, codes of ethics, and theory development (Law & Mills, 1998): Occupational therapists must show respect for the client and his or her family, and the choices and priorities they define as most important for their lives. The immediate development of the relationship between the therapist and the client is also highlighted, where the client is an active participant throughout the working process and is ultimately responsible for the decisions made (A. G. Fisher, 1998, 2001a; Kielhofner, 2002; Tickle-Degnen, 2002; Townsend, 1997). Occupational therapists should also emphasize individualization and flexibility in all phases of the occupational therapy process.

The first formulation of client-centered practice within an occupational therapy framework was initiated in the 1980s by the Canadian Association of Occupational Therapists [CAOT] (Townsend, 1997). Today, several occupational therapy models of practice use a client-centered approach to formulate basic assumptions for practice (AOTA, 2002; Baum & Law, 1997; A. G. Fisher, 1998, 2001a, 2002; Kielhofner, 2002; Townsend, 1997). The client-centered approach has also been incorporated in occupational therapy practice regarding decision-making, instrument development, and development of therapeutic rapport (AOTA, 2002; Coster, 1998; Egan, Dubouloz, von Zweck, & Vallerand, 1998; Kielhofner et al., 1997; Law et al., 1998; Tickle-Degnen, 2002).

Occupational Therapy, ADL, and Persons with Mental Retardation

Occupational therapists define ADL and community living as one important area of concern when working with persons with mental retardation. The goal for occupational therapy is enabling persons with mental retardation towards higher degree of independence in everyday life and community living, and empowering these persons to a higher level of control over a single actions or activities, daily routines, or the environment (Cumming, Jones, Cumming, &
Llewellyn, 1992; Giese, 2000; Herge & Campbell, 2000; Kluver et al., 1998; Neistadt, 1986; Tannous, Lehmann-Monck, Magoffin, Jackson, & Llewellyn, 1999). The importance of performing an ADL task independently also contributes to both the external and internal view of the person as an autonomic and responsible person (Jansson, 2002; Kjellberg, 2002).

Evaluations of ADL has always been of fundamental importance in occupational therapy practice (Eakin, 1989; Law & Letts, 1989). But although one important area for occupational therapy practice today is to support persons with mental retardation to be able to function in their everyday life, the areas of evaluation and intervention of occupational performance related to ADL do not seem to be documented in literature and research. A review of all articles published in occupational therapy journals from 1985 until today’s date that included the key words mental retardation, intellectual disability, learning disability, or developmental disability, were reviewed. The sources for the review process were AMED, MEDLINE, and CINAHL. The journals selected were American Journal of Occupational Therapy, Australian Occupational Therapy Journal, British Journal of Occupational Therapy, Canadian Journal of Occupational Therapy, Occupational Therapy International, Occupational Therapy Journal of Research, Occupational Therapy in Mental Health, OT Practice, Physical and Occupational Therapy in Pediatrics, and Scandinavian Journal of Occupational Therapy. A total number of 158 papers were found, which represented 2.7% of the total number of published articles in these journals during the period. Although differences in proportions varied between specific journals, the total number of published papers related to mental retardation or developmental disabilities in occupational therapy journals seemed to have decreased during the 1990s, compared to earlier reviews (Llewellyn, 1990; Tannous et al., 1999). No published paper was found specifically evaluating ADL ability among adults with mental retardation. A number of papers focused on evaluation of occupational therapy interventions (“hands-on treatment”). The outcomes of those studies were primarily focused on body functions (WHO, 2001), and not outcomes related to occupational performance. Only a few studies specifically focused on ADL ability, community living skills, or independent living skills as outcome variables (Neistadt, 1987; Nochajski & Gordon, 1987; Paul, 1997). Moreover, only a limited number of papers focused on the development and evaluation of assessments tools and procedures
Introduction

(Hong, Smith, & Roper, 2000; Neeman, 1986; Tabatabainia, Ziviani, & Maas, 1995; Young & Chesson, 1997). Most of those assessments were designed to assess body functions and personal factors; no assessments regarding ADL ability were specifically described or validated.

Finally, there have been a limited number of descriptions of occupational therapy practice for persons with mental retardation reported in the literature. If mentioned at all, the descriptions were focused primarily on children and adolescents with mental retardation (Baloueff, 1998). An overview of occupational dysfunctions and case studies among persons with mental retardation based on the Model of Human Occupation have also been reported (Bryze & Fisher, 1995; Lyons, Kielhofner, & Kavanagh, 1985).

Mental Retardation and ADL Evaluation

Most ADL evaluations used by occupational therapists working with clients with mental retardation were developed for use by different professions (e.g., Pediatric Evaluation of Disability Inventory [PEDI] (Haley, Coster, Ludlow, Haltiwanger, & Andrellos, 1992), Scales of Independent Behavior-Revised [SIB-R] (Bruininks, Woodcock, Weatherman, & Hill, 1996)). Although these evaluations of ADL are useful for global evaluations within multidisciplinary teams, evaluating needs of supports for persons with mental retardation, or testing specific subsamples (e.g., children), they may have limitations in the more specific range of occupational therapy intervention planning processes for persons with mental retardation.

Different studies have indicated that ADL ability is an important variable in predicting level of dependence in persons with mental retardation (Maaskant et al., 1994, 1995). A limitation of those studies was that the authors only defined and measured ADL ability in relation to personal ADL (PADL). This may also be a limitation of the use of the PEDI, which was designed to assess only PADL tasks. Since ADL scales that are limited to PADL tasks tend to be narrow in range and also less sensitive to detect variations among higher functioning clients (McDowell & Newell, 1996), ADL scales that incorporate both PADL tasks and instrumental ADL (IADL) tasks may have a potential to better identify problems and detect changes in ADL performance over time among clients with various degree of
ADL disability. ADL scales that include both PADL and IADL could also have the potential to better differentiate among higher functioning clients, since IADL tasks generally are more difficult in challenge (A. G. Fisher, 2001a; Grimby et al., 1996; McDowell & Newell, 1996; Spector, Datz, Murphy, & Fulton, 1987). For example, since persons with mild mental retardation may only have minor limitations in adaptive behavior, ADL assessments that include IADL tasks may be more suitable for detecting their problems. Furthermore, persons with mental retardation may demonstrate a different set of problems in ADL than, for instance, persons with physical limitations (Kylén, 1974). It is important, therefore, that the ADL assessments used focus not just on physical aspects of ADL performance in order to have clinical utility in the field of mental retardation.

Most standardized assessments of ADL are also based on self- or proxy-report formats (e.g., SIB-R [Bruininks et al., 1996]) (A. G. Fisher, 2001a). Several studies have, however, found differences in determining functional abilities and other types of behavior among clients with mental retardation in relation to data gathering methods used (Villeponteaux & DeCoux, 1998; Yarbrough & Carr, 2000). There is also an increased recognition that observational assessments of ADL may be preferred in many instances (Guralnik, Branch, Cummings, & Curb, 1989).

Finally, ADL assessments often provide information about what ADL tasks a person can or cannot perform, but do not offer further information about the quality of the specific actions that support or limit overall task performance (A. G. Fisher, 1995, 2001a). Since most assessments used focus on the global outcome of the doing of ADL tasks (e.g., Can the person sweep the floor with a broom and use a dustpan?) rather than the process of doing (e.g., What is the quality of the doing when sweeping the floor?), the information may therefore not be directly applicable in developing specific client-centered goals and interventions. The need for assessments focusing on directly observable competencies in real life situations has also been highlighted in the field of mental retardation (Halpern, 1984; Halpern, Lehmann, Irvin, & Heiry, 1982; Reschly et al., 2002; Schalock et al., 1994). Since occupational therapists are not only interested in the outcomes of the doing, but also the process of doing, global multidisciplinary assessments could therefore then be complemented by evaluations of
Introduction

ADL, made by occupational therapists, that gather information about the quality of the sequentially performed actions over time.

In addition to the considerations described above, it is also critical that the ADL assessments used incorporate aspects about of is meaningful and important for the specific person in his or her everyday life. Since motivational aspects impact on the quality of task performance (Doble, 1988; A. G. Fisher, 2001a; Kielhofner, 2002) it is critical that the evaluations of ADL task performance support client-centered practice, especially since occupational therapy has highlighted the importance of this approach in current practice and theory development (AOTA, 2002; Coster, 1998; Egan et al., 1998; A. G. Fisher, 1998, 2002; Kielhofner, 2002; Law, 1998; Townsend, 1997).

The Assessment of Motor and Process Skills

The Assessment of Motor and Process Skills (AMPS) is an evaluation of ADL performance skill displayed in the individual actions performed as a person carries out ADL tasks that are meaningful and relevant to him or her, and that are chosen by the person (A. G. Fisher, 2001a, 2001b). The AMPS, therefore, is a client-centered assessment of the quality (and outcome) of the doing process.

The ADL motor and ADL process skills in the AMPS represent two universal taxonomies of performance used to carry out ADL tasks. The ADL motor skills are the observable actions the person uses during the performance of ADL tasks as he or she moves oneself or the task objects used. The ADL process skills are the observable actions the person enacts as he or she logically sequences the actions of the ADL task performance over time, selects and uses appropriate tools and materials, and adapts his or her performance when problems are encountered (see Table 1).

More specifically, the doing process involves constructing and completing a task performance by enacting actions, one by one. For example, when one is preparing a bowl of cereal, the person is observed to open a drawer; search for the spoons; reach for, grasp, and lift a spoon up from the drawer; reposition it in the hand; and place it in the bowl of cereal on the table. It is the quality of skill of these individual
### Table 1. AMPS ADL motor and ADL process skills

<table>
<thead>
<tr>
<th>ADL motor skills</th>
<th>ADL process skills</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stabilizes</td>
<td>Flows</td>
</tr>
<tr>
<td>Aligns</td>
<td>Moves</td>
</tr>
<tr>
<td>Positions</td>
<td>Transports</td>
</tr>
<tr>
<td>Walks</td>
<td>Lifts</td>
</tr>
<tr>
<td>Reaches</td>
<td>Calibrates</td>
</tr>
<tr>
<td>Bends</td>
<td>Grips</td>
</tr>
<tr>
<td>Coordinates</td>
<td>Endures</td>
</tr>
<tr>
<td>Manipulates</td>
<td>Paces</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Paces</td>
</tr>
<tr>
<td></td>
<td>Attends</td>
</tr>
<tr>
<td></td>
<td>Chooses</td>
</tr>
<tr>
<td></td>
<td>Uses</td>
</tr>
<tr>
<td></td>
<td>Handles</td>
</tr>
<tr>
<td></td>
<td>Heeds</td>
</tr>
<tr>
<td></td>
<td>Inquires</td>
</tr>
<tr>
<td></td>
<td>Initiates</td>
</tr>
<tr>
<td></td>
<td>Continues</td>
</tr>
<tr>
<td></td>
<td>Sequences</td>
</tr>
</tbody>
</table>

|                 | Terminates         |
|                 | Searches/Locates   |
|                 | Gathers            |
|                 | Organizes          |
|                 | Restores           |
|                 | Navigates          |
|                 | Notices/Responds   |
|                 | Accomodates        |
|                 | Adjusts            |
|                 | Benefits           |

### Table 2. Behavior observed in ADL performance in relation to AMPS skill items

<table>
<thead>
<tr>
<th>Behavior observed</th>
<th>AMPS skill items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Opens a drawer</td>
<td>Grips (ADL motor skill)</td>
</tr>
<tr>
<td></td>
<td>Moves (ADL motor skill)</td>
</tr>
<tr>
<td></td>
<td>Calibrates (ADL motor skill)</td>
</tr>
<tr>
<td>Searches for spoons</td>
<td>Searches/Locates (ADL process skill)</td>
</tr>
<tr>
<td>Reaches for spoons</td>
<td>Reaches (ADL motor skill)</td>
</tr>
<tr>
<td></td>
<td>Bends (ADL motor skill)</td>
</tr>
<tr>
<td>Grasps a spoon</td>
<td>Grips (ADL motor skill)</td>
</tr>
<tr>
<td></td>
<td>Chooses (ADL process skill)</td>
</tr>
<tr>
<td>Lifts the spoon from the drawer</td>
<td>Lifts (ADL motor skill)</td>
</tr>
<tr>
<td>Repositions spoon in hand</td>
<td>Manipulates (ADL motor skill)</td>
</tr>
<tr>
<td></td>
<td>Handles (ADL process skill)</td>
</tr>
<tr>
<td>Places the spoon in bowl</td>
<td>Calibrates (ADL motor skill)</td>
</tr>
<tr>
<td></td>
<td>Gathers (ADL process skill)</td>
</tr>
<tr>
<td></td>
<td>Organizes (ADL process skill)</td>
</tr>
</tbody>
</table>
actions that is scored in the AMPS on one or more of the AMPS skill items (see Table 2). Each item is scored on a 4-point criterion-referenced rating scale. The scoring criteria for each performance skill in the AMPS consider simultaneously the degree of efficiency, safety, ease, and independence observed.

There are 9 PADL and 74 IADL standardized tasks available in the current edition of the AMPS manual (A. G. Fisher, 2001b). Examples of AMPS tasks with various challenges are presented in Table 3.

Table 3. Selected examples of ADL tasks included in the AMPS

<table>
<thead>
<tr>
<th>Very easy tasks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brushing teeth</td>
</tr>
<tr>
<td>Eating a meal</td>
</tr>
<tr>
<td>Folding a basket of laundry</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Easy tasks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Setting a table for one or two persons</td>
</tr>
<tr>
<td>Polishing shoes</td>
</tr>
<tr>
<td>Cleaning windows</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Average tasks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Making a luncheon meat or cheese sandwich</td>
</tr>
<tr>
<td>Raking grass cuttings or leaves</td>
</tr>
<tr>
<td>Vacuuming a room</td>
</tr>
<tr>
<td>Shopping</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Harder than average tasks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Making hot cooked cereal and beverage</td>
</tr>
<tr>
<td>Making a fresh fruit salad</td>
</tr>
<tr>
<td>Cleaning a bathroom</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Much harder than average tasks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Making french toast and beverage</td>
</tr>
<tr>
<td>Making pasta with sauce, green salad and beverage</td>
</tr>
</tbody>
</table>
The design of the AMPS allows the therapist to evaluate persons using a wide range of ADL tasks that are familiar to the person, contextually relevant, and self-chosen. As noted above, the AMPS is, therefore, an appropriate ADL assessment to use in a client-centered occupational therapy evaluation process. In addition, scoring the person on the quality of the individual actions (performance skill items), results in the examiner being able to determine not only if a person has problems in performing a specified ADL task, but also (a) why the person experienced difficulty in performance (i.e., which actions performed were less effective), as well as (b) why the person performance was successful (i.e., which actions performed were effective). The AMPS, therefore, is a suitable tool for defining specific actions (occupational performance skills) in the performance of ADL tasks that are supporting and limiting the person, in order to develop specific goals and intervention strategies well targeted to enhance and support the quality of ADL task performance for persons with mental retardation.

A limitation of the AMPS is that training is required before being able to use the assessment. During the 5-day training course, the course participants learn how to administer and score the AMPS. Following the course, the participants must complete an additional 10 AMPS evaluations within a limited period of time (beginning January, 2002, within a 3 month period after attending the AMPS training course). This calibration process is to ensure that the occupational therapist is scoring the AMPS in a consistent and reliable manner.

Additional concerns expressed related to the use of the AMPS in clinical practice have been reported, including limited choice of tasks due to the standardization limits\(^2\) (McAdam, Thomas, & Chard, 2001). The time factor in administering the AMPS has also been raised as a potential limitation (Hariz, Bergenheim, Hariz, & Lindberg, 1998), although most AMPS evaluations (including interview, setting up the environment, and scoring two AMPS tasks) is performed within a 45 to 60 minutes period (Fisher, 2001a). Other limitations in the AMPS administration and scoring are the need of accessibility to

\(^2\) This critique was primarily from occupational therapists that used earlier versions of the AMPS manuals. In the more recent AMPS manuals, 8 PADL tasks and 19 IADL tasks has been standardized and added (Fisher 1999b, 2001b).
suitable computer equipment, and basic computer skills (McAdam et al., 2001).

In contrast to these limitations, the utility of the AMPS has been extensively evaluated, indicating that the AMPS is a suitable tool for assessment, guiding interventions, and measuring outcomes. The AMPS has been internationally standardized on over 50,000 persons with a variety of orthopaedic, neurological, cognitive, medical, psychiatric, and developmental disorders. The cross-cultural/ethnic validity of the AMPS measures has been demonstrated (Bernspång & Fisher, 1995b; Dickerson & Fisher, 1993; Goldman & Fisher, 1996; Goto, Fisher, & Mayberry, 1996; Magalhães, Fisher, Bernspång, & Linacre, 1996; Stauffer, Fisher, & Duran, 2000). Several studies have also validated the AMPS measures for specific diagnostic groups (Bernspång & Fisher, 1995a; Doble, Fisk, Fisher, Ritvo, & Murray, 1994; Girard, Fisher, Short, & Duran, 1999; Hartman, Fisher, & Duran, 1999; Pan & Fisher, 1994). The AMPS ability measures are, as expected, associated with the severity of impairments among different diagnostic groups (Doble et al., 1994; Robinson & Fisher, 1996), but also unique patterns of skill strengths and limitations (occupational performance skill profiles) has been determined for specific groups (Cooke, Fisher, Mayberry & Oakley, 2000; Oakley, Duran, Fisher & Merritt, 2003). Specific cutoff criteria on the AMPS motor and process scales have been determined in order to predict level of effort, efficiency, safety, and assistance in ADL performance (A. G. Fisher, 2001a). The cutoff criteria have also been shown to discriminate between individuals who are able to function independently in the community and who needs support to function in the community (Bernspång & Fisher, 1995a; A. G. Fisher, 2001a; Hartman et al., 1999). The AMPS ability measures have also been shown to be sensitive in detecting the effectiveness of interventions (Hariz et al., 1998; Kinnman, Andersson, Wetterquist, Kinnman, & Andersson, 2000; Oakley, Khin, Parks, Bauer, & Sunderland, 2002; Oakley & Sunderland, 1997; Tham, Ginsburg, Fisher, & Tegnér, 2001). However, limited research has focused on the evidence of validity when using the AMPS with clients with mental retardation.
Validity

Validity is the most fundamental consideration in test development and evaluation. Although validity in itself has been constantly preeminent among the different psychometric concepts over time, the concept has changed considerably during the 20th century. An historical overview of the evolution of the concept has been presented by Angoff (1988). From being defined by statistically and empirically based information, it has now moved towards a more theoretical and evidence-related term. Today, the concept of validity refers to the degree of which evidence and theory supports the interpretation of test scores (American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 1999). It is not the test itself that is validated, but the interpretation of test scores or measures required by proposed uses (McDowell & Newell, 1996).

In contrast to current views, most validation studies of assessments have traditionally referred to concepts of content validity, criterion-related validity, and construct validity, suggesting that there are different types of validity (APA, 1974). With the more current view that validity is a unitary concept, the AERA, APA, and NCME (1999) proposed the use of the following validity terms: evidence based on test content, evidence based on internal structure, evidence based on relations to other variables, and evidence based on response processes (i.e., the fit between the construct and the detailed nature of individual responses [scores] actually derived from examinees). A sound validation process of an evaluation tool will integrate different sources of evidence of validity into a coherent summary of the degree to which the existing evidence supports the intended interpretation of the scores for specific uses. The validation process also indicates suggestions for improvements related to refinement of definitions, revisions of the test procedures, and suggested areas that need to be studied in the future. In this thesis, the results from the four studies will be linked to these more recent concepts of validity evidence. Since some of the published studies in this thesis use the traditional concepts of validity types, a comparison

3 A fourth term, face validity, has also been used. This term, however, pertains more to public or clinical credibility of a tool, rather than the validity of the test scores or measures (McDowell & Newell, 1996; Messick, 1989).
OCCUPATION-BASED EVALUATION AND INTERVENTION

Validity of the Assessment of Motor and Process Skills
When Used with Persons with Mental Retardation

by

Anders Kottorp

Akademisk avhandling

Som med vederbörligt tillstånd av Rektorsämbetet vid Umeå Universitet, för avläggande av medicine doktorsexamen kommer att offentligt försvaras i aulan, Vårdvetarhuset, Umeå Universitet, fredagen den 23 maj 2003, klockan 13.00

Fakultetsopponent: Associate Professor Wendy J. Coster, Sargent College of Health and Rehabilitation Sciences, Boston University, Boston, USA
OCCUPATION-BASED EVALUATION AND INTERVENTION
Validity of the Assessment of Motor and Process Skills When Used with Persons with Mental Retardation

Anders Kottorp, OT Leg, MSc, Community Medicine and Rehabilitation, Occupational Therapy, Umeå University, SE-901 87, Umeå, Sweden

ABSTRACT

The ability to perform everyday life occupations is a critical component in both evaluation and intervention for persons with mental retardation (MR). While the ability to perform activities of daily living (ADL) has always been important for occupational therapy (OT) practice, there is an absence in OT literature and research with a focus on ADL and persons with MR. The overall aim of this thesis was to evaluate the validity of the Assessment of Motor and Process Skills (AMPS) for evaluation and intervention of ADL ability for persons with MR.

In order to evaluate the evidence of validity of the AMPS ability measures based on relation to level of MR, two groups of participants with MR were evaluated with the AMPS ($n=22; n=39$). The results indicated moderate relationships between ADL motor and ADL process ability measures and level of MR, despite different methods used for evaluating level of MR. The results also indicated that the results of the AMPS evaluation could be used to more directly describe and measure the consequences in performance of ADL tasks for persons with different levels of MR.

The evidence of validity of the AMPS was further examined in a study including participants with different types of developmental disabilities (e.g., MR, cerebral palsy, spina bifida) ($n=1724$). An application of many-faceted Rasch analysis was used to examine goodness-of-fit of the responses for the tasks, skill items, and participants included in the study. All tasks and all items except one demonstrated acceptable goodness-of-fit to the model on the ADL motor and ADL process scales. An expected proportion of participants demonstrated acceptable goodness-of-fit on the ADL motor scale. On the ADL process scale, a slightly lower proportion of participants than expected demonstrated acceptable goodness-of-fit. The results indicated further that persons with more severe levels of MR and persons with more limited ADL process abilities demonstrated different response patterns across tasks and possibly items.

The evidence of validity of the internal structure of the AMPS scales was also evaluated between persons with mild and moderate MR ($n=178; n=170$). Group specific ADL motor and ADL process skill item hierarchies were generated using many-faceted Rasch analyses and compared. The hierarchies of ADL motor and ADL process skill items remained stable across groups, indicating evidence of validity of the AMPS scales when used to evaluate persons with MR. The results also indicated that although persons with moderate MR demonstrated overall lower mean ADL motor and ADL process ability, they did perform some specific ADL motor and ADL process skills at a similar level as persons with mild MR.

Finally, the utility of the AMPS ability measures for detecting change were examined in an intervention study including three female participants with moderate MR. The study was based on a single case design and evaluated the effectiveness of a structured occupational therapy intervention program. Improvements were found for the participants in relation to the implementation of the program, but the pattern of changes were different between the participants and across the dependent variables. ADL process ability was the only variable that improved across all participants. The results supported the ADL process abilities as sensitive measures for detecting changes in ADL ability of persons with MR.

In conclusion, the results of these studies contribute to the evidence of validity of the AMPS ability measures and scales, specifically in relation to the evaluation of persons with MR. The finding that an OT program resulted in improved ADL process ability suggest that the results of the AMPS can be used to plan as well as evaluate outcomes of OT practice. Further research is also suggested in order to improve validity evidence and utility of the AMPS when used with persons with MR.

Keywords: mental retardation, intellectual disability, developmental disabilities, occupational therapy, activities of daily living, ADL assessment, performance skills, occupational performance, Many-faceted Rasch measurement, single case design.
of the classical and the current validity concepts are presented in Table 4.

Table 4. Comparison of classical and current validity concepts

<table>
<thead>
<tr>
<th>Classical validity types</th>
<th>Current sources of validity evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content validity</td>
<td>Validity evidence based on test content</td>
</tr>
<tr>
<td>Criterion-related validity</td>
<td>Validity evidence based on relation to other variables</td>
</tr>
<tr>
<td>Construct validity</td>
<td>Validity evidence based on internal structure</td>
</tr>
<tr>
<td>Not specified</td>
<td>Validity evidence based on response processes</td>
</tr>
</tbody>
</table>

Classical measurement statistics limitations
There are several statistical methods used to evaluate evidence of validity of test scores or measures. A major limitation of most tests used to measure human performance in general, and ADL ability more specifically, is that ordinal level data is improperly treated as interval data (Bond & Fox, 2001; W. P. Fisher, 1993; Merbitz, Morris, & Grip, 1989; Wright & Linacre, 1989). Since almost all statistics used (e.g., mean, $SE$, correlation coefficients) are based on interval data, the underlying assumptions of those statistics are violated when applied to ordinal data. Such applications may also lead to results that will mislead clinical interpretations (W. P. Fisher, 1993; Merbitz et al., 1989; Wright & Linacre, 1989). There are, however, specific techniques that can be applied to ordinal rating scales, to transform ordinal counts (scores) into linear measures, and produce sample free test statistics. One of those models is used in this thesis: the many-faceted Rasch (MFR) measurement model.

Rasch Measurement Models
In response to the awareness of limitations of traditional psychometric statistics when applied to ordinal data, Rasch measurement models have become preferred for constructing and
evaluating tests in rehabilitation and habilitation (A. G. Fisher et al., 1994). These models offer an alternative approach to test development, resulting in unidimensional linear measures (Rasch, 1980). Rasch analyses also generate statistics used to evaluate evidence of validity of the internal structure of the scales, based on formulations of expectations about what should happen when a group of persons take the test, and then confirming that the items in the test fit these expectations (Bond & Fox, 2001; A. G. Fisher, 1993; Rasch, 1980).

When developing a test using Rasch measurement models, the test developer should start with conceptualizing a variable as a single (unidimensional) construct represented by a line. Second, the developer should imagine placing people on this line based on the idea that they each have more or less of the trait or ability that is conceptualized. Third, the developer can start to build the line, by designing items of varying difficulty that are placed along the line. The range of the items from easy to hard determines the range of the test. The sensitivity of the test is determined by how many items that are positioned along the line, how close they are positioned, and how well they match the ability of the people the developer intends to test (A. G. Fisher, 1993). The proposed test is then evaluated by gathering data on people that are scored on the items defined and analysed using a Rasch computer program. Rasch computer programs, in the case of this thesis, the FACETS program (Linacre, 1987-2002) generates different types of statistics that supports the evaluation of how well the actual data fit the assertions of the model used.

The basic formula of the simple dichotomous Rasch model is:

\[
\log \left[ \frac{P_{ni}}{1 - P_{ni}} \right] = B_n - D_i ,
\]

where

- \( P_{ni} \) = Probability of a correct response from person \( n \) on item \( i \)
- \( 1 - P_{ni} \) = Probability of a wrong response from person \( n \) on item \( i \)
- \( B_n \) = Ability measure of person \( n \)
- \( D_i \) = Difficulty calibration measure of item \( i \)

Stated verbally, the simple Rasch model is as follows: The logarithm of the odds ratio between the probability of passing an item and the probability of failing an item equals the difference between the ability of the person and the difficulty of the item. More specifically, the data
analysis results in mathematically deriving item difficulty calibration (i.e., $D_i$, where the item is placed on the line) and person ability measure (i.e., $B_n$, where the person is placed on the same line). The item difficulty calibrations and the person ability measures are expressed in logits (log-odds probability units); as logits are equal-interval units, they are additive (Wright & Masters, 1982; Wright & Stone, 1979).

The simple Rasch model shown above is based on two specific assertions: (a) the easier the item, the more likely it is to be passed by any person; and (b) the more able the person, the more likely he or she is to pass harder items than is a person who is less able (Wright & Stone, 1979). There are sometimes more than two facets (items and persons) to consider when developing a test. One example of an additional facet is rater severity (Linacre, 1993; McNamara, 1996). The severity of the raters scoring a test has been shown to account for between one- and two-thirds of the variability in the data (Linacre, 1989). In order to estimate a person's ability, adjusted for rater severity, another facet must then be entered in the model. When adding facets, the model is also complemented by more assertions. For example, the Assessment of Motor and Process Skills (AMPS) is a four-faceted test where the person ability measure is adjusted for item difficulty, task challenge, and rater severity. The AMPS also uses a 4-point criterion-referenced rating scale, which then also must be taken into account in the estimations of person ability measures. This application of the MFR model of the AMPS has been described elsewhere (A. G. Fisher, 1993, 1994a, 1997b, 2001a).

Rasch Measurement Statistics

**Person ability ($B_n$) and item difficulty ($D_i$)**

If the odds of passing or failing an item is equal (50:50), resulting in an odds ratio = 1, the logarithm is calculated to zero, resulting in $B_n = D_i$; the ability of the person matches, or is equal to, the difficulty of the item. If we imagine that the person is more able than the item, we expect to see the odds of passing the item increase, and the odds of failing the item decrease, resulting in an odds ratio larger than 1, and a logarithm of the odds ratio larger than zero. This will then be equivalent to $B_n > D_i$. If, on the other hand, we imagine that the person is less able than the item, we expect to see the odds of passing the item decrease, and the odds of failing the item increase, resulting
in an odds ratio less than 1, and a logarithm of the odds ratio less than zero. This will then be equivalent to $B_n < D_i$.

**Standard error (SE)**

Each estimate for each person ($B_n$) or each item ($D_i$) is extrapolated from the available data. Therefore, there is also an error term provided with each estimate. This standard error term ($SE$) is unique for each item and each person included in the data. The size of the $SE$ is influenced by how well the data fit the model assertions, as well as how well targeted are the difficulty of the items to the abilities of the people. The size of the $SE$ will also be dependent on the amount of data available. If the person is assessed on many items, the $SE$ will often be relatively smaller than if the person is assessed on only a few items. Likewise, if the item difficulty calibrations are estimated based on large samples, the $SE$s will be smaller, reflecting the greater precision of the estimates. The $SE$ can also be used as a measure of reliability and to determine confidence intervals when comparing estimates.

**Mean square (MnSq) and standardized (z) goodness-of-fit statistics**

Rasch measurement models support estimating the positions of the persons ($B_n$) and the items ($D_i$) on the line in a way that the actual data gathered using the assessment as close as possible fit the calculated estimations of item difficulty calibrations and person ability measures. The FACETS computer program (Linacre, 1987-2002) generates statistics that can be used to evaluate the degree of fit between what the model expects and the actual data. For example, if a person performs unexpectedly on an item (unexpectedly well or poorly), this will result in a residual (error) for this person’s response on this item. The FACETS program summarizes all unexpected responses (residuals) into goodness-of-fit statistics. Where the index of fit is unsatisfactory for a specific person, a specific task, a specific rater, or a specific item, there is a warning that inferences based on the person ability measure, task challenge measure, rater severity measure, or item difficulty calibration may be inaccurate because the actual data evaluated did not meet the assertions of the MFR model of the AMPS.

There are two groups of goodness-of-fit statistics generated by the FACETS. The first are the **infit and outfit mean square (MnSq) statistics**: the relation between the observed and predicted values. If the
variability in the observed values is larger than expected (too much variability), the $MnSq$ values will be unusually high. If the variability in the observed values is less than the model predicts (too predictable), the $MnSq$ values will be unusually low. The outfit $MnSq$ statistics are quite sensitive to unexpected responses of a specific type (e.g., unexpected responses made by a person for whom a specific item is far too easy or far too difficult). This is adjusted for in the infit $MnSq$ statistics (Wright & Masters, 1982). The infit statistics are usually considered as the most informative in analysis of goodness-of-fit, as they focus on the degree of fit in the most typical observations in the data (McNamara, 1996). To be able to evaluate the infit and outfit values across response patterns from items and persons in order to detect significant variability, these are also standardized into infit and outfit $\xi$ statistics. There are different cutoff values suggested for different types of data where significant problems of fit are indicated (McNamara, 1996; Wright & Linacre, 1994).

Rasch-based Evidence of Validity

Persons who use Rasch measurement models to develop and evaluate tests use several Rasch-specific concepts of validity that can be related to the earlier current concepts of evidence of validity. A brief summary of the concepts used in this thesis is presented in Table 5.

Rasch Measurement Model Limitations

Since several of the limitations defined in classical test theory can be met with the use of modern test theory, the discussions and critique towards the latter has shifted from considerations of their advantages and disadvantages in relation to classical test models to considerations of matters as model selection, parameter estimations, and determination of model-data fit (Hambleton, 1989). That is, while it is now widely accepted that classical test theory and statistics are unsatisfactory, there remains some controversy as to what modern test theory and statistics are to be preferred. Some disadvantages of the Rasch measurement family of models have been raised in comparison to item response theory (IRT) models (Embretson & Hershberger, 1999; Hambleton, 1989). Persons supporting this view claim that the IRT models are more complex than the Rasch measurement models and will, therefore, fit a greater variety of test
data that may be rejected by more simplistic Rasch models. But the more fundamental issue underlying which of these models that is the best is grounded in two opposing measurement guidelines. The basic guidelines of the Rasch model are based on the assertions of the specific Rasch model used. These guidelines specify that the data must meet the assertions. If the data do not, then the data is somehow flawed. In fundamental contrast, IRT guidelines specify that if data do not fit the model, one need only change the model, adjusting it until the data fit.

Table 5. Comparison of classical validity concepts, current validity concepts, and Rasch validity concepts (as they apply to the AMPS)

<table>
<thead>
<tr>
<th>Classical validity types</th>
<th>Current sources of validity evidence</th>
<th>Rasch validity types</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content validity</td>
<td>Validity evidence based on test content</td>
<td>Similar</td>
</tr>
<tr>
<td>Criterion-related validity</td>
<td>Validity evidence based on relation to other variables</td>
<td>Validity evidence based on relation between generated person ability measures ($B_o$) and other variables</td>
</tr>
<tr>
<td>Construct validity</td>
<td>Validity evidence based on internal structure</td>
<td>Scale validity Item goodness-of-fit Task goodness-of-fit Differential item functioning/stability of item hierarchies across groups</td>
</tr>
<tr>
<td>Not specified</td>
<td>Validity evidence based on response processes</td>
<td>Person response validity Targeting test item difficulties to ability of sample Rater goodness-of-fit Rater severity Overall percentage of unexpected ratings $SE$</td>
</tr>
</tbody>
</table>
The relative value a researcher places on these guidelines will lead to qualitatively different approaches to data and, consequently, to different selections of models (Embretson & Hershberger, 1999). Therefore, the discussions of advantages and disadvantages between the two families of models should rather be viewed from a conceptual basis than based on specific operational characteristics (Hambleton, 1989). Although the one-parameter IRT model is mathematically equivalent to the Rasch measurement model, they are different in form and conceptualization, and will, therefore, not be equivalent in theory (Hambleton, Swaminathan, & Rogers, 1991).
Rationale

Rationale of this thesis

The ability to perform ADL occupations is a critical component in both the diagnosis and focus of interventions for persons with mental retardation. Even though ADL ability also has been an important focus and outcome for occupational therapy practice in general, the absence of literature and research with a focus on ADL ability and persons with mental retardation is striking. By using the ideas and formulations of client-centered practice in current occupational therapy frameworks, in conjunction with a focus on occupational performance (the basis of the occupational therapy profession), the following conclusions about ADL evaluations and interventions in the area of mental retardation can be made:

First, the methods used for data gathering are most commonly based on interviews of clients and/or caregivers. If ADL assessments are to contribute to the information gained in the occupational therapy data gathering process, they should emphasize the use of observational methods where the person is evaluated based on the quality of the doing. Observational methods support evaluation and determination of the occupational performance skills (the quality of actions) in ADL tasks.

Second, several persons with mental retardation may be able to perform PADL tasks independently, but show increased effort or inefficiencies during performance of IADL tasks. ADL assessments should therefore incorporate a broader range of ADL tasks for evaluating occupational performance in persons with mental retardation.

Third, a client-centered and top-down approach should be taken into account in the evaluation process. Since an essential feature of occupational therapy is that the intervention and desired outcomes are related to the person’s engagement in meaningful and purposeful occupations, it is critical that the evaluations, the interventions, and the chosen outcomes be client-centered and focused on the task performances the person finds meaningful and purposeful. It is also important that occupational therapists use a top-down approach in the evaluation process, by initially gathering information about what is of importance for the specific person – what the person needs and wants to do to be able to fulfil role expectations satisfactorily. Based
on this information, the occupational therapist then can continue to evaluate by observing and assessing the quality of the performance of those tasks identified by the person as priorities before considering the potential “causes” of the person’s problems in occupational performance (e.g., body functions, environmental factors).

Finally, it is critical that the ADL evaluation also provides information that guides the intervention planning process, supporting the occupational therapist with information useful in detecting problems, targeting goals, determining appropriate interventions, and evaluating the effectiveness of those interventions implemented. The need of functional assessments that directly evaluate occupational performances in real life contexts and provide information suitable for intervention planning, has been targeted for a long time as a need in the area of mental retardation.

The Assessment of Motor and Process Skills (AMPS) is a standardized ADL assessment developed for occupational therapy practitioners in order to meet several of the needs described above. Limited research, however, has been implemented to evaluate the evidence of validity when using the AMPS with persons with mental retardation for evaluation, planning interventions, and evaluating the effectiveness of those interventions.

Aspects of validity of a test’s scores or measures, as well as the interpretation of those, must be considered when choosing and using a test for use with a specific group of people. Validation of the AMPS ability measures should, therefore, include aspects about how comparisons can be made between persons or groups, support evidence that the measures are free of differential item functioning (item bias), and demonstrate how useful they are in detecting change. In order to evaluate the validity of the AMPS scales and ability measures when used with persons with mental retardation, methods of modern test theory, specifically a Rasch measurement model, are appropriate and preferred.
The overall aim of this thesis was to evaluate the validity of the AMPS for evaluation and intervention of ADL ability for persons with mental retardation. The specific research aims were:

- To evaluate the evidence of validity of the AMPS ability measures based on relationships to level of mental retardation (Study I and Study III).

- To evaluate the evidence of validity based on the response processes of the AMPS when used with persons with different types of developmental disabilities, including mental retardation (Study II).

- To evaluate the evidence of validity based on the internal structure of the AMPS when used with persons with different levels of mental retardation (Study II and Study III).

- To evaluate AMPS item difficulty hierarchies in order to compare and predict relative and actual ADL performance skills in persons with different levels of mental retardation (Study III).

- To evaluate the usefulness of the AMPS ability measures for detecting change when evaluating the effectiveness of occupational therapy interventions to improve ADL ability for persons with mental retardation (Study IV).
# METHODS

## Participants

The participants in this thesis were primarily adults with mental retardation involved in four different studies (I to IV). An overview of the participant demographics in the different studies is presented in Table 6. Additional details about the participants in each study are discussed in more detail below.

<table>
<thead>
<tr>
<th>Study</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of participants</td>
<td>61</td>
<td>1724</td>
<td>348</td>
<td>3</td>
</tr>
<tr>
<td>Overlap in study groups</td>
<td>None</td>
<td>Also in study III</td>
<td>Also in Study II</td>
<td>None</td>
</tr>
<tr>
<td>Gender M/F</td>
<td>31/30</td>
<td>897/827</td>
<td>155/193</td>
<td>0/3</td>
</tr>
<tr>
<td>Diagnosis</td>
<td>Mental retardation, specified as mild, moderate, and severe/profound; some with dual diagnoses</td>
<td>Different types of developmental disabilities</td>
<td>Mental retardation specified as mild and moderate</td>
<td>Moderate mental retardation</td>
</tr>
<tr>
<td>Mean age (range)</td>
<td>35.7 years (18-67)</td>
<td>27.7 years (3-93)</td>
<td>37.1 years (16-76)</td>
<td>24, 26, and 30 years</td>
</tr>
</tbody>
</table>

### Study I

The first study (I) was a collaborative study between Sweden and United States, and included 61 adults with the diagnosis of mental retardation. At the time of the study, the participants were all participating in community-based activity programs. All subjects were ambulatory and able to independently move around in their
immediate environments without assistance of other persons. Ten of
the participants had dual diagnoses of mental retardation and physical
deficits (e.g., cerebral palsy, seizure disorders) affecting occupational
performance. The participants represented a variety of severities of
mental retardation, and were selected based on personal knowledge of
the first and fourth authors of the study, who also conducted all of
the AMPS assessments.

Study II and III
The participants in Study II and III were gathered from the AMPS
database at Colorado State University, Ft Collins, Colorado, USA.
The participant data were initially selected based on the diagnosis of
the person. For study II, all participants with the diagnosis of autism,
cerebral palsy, mental retardation, spina bifida, specific developmental
disorders related to impairments of learning (e.g., attention deficit
disorder, dyslexia), and other developmental disorders, including dual
developmental diagnoses (e.g., mental retardation and cerebral palsy)
and unspecified developmental delays were initially included. The data
for each participant were then examined and systematically excluded
if (a) the participant had been co-scored by more than 10 raters for
purposes of rater calibration, as inclusion of such data would lead to
over-representation of that participant in the responses (scores)
derived, or (b) the data had previously been associated with rater
scoring error as evidenced by artificially high ADL motor or ADL
process ability measures, or (c) the participant had been scored by a
rater that did not meet the criteria of scoring the AMPS in a reliable
manner, as evidenced by unexpectedly high rater goodness-of-fit
statistics. Finally, data were also excluded for those AMPS tasks not
performed by at least 10 of the participants to be able to ensure
reasonably stable estimations of goodness-of-fit statistics for those
tasks. This method of data selection is consistent with earlier studies
that utilized existing participant data in the AMPS database (Stauffer
et al., 2000). A total of 1724 participants met the criteria for inclusion
in Study II. The distribution of participant demographics is presented
in Table 7.

In addition to the selection criteria described in Study II, the
additional selection criteria for Study III were that the participant (a)
had the single diagnosis of mental retardation, specified as mild or
moderate mental retardation, with no identified additional major
physical deficit (e.g., cerebral palsy, blindness), and (b) was 16 years of
Methods

age or older. These additional criteria were intended to minimize the risk of confounding variables, since both younger age and combinations of diagnoses may impact on the generated skill item profiles (A. G. Fisher, 2001a; Poulson, 1996). A total number of 348 participants were included in Study III.

Table 7. World region of included participants in study II and III

<table>
<thead>
<tr>
<th>World region/country</th>
<th>Study II (n, %)</th>
<th>Study III (n, %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>North America</td>
<td>234 (13.6)</td>
<td>44 (12.6)</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>470 (27.3)</td>
<td>137 (39.3)</td>
</tr>
<tr>
<td>Scandinavia</td>
<td>623 (36.1)</td>
<td>133 (38.2)</td>
</tr>
<tr>
<td>Netherlands</td>
<td>127 (7.4)</td>
<td>2 (0.6)</td>
</tr>
<tr>
<td>Australia/New Zealand</td>
<td>167 (9.7)</td>
<td>23 (6.6)</td>
</tr>
<tr>
<td>Asia</td>
<td>58 (3.4)</td>
<td>6 (1.7)</td>
</tr>
<tr>
<td>Israel</td>
<td>45 (2.6)</td>
<td>3 (0.9)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>1724 (100)</strong></td>
<td><strong>348 (100)</strong></td>
</tr>
</tbody>
</table>

**Study IV**

The potential participants for Study IV were persons with mental retardation who were referred to the Disability Services in the county of Västerbotten and met the following criteria: (a) they lived alone in their own apartments without support or with only minimal support from caregivers (i.e., supporting staff), and (b) they had an identified need and desire to develop skills in performing everyday life activities within their homes. Three females with moderate mental retardation participated in the study. They were 24, 26, and 30 years of age and lived in single apartments with two rooms and a kitchen in the community. They received support from caregivers with instrumental ADL tasks, but to different extents. The supporting interventions from caregivers for the participants were both verbal and physical.
Methods

Assessments, Procedures, and Interventions

The Assessment of Motor and Process Skills
All participants in all four studies were assessed by occupational therapists trained and calibrated to reliably administer the AMPS according to the standardized criteria in the AMPS manuals (A. G. Fisher, 1995, 1997a, 1999, 2001a, 2001b). Although different versions of the AMPS manual were used over time when assessing the participants, the administration criteria have remained essentially the same. The AMPS administration process included an initial interview to determine what types of ADL tasks the participant usually did or had experience performing. If the participant had a severe disability, resulting in difficulty communicating with the therapist, a caregiver supported the participant in the interview process. During the interview, the therapist considered which ADL tasks the participant usually did that were of an appropriate challenge to him or her and that matched the standardized tasks in the AMPS manual. The therapist then presented two or more appropriate and relevant tasks to the participant to choose from for the assessment. When appropriate, the participant was given the opportunity to practice less familiar tasks a few times, to ensure that he or she had the experience and an understanding of the essential goal of the task. The assessment environment was set up by the participant and therapist to be as natural as possible and to ensure participant familiarity. Examples of administering the AMPS with clients with mental retardation have been presented elsewhere (Bryze & Fisher, 1995; Kottorp & Fisher, 2001). After the participant had performed each of the chosen tasks, the participant’s performance was scored on each of the ADL motor and ADL process skills.

Assessment of level of mental retardation
In study I, the level of mental retardation of the sample from USA was assessed using standardized assessments of intelligence and adaptive behavior (Bruininks, Woodcock, Weatherman, & Hill, 1984). The Swedish sample was assessed through informal observations and semi-structured interviews. The assessment was guided by non-standardized test manuals (Granlund & Olsson, 1987, 1988; Terneby, 1985), using the model developed by Kylén (1974, 1981, 1985). This model uses a Piagetian framework (Piaget, 1953, 1955) to incorporate different aspects and categories of cognition into four distinct levels of abstraction. This can then be presented as a profile of a person’s
cognitive functioning, presented as levels of mental abstraction (level of mental retardation). The Swedish classifications made in Study I were summarized into a single global level of mental retardation, specified as A-level, B-level, and C-level, where A-level represented the most severe form of mental retardation. This classification system is related, but not identical to, the international levels of mental retardation described in ICD-10 (WHO, 1992) or DSM-IV (American Psychiatric Association, 1994) used to classify the sample from USA in study I (Kylén, 1981, 1985). The level of mental retardation in Studies II, III, and IV were classified according to the ICD-10 classification system (WHO, 1992). The level of mental retardation was obtained from the AMPS database or from the participants’ records.

**The Assessment of Awareness of Disability**

The Assessment of Awareness of Disability (AAD) was used in Study IV in conjunction with the AMPS. The AAD is a new assessment designed to measure the discrepancy between the person’s observed ADL skills (actual quality of occupational performance) and his or her self-reported description of ADL skills (perceived quality of occupational performance – awareness of disability, an underlying body function). The current version of AAD (Kottorp & Tham, 2002) consists of an interview guide with 12 questions. The questions relate to both global aspects of ADL performance and more focused questions targeting specific areas of skilled ADL performance.

When administering the AAD, the therapist first observed the participant perform each of two ADL tasks from the AMPS battery of 83 task choices (A. G. Fisher, 2001b). Directly after each task performance, the participant was interviewed about his or her perceptions of his or her performance. The focus of the interview was on different aspects of the performance (e.g., How was it for you to move or transfer when you [task name]? When you think about [task name], did you perform the task as we decided beforehand?). The therapist took detailed notes of the participant’s answers. After scoring the AMPS assessment, the therapist then compared (a) the observed problems in ADL performance that had been scored using the AMPS with (b) the participant’s described problems from the AAD interview. For example, the participant’s perceptions of his or her ability to move and/or transfer in the environment was compared to the participant’s scores on the AMPS skill items Stabilizes, Aligns,
Walks, Moves, Transports, and Navigates. The therapist then judged and scored the degree of discrepancy on each question using a 4-point criterion-referenced rating scale, where 4=no discrepancy, 3=a minimal discrepancy, 2=a moderate and detectable discrepancy, and 1=a severe discrepancy (Kottorp & Tham, 2002; Tham, Bernspång, & Fisher, 1999). The AAD had been found to show acceptable evidence of scale validity and overall person response validity (Tham et al., 1999).

A MFR measurement model (Linacre, 1989, 1993) has also been used to develop and evaluate the validity of the AAD ability measures (Kottorp & Tham, 2002; Tham et al., 1999). Similar to the AMPS ability measures, the AAD measures were adjusted for the severity of the rater that scored the AAD, the difficulty of the questions the participant was scored on, and the challenge of the tasks that the participant performed (see A. G. Fisher, 1993, 1994, 1997 for further details).

**Occupational therapy intervention program**

The occupational therapy intervention program used in Study IV was based on the Occupational Therapy Intervention Process Model (OTIPM): a professional reasoning model used to guide a client-centered, top–down evaluation and intervention process where the evaluations and interventions are grounded in the life situations and occupations that are relevant and meaningful to the person (A. G. Fisher, 1998, 2001a). The process was implemented as follows:

After gathering comprehensive information about each participant’s global performance context (including personal, environmental, and cultural aspects), and establishing a collaborative relationship with the participants and the caregivers, the occupational therapist had the participants and the caregivers identify and prioritize the occupational performances that were of the greatest concern to them, and that they wanted to target as priorities for improvement. The occupational therapist next administered the AMPS in order to gather specific information about the ADL actions that were effective and ineffective during task performance. Finally, the AAD was administered to support the clinical reasoning process about the potential reasons for the participant’s ineffective actions. That is, by clarifying what aspects of each participant’s task performance were most associated with greater or lesser awareness of disability, the occupational therapist
Methods

hoped to gain a better understanding of where the participant and therapist more or less agreed on the problems experienced and observed, in order to support and facilitate the intervention planning process. The participant’s performance of ADL tasks was also videotaped for use in the intervention program.

The interventions involved implementing individualized occupational therapy programs, and could be divided into adaptive interventions to compensate for ineffective actions in ADL performance, and restorative interventions to develop/enhance more effective actions in ADL performance. The adaptive interventions could be further categorised as modifications of task or physical environments, provision of adapted equipment, and teaching compensatory techniques. The restorative interventions could be further categorised as being focused on developing body functions (e.g., enhancing awareness of disability), and developing more skilled ADL performance routines. The intervention program sessions were 8 and 10 per participant.

Data analyses and Designs

Rasch measurement statistics
Each participant’s ADL motor and ADL process item raw scores in Study I, II, and III were analyzed using the FACETS computer program (Linacre, 1987-2002). In the development of the AMPS, a criterion of $MnSq \leq 1.4$ with an associated $\zeta < 2$ has been used to indicate acceptable goodness-of-fit to the MFR model of the AMPS. It is generally expected that 5% of the participants, the ADL tasks, or the ADL skill items will fail to meet this criterion by chance at $\zeta < 2$. Conversely, 95% of the participants, the ADL tasks and the ADL skill items were expected to meet his criterion. Since the ADL motor scale only had 16 items, the goodness-of-fit criterion for the ADL motor items was set as no more than one item exhibiting misfit (Goto et al., 1996; Magalhães et al., 1996). Finally, the overall proportions of misfit of individual AMPS skill item ratings was expected to be less than 1% at $\zeta \geq 3$ or $\zeta \leq -3$ (Fisher, 1994; Bernspång & Fisher, 1995b). The individual ratings are the item by item raw scores the rater assigns to each of the ADL skills for each ADL tasks the participant performs. One AMPS observation of two ADL tasks, therefore, involves 72 raw skill item scores (32 motor, 40 process).
In study II, the item difficulty calibrations, rater severity calibrations, and task challenge calibrations were anchored at preestablished values from the MFR model for the AMPS (Fisher, 2001a). The MFR analysis could be used, therefore, to determine how participants with developmental disabilities fit the preestablished MFR model for the AMPS.

To examine for clinically detectable differences in skill item difficulty calibrations in Study III, a method used earlier by several authors was adopted, where the criteria for a clinically detectable difference in relative and actual skill item difficulty calibration values was set at $>0.43$ logit. This is equivalent to a 95% confidence interval when the $SE$ for the skill item difficulty calibration value is 0.15 logit (Bernspång & Fisher, 1995b; Cooke, Fisher, Mayberry, & Oakley, 2000; Duran & Fisher, 1996; Goldman & Fisher, 1996; Magalhães et al., 1996; Oakley, Duran, Fisher, & Meritt, in press; Stauffer et al., 2000).

**Statistical analyses**

For all other statistical analyses, the Statistical Package for the Social Sciences (SPSS versions 9.0 and 10.0) was used. In order to examine the relationships between levels of mental retardation and ADL motor and ADL process ability measures in Study I, Spearman’s rank correlation statistics and Kruskal-Wallis nonparametric tests were used. The choice of statistics was based on small study samples and ordinal data (levels of mental retardation). In study II, independent $t$ tests and Chi-square tests were used to specifically examine for group differences in the generated ADL motor and ADL process ability measures. In study III, the skill item difficulty calibration hierarchies were compared using standardized difference ($Z$) (Wright & Masters, 1982; Wright & Stone, 1979). Independent $t$ tests were used to evaluate for group differences. In Study IV, lag-1 autocorrelation statistics were used to evaluate the impact of serial dependency in the datasets (Barlow & Hersen, 1984; Kazdin, 1982; Ottenbacher, 1986). Since the analyses indicated significant autocorrelations in the datasets, no additional statistics were used. In order to evaluate and compare the relationships between levels of mental retardation in ADL motor and ADL process ability measures from Studies I and III, effect sizes ($d$) were computed between pair of groups by level of mental retardation (Dawson-Saunders & Trapp, 1994).
Methods

Specific study designs

Studies I, II, and III were based on nonexperimental, descriptive comparison designs. The samples compared in Study I and III were not matched for age and gender, but no differences in mean age and gender distribution were found between groups in Study III. In study IV, a single case design was used, that included a baseline phase (A), an intervention phase (B), and a follow-up phase (C) (Barlow & Hersen, 1984; Kazdin, 1982; Ottenbacher, 1986). Since the occupational therapy intervention program included introducing environmental adaptations and adaptive equipment as intervention strategies, it was considered unethical to remove the adaptations from the clients after they had learned how to use them effectively. The intervention program, therefore, was implemented during the intervention phase (B), and in the follow-up phase (C), a withdrawal of the direct interventions from the occupational therapist was made, but environmental adaptations were not removed.

Ethical Considerations

Study I was approved by the Ethical Committee, Medical Faculty, University of Linköping, Linköping, Sweden prior to beginning. Approvals for study II and III were obtained before initiation of the studies from both the Ethical Committee, Medical Faculty, Umeå University, Umeå, Sweden and the Office of Regulatory Compliance at Colorado State University, Fort Collins, Colorado, USA. Approval for Study IV was obtained from the Ethical Committee, Medical Faculty, Umeå University, Umeå, Sweden before initiation of the study. Verbal and written information about the study was initially offered to three potential participants and their guardians, and signed permissions were obtained from all participants and their guardians after they volunteered to participate in the study. Although three participants were included in the study, only two participants completed the participation during all phases of the study. The third participant withdrew from the study, but agreed to allow usage of her data.
Results

Evidence of Validity of the AMPS Ability Measures Based on Relationships to Levels of Mental Retardation

The results from Study I revealed a moderately strong and significant relationship between levels of mental retardation and ADL process ability in the Swedish sample. The data revealed a similar pattern in the sample from USA, although the correlation was lower and not statistically significant. The correlation between levels of mental retardation and ADL motor ability was moderate and significant in the Swedish sample, but again lower and nonsignificant in the US sample. When combining both samples, the correlations found were moderately strong and significant for both ADL motor ability and ADL process ability, indicating an expected moderate relationship between levels of mental retardation and ADL ability, despite the use of different evaluation methods of level of mental retardation.

In study III, the results revealed further evidence of validity of the AMPS ability measures based on relationships between levels of mental retardation and ADL ability. More specifically, significant differences in both mean ADL motor ability and mean ADL process ability between participants with mild and moderate mental retardation were found.

In Tables 8 and 9, the means of the ADL motor and ADL process ability measures by group (level of mental retardation) for Studies I and III are presented. In Figures 1 and 2, boxplots of the distributions of the ADL motor and ADL process ability measures by group are shown. In Table 10, the effect sizes ($d$) between groups (level of mental retardation) are presented, indicating moderate to large effect sizes for both ADL motor ability and ADL process ability. The results also revealed that the effect sizes were lower but still moderate for both ADL motor ability and ADL process ability measures with the larger samples from Study III ($n=380$). Despite the differences in effect sizes and distributions, the mean ADL motor ability and mean ADL process ability of the samples were similar (see Tables 8 and 9).
### Table 8. ADL motor ability by levels of mental retardation

<table>
<thead>
<tr>
<th>Study</th>
<th>Level of MR</th>
<th>n</th>
<th>Range</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Mild (USA)</td>
<td>5</td>
<td>0.93 - 2.99</td>
<td>2.16</td>
<td>0.86</td>
</tr>
<tr>
<td>I</td>
<td>C-level (SE)</td>
<td>16</td>
<td>0.93 - 2.62</td>
<td>2.25</td>
<td>0.76</td>
</tr>
<tr>
<td>III</td>
<td>Mild</td>
<td>178</td>
<td>-0.58 - 3.83</td>
<td>2.14</td>
<td>0.95</td>
</tr>
<tr>
<td>I</td>
<td>Moderate (USA)</td>
<td>14</td>
<td>-0.92 - 2.65</td>
<td>1.76</td>
<td>0.65</td>
</tr>
<tr>
<td>I</td>
<td>B-level (SE)</td>
<td>17</td>
<td>0.07 - 3.10</td>
<td>1.48</td>
<td>0.92</td>
</tr>
<tr>
<td>III</td>
<td>Moderate</td>
<td>170</td>
<td>-3.39 - 3.91</td>
<td>1.59</td>
<td>1.10</td>
</tr>
<tr>
<td>I</td>
<td>Severe/profound (USA)</td>
<td>3</td>
<td>0.35 - 1.14</td>
<td>0.87</td>
<td>0.45</td>
</tr>
<tr>
<td>I</td>
<td>A-level (SE)</td>
<td>6</td>
<td>-0.27 - 1.26</td>
<td>0.52</td>
<td>0.63</td>
</tr>
<tr>
<td>III</td>
<td>Severe/profound</td>
<td>32</td>
<td>-0.62 - 2.59</td>
<td>0.92</td>
<td>0.81</td>
</tr>
</tbody>
</table>

### Table 9. ADL process ability by levels of mental retardation

<table>
<thead>
<tr>
<th>Study</th>
<th>Level of MR</th>
<th>n</th>
<th>Range</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Mild (USA)</td>
<td>5</td>
<td>0.27 - 1.58</td>
<td>0.75</td>
<td>0.50</td>
</tr>
<tr>
<td>I</td>
<td>C-level (SE)</td>
<td>16</td>
<td>0.40 - 1.30</td>
<td>0.86</td>
<td>0.33</td>
</tr>
<tr>
<td>III</td>
<td>Mild</td>
<td>178</td>
<td>-0.77 - 2.79</td>
<td>0.79</td>
<td>0.60</td>
</tr>
<tr>
<td>I</td>
<td>Moderate (USA)</td>
<td>14</td>
<td>-0.59 - 1.10</td>
<td>0.40</td>
<td>0.46</td>
</tr>
<tr>
<td>I</td>
<td>B-level (SE)</td>
<td>17</td>
<td>-0.55 - 0.76</td>
<td>0.19</td>
<td>0.39</td>
</tr>
<tr>
<td>III</td>
<td>Moderate</td>
<td>170</td>
<td>-2.90 - 1.97</td>
<td>0.20</td>
<td>0.85</td>
</tr>
<tr>
<td>I</td>
<td>Severe/profound (USA)</td>
<td>3</td>
<td>-0.91 - 0.23</td>
<td>-0.42</td>
<td>0.59</td>
</tr>
<tr>
<td>I</td>
<td>A-level (SE)</td>
<td>6</td>
<td>-2.62 - -0.45</td>
<td>-1.36</td>
<td>0.73</td>
</tr>
<tr>
<td>III</td>
<td>Severe/profound</td>
<td>32</td>
<td>-2.85 - 1.32</td>
<td>-0.48</td>
<td>0.81</td>
</tr>
</tbody>
</table>
Results

Note. S/P MR = severe/profound mental retardation, Mo MR = moderate mental retardation, Mi MR = mild mental retardation

Figure 1. Boxplots of ADL motor ability measures by level of mental retardation for Study I and Study III

Note. S/P MR = severe/profound mental retardation, Mo MR = moderate mental retardation, Mi MR = mild mental retardation

Figure 2. Boxplots of ADL process ability measures by level of mental retardation for Study I and Study III
Table 10. Effect sizes (d) between groups by level of mental retardation in ADL motor ability and ADL process ability

<table>
<thead>
<tr>
<th>Level of mental retardation</th>
<th>Mild</th>
<th>Moderate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study</td>
<td>I</td>
<td>III</td>
</tr>
<tr>
<td>ADL motor ability</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderate</td>
<td>0.82</td>
<td>0.53</td>
</tr>
<tr>
<td>Severe/profound</td>
<td>2.50</td>
<td>1.47</td>
</tr>
<tr>
<td>ADL process ability</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderate</td>
<td>1.38</td>
<td>0.80</td>
</tr>
<tr>
<td>Severe/profound</td>
<td>2.70</td>
<td>1.19</td>
</tr>
</tbody>
</table>

Evidence of Validity Based on Response Processes and Internal Structure

Fit among tasks and items
When examining the responses on the AMPS from the sample of persons with different types of developmental disabilities in Study II, all tasks (n=45) included in the analysis demonstrated acceptable goodness-of-fit to both ADL motor and ADL process scales. When examining the skill item goodness-of-fit statistics, all ADL process skill items demonstrated acceptable goodness-of-fit to the measurement model, and one item (Paces) failed to demonstrate acceptable goodness-of-fit on the ADL motor scale. Since none of the tasks, and no more than one skill item, failed to demonstrate acceptable goodness-of-fit, we concluded that the tasks and items showed acceptable scale validity when the AMPS was used to test participants with developmental disabilities.

Fit among participants
When examining the fit of the 1724 participants with developmental disabilities in Study II, 1639 (95.1%) demonstrated acceptable goodness-of-fit on the ADL motor scale, indicating person response validity on this scale. On the ADL process scale, 1586 (92.0%) demonstrated acceptable goodness-of-fit. Since the number of participants that failed to demonstrate acceptable goodness-of-fit on the ADL process scale exceeded the expectations, a supplementary
evaluation of goodness-of-fit was performed, examining the overall percentage of unexpected item ratings (A. G. Fisher, 1993, 1994a; Goto et al., 1996). When examining the total number of unexpected item ratings on the ADL process scale, only 390 out of 72,306 ratings were unexpected (0.5%). Although the percentage of participants who failed to demonstrate acceptable goodness-of-fit was higher than expected on the ADL process scale, the overall proportion of unexpected ratings for the AMPS process scale was acceptable, supporting evidence of validity based on a more detailed rating by rating goodness-of-fit analysis.

Nevertheless, because a slightly higher than expected percentage of participants failed to demonstrate acceptable goodness-of-fit to the MFR model on the ADL process scale, further examinations were performed in order to determine the source of the disturbance in the AMPS measurement system when it was used to evaluate participants with developmental disabilities. When comparing the data between the 138 participants who failed to demonstrate acceptable goodness-of-fit on the ADL process scale and the rest of the sample (n=1586), there was a significant difference between the groups related to ADL process skill ability. The participants who failed to demonstrate acceptable goodness-of-fit on the ADL process scale had significantly lower mean ADL process ability (M=-0.32 logit) than the rest of the sample (M=0.57 logit).

When examining the groups in relation to age and gender, the participants who failed to demonstrate acceptable goodness-of-fit did not differ significantly from those that did demonstrate acceptable goodness-of-fit. Age and gender were, therefore, not considered the source of disturbance causing the higher percentage of misfit among clients with developmental disabilities. When examining the participants who failed to demonstrate acceptable goodness-of-fit by diagnosis, participants with autism, and moderate or severe mental retardation were most likely to fail to demonstrate acceptable goodness-of-fit.

When examining the unexpected responses on individual items on the ADL process scale, 103 out of 390 unexpected ratings (26.4%) were related to the skill item Uses. When excluding the skill item Uses from the MFR model, the proportion of participants who failed to demonstrate acceptable goodness-of-fit was only reduced from 138 to
133 (4%), suggesting that Uses was not the source of disturbance. Finally, the participants who failed to demonstrate acceptable goodness-of-fit were evaluated in relation to the tasks performed. Approximately 60 participants (40%) of the 138 participants who failed to demonstrate acceptable goodness-of-fit performed either unexpectedly well or poorly on one of the included tasks in the AMPS evaluations.

Considering all of these secondary analyses, lower ADL process ability, specific diagnoses, and/or unexpected variations in specific task performances were the most likely sources of the higher than expected rate of participants who failed to demonstrate acceptable goodness-of-fit on the ADL process scale of the AMPS.

Further Evidence of Validity Based on Internal Structure

In Study III, the separately generated relative and actual ADL motor and process skill item difficulty calibrations were compared between participants with mild and moderate mental retardation. Since the participants with mild mental retardation demonstrated acceptable levels of goodness-of-fit on the AMPS process scale, compared to the participants with moderate mental retardation who did not (Study II), one could expect these two groups to have different ADL process skill item difficulty calibrations (differential item functioning), which could be a threat to validity. The comparison of the relative performance skill item calibration values revealed that all ADL motor skill items had a similar relative position in the hierarchies except one. The skill item Endures was relatively easier for participants with moderate mental retardation. Since only 1 item out of 16 had a clinically detectable difference in relative position on the ADL motor scale, the overall skill item hierarchy were judged to remain stable between participants with mild and moderate mental retardation.

On the process scale, all process skill items except one again remained stable between groups; the skill item Uses was relatively easier for participants with moderate mental retardation. These results indicated, contrary to the expectations, that the overall ADL process skill item hierarchy also remained stable between participants with mild and moderate mental retardation.
Results

The results indicated that overall, the ADL skill items that were relatively easier to perform for clients with moderate mental retardation also were relatively easier for clients with mild mental retardation despite any existing differences in overall ADL ability. That is, the linear hierarchical order of each specific skill followed a similar pattern in each group, but was centered (when adjusted for their mean difference in ADL ability) at different levels of ability, indicating evidence of validity based on internal structure of the AMPS, when employed with either persons with mild or moderate mental retardation.

Group Profiles in Actual ADL Performance Skills

To determine if participants with moderate mental retardation actually demonstrated a greater challenge in performing all of the ADL motor skill items than participants with mild mental retardation, the relative skill item calibration differences on all ADL motor skill items were adjusted into actual skill item difficulty calibration differences (Cooke et al., 2000; Oakley et al., in press) (Study III). These adjustments of the relative hierarchies (reported previously) enabled determination of which ADL skills that were not affected, and which skills that were to a greater degree affected with the presence of more severe forms of mental retardation. Such evidence provided clinically important information for predicting which ADL skills are more critical to target for interventions by severity of mental retardation, and which skills are stable across levels of mental retardation. These results, therefore, could also be viewed as a form of validity.

The ADL motor skill items Endures, Calibrates, Coordinates, and Manipulates, showed no clinically detectable differences in actual challenge between participants with mild and moderate mental retardation, suggesting that these skills were equally difficult for persons in both groups despite different levels of ADL ability. Differences in the actual challenge of all of the remaining 12 motor skills suggested that they were more difficult to perform for participants with moderate mental retardation than for participants with mild mental retardation.

Similarly, when adjusting the relative difference based on the overall mean difference in ADL process ability between the groups, the ADL
process skill items Uses, Handles, Navigates, and Notices/Responds, also showed no clinically detectable differences in actual difficulty between participants with mild and moderate mental retardation, suggesting that these skills were also equally challenging for both groups, despite different levels of ADL ability. Differences in the actual challenge of the remaining ADL process skills suggested that they were also more difficult to perform for clients with moderate mental retardation than for clients with mild mental retardation.

The Usefulness of the AMPS Ability Measures for Detecting Change

In order to evaluate the effectiveness of an occupational therapy intervention program, that used the information gained from the AMPS as a guide to intervention planning, a single case design with repeated measures over time were implemented for the three participants (Study IV). Improvements were found for all participants in relation to the implementation of the program, but the pattern of changes were different between the participants and across the dependent variables. One participant showed a detectable increase in ADL motor ability and ADL process ability, but not in awareness of disability. Another participant showed a detectable increase in ADL process ability and awareness of disability, but not in ADL motor ability. The last participant showed a detectable increase in ADL process ability only. ADL process ability was the only variable that improved across all participants. For the two participants that remained in the study across all phases, the improvements remained at 9- and 10-months follow-up. These findings provided evidence that the AMPS process scale was a sensitive measure of change.

An additional finding also emerged. That is, the results from the study also indicated that although the intervention program included enhanced feedback using video in order to improve awareness of disability, there were no clear patterns of such effects across the participants, despite the fact that all three improved in ADL process ability. This suggested that awareness of disability might not be strongly related or casual to improvements in ADL ability for persons with mental retardation.
DISCUSSION

Methodological Considerations

The overall aim of this thesis was to evaluate the validity of an occupation-based, client-centered assessment of ADL ability: the AMPS when used with persons with mental retardation. Before proceeding to discussing the main findings, there are methodological limitations that may impact on the conclusions made. Evidence of validity must always be judged in relation to the sample being evaluated. Since different methods of sample selection were used in Studies I, II, and III, the limitations of each of those will be discussed initially.

Sample selection, representation, and size

Studies II and III used participants with developmental disabilities (e.g., mental retardation) selected from the AMPS database. The use of databases in research has several advantages. It is a time-efficient method in gathering appropriate and large samples; large samples more easily support evaluation of evidence of validity related to differential item functioning across raters, countries, and ethnicities. There are, however, also limitations when using data that were not primarily collected for the purpose of a specific study. First, knowledge of the included participants is limited to the information gathered in the database. In these studies, limited information was available about the aetiology of the participants' diagnoses (e.g., specific syndromes). This could be a limitation in evaluating the results from Study III if participants with specific syndromes are unequally represented in the groups compared.

Second, research has shown that familiarity of environmental surroundings impacts the ability to perform ADL tasks (Darragh, Sample, & Fisher, 1998; Nygård, Bernspång, Fisher, & Winblad, 1994; Park, Fisher, & Velozo, 1994). In Study I, where the first and fourth author collected the data on the participants, information about environmental familiarity was directly obtained and could be considered in the interpretation of the results. Information about the familiarity of the environmental surroundings for the AMPS evaluations was not available from the database and, therefore, could not be controlled for as a potential extraneous variable in Studies II and III. Although the AMPS administration specifically addresses
having the person set up the environment and ensuring familiarity to minimize the impact of lack of environmental familiarity on the ADL performance of the person (A. G. Fisher, 1995, 1997a, 1999, 2001a, 2001b), it is unknown how well the occupational therapists managed to administer the AMPS according to these criteria.

Since none of the samples in Studies I, II, and III were systematically included based on random selection from a larger sample of participants with the specific diagnoses, the generalizability of the results may be limited. While one could argue that the entire available sample was selected, minimizing the risk for error, there remains reason to be cautious. As an example, participants with mild mental retardation may not be represented proportionally in the study samples, since these participants may not be in need of professional services and, therefore, not available for evaluation and interventions by occupational therapists. Similarly, participants with severe and profound mental retardation are not represented equally in proportion, probably because the clinical focus of evaluation and intervention for these participants is not concerning ADL ability specifically. In both instances, the AMPS would not have been administered. However, the samples do probably appropriately represent participants with mental retardation that are concerned about their ADL performances and receive occupational therapy services.

The issue of representation and selection must also be considered in relation to difficulties in defining samples of persons with mental retardation in more general terms. Since there are a variety of evaluation methods and diagnostic procedures used for classification, the variations in prevalence rates vary considerably between different studies, even within a specific region (Kebbon, 1987; Sonnander, Emanuelsson, & Kebbon, 1993). There are also differences in proportions between persons with mild mental retardation and more severe forms of mental retardation comparing samples from Sweden and USA. One issue that might impact on these differences in prevalence rates may be the different support systems available to persons with disabilities. As an example, in the Swedish welfare system, several services are available to persons without any specific diagnosis. Persons with mild mental retardation may, therefore, benefit from these services without being labeled with the diagnosis of mental retardation (Sonnander et al., 1993).
The subsamples compared in Studies I and III were also not matched according to age and gender. In Study III, the subgroups were evaluated for such biases, with no significant differences found in mean age and gender distribution between groups. Further supporting the probability that neither age nor gender were potential confounding variables in these studies, are several studies supporting stable ADL ability measures between persons aged 16 and approximately 50 years of age (Hayase et al., in press), and across gender groups (Duran & Fisher, 1996; Merritt & Fisher, in press).

In Study I, the sample sizes were rather small, and in combination with extreme measures, the correlation coefficients may be artificially inflated. By performing the same analysis with a larger sample, evaluations of such effects may be judged. For example, the effect sizes between level of mental retardation and ADL motor and ADL process ability measures were smaller, but still moderate with the larger sample from Study III. When examining the means and SDs between groups, the group means were similar despite the differences in sample sizes, but the SDs were larger in the group from Study III. These results suggest that the sample used in Study I was probably representative in relation to means in ADL ability by level of MR, but the variations in each group was considerably larger in the sample from Study III than exhibited in the smaller sample from Study I. It may be critical to carefully consider sample sizes when evaluating evidence of validity of an assessment in relation to other variables, since small sample sizes can be threats to the magnitude and significance of correlation coefficients.

Contrary to this is a possibility that the greater variations in Study III were due to error in coding the level of mental retardation and additional diagnoses of the participants. That is, in Study I, the first and fourth authors had direct knowledge of each participant's level of mental retardation and additional diagnoses. For all studies based on data in the AMPS database, there was no way to confirm accurate diagnosis. For example, the possibility that error did exist was suggested in the distribution of the data, where some participants' ADL motor ability measures were close to, or below, 0.0 logit, indicating marked limitations across all the ADL motor items (Fisher 2001a). Such marked limitations in ADL motor ability are more commonly found in persons with physical disorders, as for the participants in Study I with additional physical deficits. This may
therefore suggest that the participants in Study III with mental retardation also had additional physical diagnoses that either were overshadowed by the mental retardation, or were not correctly entered in the database by the occupational therapist (rater scoring error).

Similarly, when comparing the participants’ ADL process ability measures in relation to the cutoff criteria of 1.0 logit, that has been used to predict the need for assistance to function in the community (Bernspång & Fisher, 1995a; Fisher, 2001a, Hartman et al., 1999), some of the ADL process ability measures for participants with severe or profound mental retardation were close to, or above, this cutoff criteria. An ADL process ability measure close to 1.0 logit suggests that a person probably can or will be able to function independently in the community. This is a further indication of the presence of rater scoring error, where the severity of mental retardation may be incorrectly entered, which was not controlled for in this study.

Both of these potential rater scoring errors in Study III may, therefore, be one explanation of the larger distribution in the dataset for this study. It is therefore difficult to propose which of these two studies is most representative of persons with mental retardation in general, since smaller sample size versus higher rate of rater error are both threats to evidence of validity.

**Choice of specific Rasch statistics and criteria**
The FACETS computer program generates both infit and outfit goodness-of-fit statistics, but only the infit $MnSq$ and infit $\xi$ statistics were used for evaluating goodness-of-fit in Study II. Different opinions about the choice of statistics and criteria to evaluate goodness-of-fit in Rasch models has been discussed in more recent papers. Some authors have proposed, based on results from constructed data, that the outfit $MnSq$ has more power than the infit $MnSq$, and that the $\xi$ is less sensitive to sample size than are $MnSq$ statistics overall (Smith, 1991, 2000; Smith, Schumacker, & Bush, 1998). We chose to focus on the weighted goodness-of-fit statistics (infit $MnSq$ and $\xi$) in Study II, since we did not want to reject an item, person, or task as misfitting based on a few unexpected responses made by persons for whom the item or task was inappropriate (McNamara, 1996; Wright & Masters, 1982). We were also aware of
variations in human performance in general, and especially when scoring the doing of everyday life occupations. We used a more conservative approach for infit MnSq criteria, however, than has been suggested for clinical observational data (Wright & Linacre, 1994).

We set our criteria as \( MnSq \leq 1.4 \) with an associated \( \gamma < 2 \) to indicate acceptable goodness-of-fit to the MFR model when evaluating for too much variability in the scores. Although low \( MnSq \) and \( \gamma \) statistics also may target potential threats to validity, we did not primarily consider these in relation to the AMPS, based on the following reason. Low goodness-of-fit statistics indicate less variability in the observed values (Bond & Fox, 2001; McNamara, 1996). This can either be represented as a Guttman pattern across items, where raw scores too predictably decrease from 4 to 1 as items gets progressively harder, or as similar scores across all items, regardless of item difficulty. Although both of these patterns will indicate lower goodness-of-fit, they have not been considered as major threats to validity in the AMPS. The first pattern is almost never seen in AMPS raw scores. The latter pattern is sometimes observed on the ADL motor scale when the AMPS is used to evaluate persons with moderate to severe physical disabilities. In that case, the pattern of scores is typically logical, but does suggest that the ADL motor scale may not differentiate well among these persons (A. G. Fisher, personal communication, March 3, 2003).

**Single case research design limitations**

The single case design used in Study IV was chosen in order to specifically evaluate an occupational therapy intervention program over time in a limited number of participants. Although a number of different designs are used in single case research, they all share certain characteristics, including a dependent variable that is measured repeatedly, and an independent variable that is the intervention evaluated (Barlow & Hersen, 1984; Kazdin, 1982; Ottenbacher, 1986). The implementation of the single case design in Study IV does exhibit some limitations that may impact on conclusions made.

First, the number of evaluations in each phase was too few to allow for separate analyses of autocorrelations, or serial dependency in the dataset. Since the intervention program included aspects of learning new or compensatory methods of doing, a gradual improvement during intervention actually could be expected in some outcome variables and, therefore, also result in some degree of autocorrelation.
Discussion

But since the number of data points was too limited, especially in the baseline phase, it could be the fact that although autocorrelation was present in the total dataset, such serial dependency did not exist in the baseline phase. Research has supported that relying exclusively on visual inspection in single case research may lead to inconsistent interpretations of change (Ottenbacher, 1986). In addition, the presence of autocorrelation in a dataset often results in a Type II error in the visual interpretation (Bengali & Ottenbacher, 1998). In order to improve the analysis of intervention programs, not relying on visual inspection only, future single case studies should attempt to expand the number of assessments in each phase, to allow for evaluation of the degree of autocorrelation in each specific phase. The problem will remain, however, of withholding interventions longer than is ethically possible. This intervention program was based on a client-centered top-down approach where the participant's concerns for improved occupational performance was in focus. It was therefore difficult to argue for more assessments during the baseline phase, compared to the ethical dilemma of withholding interventions from a motivated client.

Second, the impact of repeated use of client-centered evaluations (such as the AMPS and the AAD) may in themselves also support improvements in performance. Such effects may have been present in Study IV. If so, these effects should be considered, and adjusted for, in the evaluation of the effectiveness of the intervention program itself. However, occupational therapists use the performance of occupations as both the means and goals (ends) for interventions. Although there is a risk of using repeated measures from a research perspective, the potential improvement by doing and reflecting upon occupational performance in evaluation may, from a clinical perspective, be a positive side-effect.

Main findings

The results from these studies suggest evidence of validity of the AMPS when used with persons with mental retardation. The findings will, therefore, support the multidisciplinary team with a client-centered ADL evaluation that contributes valid information about the ADL performance skills in persons with mental retardation. This information may also support the specific targeting of interventions in order to improve ADL functioning in these persons. In addition, the
results from this thesis confirm the effectiveness of those interventions, by directly measuring the magnitude of improvements in ADL performance. The results from the different studies also indicate clinical applications and suggest areas to be studied in more depth in the future.

**Relationships to level of mental retardation**

The evidence of validity of the AMPS ability measures based on relationships to level of mental retardation indicated moderate correlations between the ADL motor and ADL process ability measures and level of mental retardation. The results were in accordance with other studies examining relationships between underlying body functions and ADL ability (Doble et al., 1994; Robinson & Fisher, 1996). Still, the level of mental retardation in itself is not a sufficient predictor of ADL performance, indicated both by the range of ADL motor and ADL process ability measures within each group, but also by evaluating the coefficient of determination, or $r^2$ (Dawson-Saunders & Trapp, 1994). The level of mental retardation only predicted 28% of the variation in ADL motor ability, and 52% of the variation in ADL process ability in the total sample from Study I ($n=61$). This suggests that other factors besides the level of mental retardation impact on the ADL performance of persons with mental retardation, but also that the AMPS does not predict a person's level of mental retardation, nor that the level of mental retardation solely can predict a person's ADL ability. The occupational performance of a person with mental retardation is a complex set of interactions between environmental and intrinsic factors, including motivational, habitual, and cognitive capabilities (Kylén, 1979, 1981; Lyons et al., 1985). These findings suggest, therefore, that the AMPS can be used as an evaluation to describe and measure the consequences on ADL performance skills for a person with mental retardation. In order to understand the potential reasons for strengths and limitations in ADL skills for a person with mental retardation, several additional models and evaluations may then further supplement the clinical reasoning process.

There have been concerns that the definitions of levels of mental retardation are artificial, with limited utility in clinical practice (Luckasson & Reeve, 2001; Reiss, 1994). This criticism may also in part be reflected in the results from this study, as indicated by the large overlap between groups in both ADL motor and ADL process
ability measures. Several authors have, however, also emphasized the clinical differences between persons with mild and more severe forms of mental retardation (Accardo & Capute, 1998; Greenspan, 1999; MacMillan et al., 1993). The results from Study I and Study III do also, in part, support such clinical differences in overall mean ADL abilities between these groups. Since ADL functioning is included in the concept of adaptive behavior (Reschly et al., 2002; Widaman et al., 1991; Widaman & McGrew, 1996), and therefore a part of the diagnosis of mental retardation, it is critical that an ADL assessment exhibit relationships to such categorization in order to support evidence of validity. The engagement in IADL and community tasks is also strongly related to levels of adaptive behavior (Felce & Emerson, 2001a). The results of this thesis supported the expected moderate correlations between the ADL motor ability measures and ADL process ability measures of the AMPS with levels of mental retardation, demonstrating evidence of validity in relation to levels of mental retardation.

Response processes and internal structure

The analysis of response processes on the AMPS for persons with different types of developmental disabilities, including persons with mental retardation, indicated that the AMPS tasks and items showed acceptable scale validity. The results suggest that the constructed AMPS task hierarchy and skill item hierarchy are valid when used to assess persons with different types of developmental disabilities. That is, the assertions formulated in the MFR model of the AMPS are valid for persons with developmental disabilities regarding ADL tasks and skill items (e.g., the ADL tasks that are generally easier for persons are also easier for persons with developmental disabilities; the ADL motor skill items that are generally harder for persons are also harder for persons with developmental disabilities). Based on the evidence of scale validity of the AMPS, the hierarchies of tasks and items can therefore be used to predict ADL performance of persons with developmental disabilities, including persons with mental retardation (A. G. Fisher, 1994a).

When examining the distribution of participants in relation to ADL motor and ADL process ability, the results also indicated that the AMPS can be used to evaluate persons with developmental disabilities over a large continuum of ADL ability. This supports the inclusion of both PADL and IADL tasks when evaluating persons with mental
Discussion

retardation on different levels and different abilities. The absence of ceiling and floor effects in the sample also indicated that the AMPS scales are adequately targeted to the ability of the sample, indicating further evidence of validity of the AMPS.

Further examination of ADL motor and ADL process ability measures in the sample suggested that persons with developmental disabilities in general, and persons with mental retardation specifically, do exhibit both ADL motor skill limitations and ADL process skill limitations in occupational performance. Despite the earlier methodological considerations related to rater scoring error in persons with mental retardation and marked limitations in ADL motor skills, the clinical results from these studies indicate that problems in ADL motor skills, to different extents, are impacting on the quality of the doing in persons with developmental disabilities. Even though these problems might be minor for many persons with mental retardation, with only a minimal increase in effort in ADL performance, it is important to bear in mind that persons with mental retardation also may have additional diagnoses that impact on ADL motor skills (e.g., cerebral palsy) (Larson et al., 2001). In addition, persons with mental retardation have difficulties in adapting to problems in general. Therefore, motor limitations must not be neglected in evaluation of ADL skills for persons with mental retardation, since these will contribute to the overall adaptive limitations exhibited in ADL performance. The major limitations in ADL performance for persons with mental retardation, however, are related to inefficiencies (inappropriate use of time and energy), safety risks, and need for assistance in ADL performance. Such problems are operationalized, and measured, as limitations in ADL process skills in the AMPS.

When examining the response processes of the participants in Study II, a slightly higher than expected rate of participants did not demonstrate acceptable goodness-of-fit to the model on the ADL process scale. When examining this subgroup of participants, we concluded that lower ADL process ability, specific diagnoses (moderate mental retardation, severe mental retardation, and autism), and/or unexpected variations in specific task performances were the most likely sources of the disturbance in the measurement system. One reason for a participant to receive high goodness-of-fit statistics is that the participant performs unexpectedly well on a more
Discussion

challenging task, or unexpectedly poorly on a less challenging task. Exceptional skills are occasionally displayed in persons with developmental disabilities, and referred to as savant behavior, or splinter skills (Miller, 1999; O'Connor & Hermelin, 1988). Such skills are, however, not described in relation to ADL tasks specifically, but rather algebraic calculation tasks or arts. Clinical experience has though shown that some persons with more severe forms of mental retardation and/or autism can be unexpectedly skilled performing specific ADL task that may have been learned at home or in a sheltered workshop, although their general ability in other ADL tasks is more limited. This could explain the higher proportion of misfit among persons with more limited ADL process ability and severe forms of mental retardation and/or autism. This suggests some disruption in the measurement process when the ADL process scale is used to test persons with more severe developmental disabilities. Although the results from Study II did not indicate that specific ADL tasks were disrupting the measurement system, future evaluation of differential item functioning in specific ADL tasks for persons with more severe developmental disabilities will support or reject such conclusions.

Study II did not specifically evaluate the response processes from persons with mental retardation, which would have been more in line with the overall aim of this thesis. The information gained from evaluating a larger sample with persons with developmental disabilities, however, supports the clinical utility of the AMPS as a general assessment of ADL skills in habilitation services. The results also indicated that the unexpected responses found were not only linked to persons with mental retardation, but rather represented a larger phenomenon across developmental disability groups.

Since persons with moderate and severe mental retardation did not demonstrate acceptable levels of goodness-of-fit to the MFR model on the ADL process scale, this could indicate that the skill item hierarchy does not remain stable among these groups. This hypothesis was also, in part, supported by the higher amount of unexpected responses related to the skill item Uses found (although further examination of this item did not seem to be the source of disturbance to the measurement system). The results from Study II indicated further examination of the internal structure of the AMPS for persons
Discussion

with different levels of mental retardation, which was implemented in Study III.

The examination of the relative hierarchy in Study III suggested that the linear hierarchical order of ADL motor and ADL process skills did follow a similar pattern in each group, supporting further evidence of validity based on internal structure of the AMPS. These results did therefore not support the hypothesis of differences in skill item hierarchies between persons with mild and moderate mental retardation, as indicated in Study II. So, the indications of differential item functioning in the AMPS process scale found in Study II were not related to these specific diagnoses in skill item hierarchies.

**Actual skill profiles in ADL performance**

The results from Study III also supported determinations of which skills actually are more or less affected in ADL performance for persons with different severities of mental retardation. The ADL motor skill items Calibrates, Coordinates, Endures, and Manipulates showed no clinically detectable differences in actual challenge between participants with mild and moderate mental retardation. Clinical experience and research have suggested that fine motor functions and postural control are commonly the only body functions that are impaired in persons with mild physical deficits (A. G. Fisher, 1994a; A. G. Fisher, Murray, & Bundy, 1991). Although the fact that this literature primarily focuses on body functions rather than occupational performance skills (Activity dimension), it is possible that the goal-directed task actions Calibrates, Coordinates, Flows, Grips, and Manipulates are among the ones that are most readily affected by mild physical impairments. Since the mean of ADL motor abilities for the participants with mild and moderate mental retardation were 2.14 and 1.59 logits, respectively, there was an indication that both groups of persons with mild and moderate mental retardation exhibited only minimal or no increase in effort when performing ADL motor actions (A. G. Fisher, 2001a). The results from this study suggested, therefore, that the fine ADL motor skills and the positioning of body to promote efficient arm movements (Positions) may be primarily critical to target in order to minimize the effort in ADL performance for persons with mental retardation. Furthermore, the results from Study III also suggested that the fine ADL motor skills were equally difficult for both groups,
Discussion

despite different levels of ADL ability and levels of mental retardation.

The results from Study III were also in accordance with those of Oakley et al. (in press), comparing participants with dementia of Alzheimer’s type with a group on nondisabled controls. Their results indicated that although both cognitive and physical body functions of persons with dementia of Alzheimer’s type decline over time, they remained equally skilled in regulating the force and speed of movements (Calibrates), enduring throughout the performance of ADL tasks (Endures), and grasping objects efficiently (Grips).

The findings that the ADL process skill items Handles, Navigates, Notices/Responds, and Uses also showed no clinically detectable differences in actual difficulty between participants with mild and moderate mental retardation, suggested that these skills were also equally challenging for both groups, despite different levels of ADL ability and levels of mental retardation. An explanation of similar results found for persons with dementia of Alzheimer’s type has been suggested by Cooke, et al. (2001), suggesting that such patterns may be related to “sparing” of procedural knowledge (knowing how to do) (Bäckman, 1992; Dick, 1992), in the presence of loss of declarative knowledge (knowing what to do). When comparing persons with dementia of Alzheimer’s type to a group of nondisabled controls, the ADL process skills that were least affected by the overall loss of ADL process ability in persons with dementia of Alzheimer’s type were Handles, Navigates, and Uses (Cooke, et al., 2001). The results from Study III suggest a similar pattern in ADL performance when comparing persons with mild and moderate mental retardation. Although persons with moderate mental retardation have more limited underlying cognitive functioning and are more limited in ADL process ability in general than persons with mild mental retardation, they exhibit equivalent ability to use and support tools and materials appropriately, and respond appropriately to task related cues in the environment, as do persons with mild mental retardation.

Although there might be some concern about comparing persons with mental retardation with persons with dementia of Alzheimer’s type, the similarities in the results is still striking. While research has supported, from a neuropsychological view, that mental decline in clients with Down syndrome does not reproduce the cognitive
profiles (Body function dimension) of persons with dementia of Alzheimer's type (Caltagirone, Nocentini, & Vicari, 1990), the results from this thesis suggest similarities in skill profiles in occupational performance (Activity dimension) between persons with mental retardation and persons with dementia of Alzheimer's type. Despite the fact that these groups differ in relation to age, aetiologies, related disorders, and prognosis, their limitations in cognitive functioning are associated with similar profiles in ADL motor and ADL process skills.

However, the similarities found between persons with mental retardation and persons with dementia of Alzheimer's type may be more related to groups on similar levels of ADL process ability, than only the severity of cognitive functioning for these specific diagnoses. That is, progressively lower ADL motor ability and ADL process ability may impact on profiles in ADL skills in similar patterns, despite the reason for limitations in ADL ability. Future research, including heterogenous samples on different ADL process ability levels, may provide or reject evidence for such conclusions.

_Evaluation of occupational therapy interventions_

There is an increased awareness today of the need to provide evidence that the services offered to our clients are the best possible (Holm, 2000; J. Rogers & Holm, 1994; Tickle-Degnen, 1999). It is therefore crucial that occupational therapists clearly can describe for others what the interventions consist of, how they are to be implemented in order to achieve the best possible outcome, and link the interventions to clinically relevant outcomes that demonstrate objective evidence of the effectiveness of those. Although the evaluation of independent living skills training program has not always been positive (Lozano, 1993), the results from this thesis, evaluating a client-centered occupational therapy program directly in ADL tasks, did indicate sustained improvements in ADL process ability in the participants. In addition, the intervention program in this study had positive effects in participants with moderate mental retardation, even though the level of mental retardation in other studies has shown to be a major constraint to achieve a higher level and engagement in everyday life tasks (Felce & Emerson, 2001b).

The results from Study IV indicated that the AMPS could be used to plan interventions, and also detect improvements in ADL process
ability. Although differences were found in mean ADL motor ability and awareness of disability, these changes across participants were not verified in the visual analysis of the data. The differences in impact on the outcome variables may be due to the initial abilities of the participants, but also the different types of specific interventions provided within the program.

When considering the type of intervention, it was interesting to note that although the occupational therapy interventions included enhanced feedback in order to improve the awareness of disability, there was no clear pattern indicating such effects across the participants. In addition, an interesting finding was that all three participants improved in ADL process ability, despite the fact that only one showed a clear change in awareness of disability. This suggests that awareness of disability may not be as strongly related, or casual, to improvements in ADL ability as is often assumed. It should be noted, however, that the findings of this study are in contrast to those of Tham et al., (2001), who did find improvements in both awareness of disability and ADL ability after an occupational therapy intervention program for persons with unilateral neglect.

A critical component in outcome studies can be the evaluation of transfer of effects into other settings or task performances than those specifically targeted (Langone, Clees, Oxford, Malone, & Ross, 1995). Such effects were not directly evaluated in this study. Since the intervention program was based on a client-centered approach, there was no logical reason from the perspective of the client to evaluate the performance in other ADL tasks that were not based on her own choices. In addition, since the interventions included modifications of task or physical environments, and provision of adapted equipment, improved skills in settings or situations other than those where the interventions were provided were not directly expected. This does not mean that it is not critical to consider such effects, if the outcomes are expected to transfer to other settings or task performances. Such evaluations would also enhance the evaluation of the effectiveness of occupational therapy interventions. The AMPS is also a suitable tool for evaluating transfer effect due to the large number of available ADL task options, and the MFR model which allows comparisons of persons' ADL ability measures across different tasks (A. G. Fisher, 1993, 1994a; Linacre, 1993).
A major limitation of Study IV was the amount of variability in the data. Even though differences in outcomes were present for all participants during and after the implementation of the intervention program, given the variability, the results should be viewed with caution. Despite this, the results from this study suggest that the AMPS is sensitive enough to detect changes in performance. While a sensitive measure is more likely to detect change, it may also become harder to separate the effects from interventions from those caused by changes in motivation, mood, environmental conditions, or other factors. Since it can be expected that persons vary in ADL performance between two occasions (Ellison, Fisher, & Duran, 2001; A. G. Fisher, 2001a; Kirkley & Fisher, 1999), a conservative approach to the visual analysis was used in this study. It may be critical to determine the expected variability over time in ADL performance specifically for persons with mental retardation in order to provide a better basis for judging the stability in ADL performance when no intervention is provided, as well as detecting real changes that are due to intervention. The alternative, using tools that are less sensitive for detecting minor differences in performance, would likely lead to more stable measures over time, but also result in failure to detect minor, but real, changes in ADL ability for persons with mental retardation.

Clinical applications

The purpose of evaluation of a specific test should not only be to "prove that this test is good," but rather demonstrate "how this test can be improved even further" (Ovretveit, 1998). A sound validation process should also include suggestions for improvements (AERA, APA, & NCME, 1999). In the case of this thesis, this means discussing the potential threats to validity evidence of the AMPS when used with persons with mental retardation, and how to minimize those in the clinical administration process. It is also critical that the various forms of validity evidence are integrated into a coherent summary that supports the proposed use of the AMPS with persons with mental retardation. Some of the clinical applications from the thesis are summarized below.

- The AMPS can be used in order to evaluate the consequences of mental retardation on the quality of performance of ADL tasks.
• The AMPS can also be used as a valid ADL assessment across a wide span of developmental disability groups, enabling comparisons of ADL performance skills between specific diagnostic groups.

• The occupational therapist must be cautious in suggesting ADL tasks options, by carefully conducting an interview with the person, but also caregivers and/or family members, in order to avoid offering tasks that are over-learned. This is especially important when evaluating persons with more severe forms of developmental disabilities. Although determination of such exceptional skills in performing ADL tasks may be clinically valuable, it will invalidate the generated ADL ability measures, and also delimit the predictive validity of the AMPS.

• By evaluating the ADL performance at the level of skills for a person with mental retardation, detailed information about strengths and limitations may be obtained. By evaluating the profiles of ADL motor and ADL process skills, further information for planning interventions can be obtained for groups on different levels of mental retardation.

• The ADL skill items that are found on the upper (positive) end of the ADL skill item calibration hierarchy can be viewed as relative strengths in ADL performance and may, therefore, not be the primary target for interventions in order to improve ADL performance for a specific person or group.

• The ADL skill items that are on the lower (negative) end of the ADL skill item calibration hierarchy are the most difficult to perform efficiently, safely, independently, and without effort and may, therefore, be the primary targets in order to improve overall ADL performance for a specific person or group.

• Both ADL motor skill limitations and ADL process skill limitations are critical to consider in the evaluation and intervention of persons with mental retardation, even if the persons do not have any diagnosed physical disorders.
Some specific ADL motor skills and ADL process skills are equally challenging for persons with mild and moderate mental retardation, despite differences in general ADL ability.

The ADL adaptation skill items are the most difficult to perform for persons with mental retardation. This supports interventions related to introducing adaptations of the physical and/or social environments, introduction of different types of technical aids, simplifying task performances, or the teaching of alternative compensatory strategies, all in order to support the person to more effectively compensate for the observed problems. Interventions can also be focused on further developing more effective adaptation skills, by giving the person opportunities to practice and develop habits and routines in the performance of ADL tasks.

The AMPS can be used to evaluate the effectiveness of such interventions for persons with mental retardation, by generating valid ADL motor and ADL process ability measures.

Although the level of mental retardation can't be changed by treatment, ADL ability can be improved in persons with mental retardation by occupational therapy interventions.

Future research

The results from these studies indicate several areas for future research. First, the findings that the ADL skill item hierarchies shared similar characteristics between different diagnostic groups (mental retardation and dementia of Alzheimer's type) indicate that ADL skills may be viewed from a different perspective than the traditional diagnosis-related one. Future research studies could therefore focus on generating further information about the ADL skill profiles across different ADL ability levels and diagnostic groups. Such information could also support the development and implementation of intervention strategies across diagnostic groups.

Another area for future research is the evaluation of interventions in order to improve ADL ability in persons with mental retardation. The
results from these studies will generate information that supports a systematic approach in intervention planning, where the development of specific interventions are specifically targeted to the ADL skills that are determined to be the most critical for performance. The implementation of such interventions must also be evaluated in relation to their effectiveness. This thesis offers one approach in evaluating the effectiveness of interventions in ADL tasks. A single case design can be used not only to evaluate specific outcomes, but also serve as a basis for improvements of the interventions offered. Such designs can, therefore, initially support the refinement of an intervention program, which then can be implemented in a context that supports a higher level of generalizability of results (e.g., randomized controlled trials).

Another critical aspect in relation to future research focused on the evaluation of interventions is the variability of ADL performance over time. The results from this thesis indicate that variability must be considered carefully in persons with mental retardation when evaluating outcomes. Future research may, therefore, specifically evaluate the stability of ADL performance over time for persons with disabilities when no intervention is implemented, as a basis for evaluating variability in future intervention research studies. The issue of variability over time also has important clinical implications.

Occupational therapy practitioners have a long tradition in offering interventions that promote independence and minimize the need of assistance in occupational performance for persons with disabilities in general. It is therefore critical to identify which variables are most critical for predicting the need of assistance. Further examination of the use of the ADL motor and ADL process ability measures in predicting need for assistance for persons with mental retardation to function in the community is therefore also an important area for future research.

There is a need in occupational therapy for developing new assessments that focus on phenomena closely linked to occupational performance. Such development should also preferably be based on modern test theory (e.g., Rasch models) rather than classical test theory, in order to overcome the threats to validity evidence of the latter. The development of new and relevant occupational therapy assessments should also incorporate a client-centered focus. Future
research should, therefore, also focus on providing further evidence of validity of the AAD, including the utility of the use of the AMPS and AAD in a clinical intervention planning process.

Finally, ADL ability for persons with mental retardation must be linked to a wider perspective of meaning, participation, and quality of life. Since occupational therapy practitioners are focusing on improvements in the doing of meaningful occupations, it is important that future occupational therapy research use several methods (quantitative and qualitative) for evaluation of the effectiveness of our interventions.
CONCLUSIONS

- The AMPS ability measures exhibit expected relationships to level of mental retardation, supporting evidence of validity when used with this population. Furthermore, the AMPS tasks and items showed acceptable scale validity when used with a large sample of persons with different types of developmental disabilities, including persons with mental retardation.

- Evaluation of the individual responses from the AMPS evaluations supports overall person response validity of the AMPS when used with persons with developmental disabilities. Some concerns when using the AMPS with persons with more severe forms of developmental disabilities and limited ADL process abilities are suggested, and must be considered in the administration of the AMPS for these persons.

- The overall relative skill item hierarchies remain stable between participants with moderate and mild mental retardation, indicating evidence of validity based on internal structure of the AMPS. Although persons with moderate mental retardation demonstrate more limited ADL ability in general, compared to persons with mild mental retardation, they are equally skilled in performing some specific ADL motor and ADL process items as are persons with mild mental retardation.

- An occupational therapy intervention program for persons with mental retardation improved ADL process ability, and the effects remained at long-term follow-up. The ADL process ability measures can be used to detect minor changes in ADL ability for persons with mental retardation.
ACKNOWLEDGEMENTS

My years in this research process have been a time of personal enrichment to me that I now can look upon with gratitude. I wish to thank all you who have been supporting me and, in various ways, contributed to this thesis. I wish to direct my warmest gratitude particularly to my two supervisors:

Birgitta Bernspång, my main supervisor, colleague, and friend, for your whole-hearted engagement, supervision, and support throughout this process. Your encouragement, your concern about me, as a doctoral student and person, and your commitment to develop occupational therapy praxis and research has rendered you my greatest appreciation.

Anne G. Fisher, my second supervisor, colleague, and friend, for your total commitment and endless support in the research process. The excellence in your supervision has stimulated my growth and development. Your deep knowledge in research and devotion to occupational therapy has rendered you my admiration and respect.

I also want to thank the following persons who substantially have contributed to the thesis:

Kim Bøyce, my research colleague in Study I. Your generous support and guidance, in combination with your personal and professional commitment to persons with mental retardation, made the collaboration process a pure enjoyment.

Monica Hälgren, my research colleague in Study IV. I will always remember your excellent professional skills in implementing occupational therapy for persons with mental retardation, but also for your respect and empathy for the participants’ own experiences.

I also want to thank the following persons, who have supported and stimulated me in different ways:

My colleagues at Dagcenter Ekholmen, Linköping, who initially stimulated me in linking theory and research into clinical praxis.
Acknowledgements

All my colleagues at Disability Services, Västerbotten läns landsting, especially Agneta Nilsson and Birgitta Bäckström, for friendship, continous concern, and for always wishing me welcome back. To Sofie Edberg and Ann-Sofie Kaati for providing excellent support and working conditions.

All colleagues at the Department of Occupational Therapy, Umeå University, Umeå, especially the research group, for challenging discussions, valuable seminars, and nice talks.

The AMPS National and International Faculty and staff, for enjoyable collaborations and supporting friendships along the way.

Kerstin Tham, Karolinska Institutet, for past, present, and future collaboration in the development of AAD.

To all occupational therapists and participants engaged in data collection for the studies, especially the three female participants who were engaged in the intervention study (Study IV).

Finally, I will also thank my family and close friends who always have been there during all these years. Hanna, Emilia, Niklas, Ingrid, and the rest of my family. My love to all of you.

This work was supported by grants from the Vårdal foundation, Sävstaholmsföreningen, the County of Västerbotten, the Disability Services, Västerbotten, and the Swedish Association of Occupational Therapists (FSA).
References


References


References


73


measurement: Implications for the field of mental retardation (pp. 15-42). Washington, DC: American Association on Mental Retardation.


