Contents lists available at ScienceDirect

Developmental Review

journal homepage: www.elsevier.com/locate/dr

A formal framework for modelling the developmental course of competence and performance in the distance, speed, and time domain

Dietrich Albert [a], Michael D. Kickmeier-Rust [a,*], Fumiko Matsuda [b]

[a] *Department of Psychology, University of Graz, 8010, Graz Austria*
[b] *Fukuyama University, Japan*

ARTICLE INFO

Article history:
Received 20 August 2007
Revised 6 May 2008
Available online 17 July 2008

ABSTRACT

The developmental course in the distance–speed–time domain is still a matter of debate. Traditional stage models are contested by theories of continuous development and adaptive thinking. In the present work, we introduce a formal framework for modelling the developmental course in this domain, grounding on Competence-based Knowledge Space Theory. This framework, as a more general case, widely includes assumptions and facets of previous models and covers empirical findings collected based on different experimental paradigms. By a distinction of latent competences and observable performance, model validation is not bound to a certain experimental paradigm and no one-to-one correspondence between competences and tasks is required. Therefore, the framework has the potential to bridge the gap between stage models and models of continuous development. The approach also precisely defines misconceptions, for example overgeneralization, and empirically investigates their occurrence. In the present work, we established a prototypical model for the development of understanding the distance–speed–time system. We extended this model with definitions based on different perspectives of overgeneralization. The assumptions of the model and its extensions were examined on the basis of the results of two empirical investigations using six judgment task types. The results yielded a reasonably good fit of model and data. No evidence was found for the occurrence of overgeneralization in this domain. The theoretical model and empirical results are discussed with respect to their relationship to other developmental models and theories.

* Corresponding author. Fax: +43 3163809806.
E-mail address: michael.kickmeier@uni-graz.at (M.D. Kickmeier-Rust).

0273-2297/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.dr.2008.05.001

Introduction

A considerable amount of theoretical work and empirical research has addressed children's development of understanding distance, speed, and time (DST) concepts and their interrelations. The developmental course in this domain is not only important from the theoretical perspective of cognitive development, for instance it is an integral part of Piaget's theory, it is also important in the integration of distance, speed, and time information as they are required in many situations in everyday life; for example, to enable a child to cross a street safely before an oncoming car. Moreover, as Strauss (1998) emphasized, the knowledge about the developmental course is crucial for planning and timing a sequence of teaching these concepts.

From a traditional point of view, cognitive development is seen as a process of "discrete shifts in the nature of cognition" (Anderson & Wilkening, 1991, p. 24). Piaget (1969, 1970, 1971) was the originator of such *stage models*. He suggested that in any developmental stage, children can acquire only those competences that comply with a set of available internal operations needed to process information about the world around them. Moreover, he assumed that the development of operations occurs gradually with age and that the nature or the level of cognition would have the same quality within each developmental stage. Empirical observations have revealed that such assumptions are too strict. Apparently, children possess very basic competences about most concepts (e.g., time) almost from birth. Moreover, the development of a concept occurs rather continuous and it is not necessarily linked to the development of other concepts (cf. Case & Sowder, 1990). As a consequence, *neo-Piagetian* theories have been developed that addressed these limitations (e.g., Case, 1985; Demetriou, 1988).

More recent theories of cognitive development such as *assemblage theory* (Anderson & Wilkening, 1991) and the *overlapping wave model* (Siegler, 1996) emphasized the important role of adaptive, variable cognition. Not only one level of thinking (e.g., concrete operations) is available at one time, children are able to adapt cognitive strategies and knowledge to situational requirements, fully utilizing their potential of possible solution strategies.

Paradigms of cognitive development

A variety of previous research on children's acquisition of distance, speed, and time concepts was based on the logical-operational paradigm of Piaget. A central assumption was that the time concept is derived from the "more basic" concepts, distance and speed (Piaget, 1971). In a prototypical experimental situation, children were shown two parallel rail tracks with a toy train on each track. The trains ran with different speeds, for different times, over different distances. Children were asked to point out the train that ran faster, for a longer time, or over a longer distance. Piaget's choice task had three substantial methodological limitations: (a) Such tasks cannot directly reveal children's understanding about the relations among the three concepts. Wilkening (1981, 1982), argued that such tasks do not require attending more than one dimension in order to infer the correct answer. (b) Piagetian tasks cannot distinguish whether children's answers are based on representative thinking or on perception (Fraisse & Vautrey, 1952). (c) These tasks lack comparability between different experimental situations (Levin, 1983).

To overcome these limitations, authors such as Crépault (1979, 1980) or Montangero (1977, 1979) revised the Piagetian methodology to focus on non-metric relations. In a prototypical experiment, children were presented two toy houses and two toy cars. Children were told that one car needs a whole day to drive from one house to the other and that the other car needs a half day. Then children were asked which car needed more time from one house to the other, which car drove faster, and whether the distances both cars drove were the same. Wilkening (1982) argued that such tasks were directly derived from the Piagetian paradigm, inheriting the choice between two alternative objects or events and requiring the understanding of relational terms and comparative adjectives (e.g., more, less, same, or different).

Wilkening (1982) and Wilkening and Anderson (1982) introduced the *functional measurement* method to overcome the limitations of Piaget's experimental paradigm. Wilkening presented three animal with different speeds: a turtle, a guinea pig, and a cat. Children were told that these animals were fleeing from a dog that barks either for two, five, or eight seconds. Then children were asked how far one animal ran during a given time interval of barking. In contrast to the Piagetian paradigm, this method does not require a comparison of two objects or events, therefore demands on cognitive processing capacities are rather low. In addition, this method does not require full understanding of relational terms.

Also Matsuda (1994) presented a revised experimental paradigm, addressing some of the mentioned limitations by representing the three concepts as close as possible with respect to concreteness of representation mode and low demands on memory. In a prototypical task, children were presented three alternative toy trains, each travelling at a different speed along a track. Three toy stations were presented along the track in different distances to the trains' starting point. Children were shown that a specific train reached one of the three train stations within a given time interval indicated by a train whistle. Then children were asked which of the two other trains (a faster or a slower one) could reach a specific station.

Discrete stages vs. continuous development

To date, the course of children's acquisition of DST concepts as well as the understanding of the interrelations among these concepts is still a matter of debate. According to Piaget (1969, 1970, 1971), the development of understanding the DST system is a four-stage process: sensorimotor stage, preoperational stage, stage of concrete operations, and stage of formal operation. Along these stages, various new competences are acquired. A similar assumption came from Levin (1979). She suggested that understanding the relation of time and speed develops in two stages. In the first stage, children's judgments are based on the rule "longer time means faster," indicating the direct relation "more is more." In the second developmental stage, children attribute correctly a shorter duration to a faster object (inverse relation).

In a succeeding model, Levin (1992) proposed five stages of development. First, children are able to understand distance and speed concepts but not time values. Second, children assume direct relations between each pair of the DST system, while the respective third concept is ignored. Third, children understand the inverse relationship between time and speed, still ignoring the third concept. Forth, children begin to consider all three concepts, but coordination is not fully mature. Finally, in the fifth stage, the integration of the distance–speed–time system is completed and children can correctly derive one concept from the others. According to Levin, this five-stage process begins about the age of four and generally is completed by adolescence.

In a more recent study, Matsuda (2001) proposed a six-stage model similar to the model of Levin (1992). In a first stage, children are able to discriminate correctly between time, distance, and speed concepts. In a second stage, children are able to understand the direct relations between time and distance as well as distance and speed, but they are limited in their ability to verbalize their reasoning processes. In a third stage, children begin to understand the inverse relation between time and speed and, somewhat later, between speed and time. According to Matsuda, children still consider only two concepts, ignoring the third. Moreover, children still have difficulties to coordinate the direct and inverse relations. In the forth stage, children almost fully understand the relationships between the three concepts, still, their representations are unstable and based on two-by-two relations. In a fifth stage, children begin to consider the triadic DST system but are not fully conscious of it. Finally, in the sixth stage, children are able to refer consciously to the triadic system including direct and inverse relations.

The application of the *functional measurement* revealed results that were quite contrary to those of preceding research. Friedrich Wilkening (Krist, Fieberg, & Wilkening, 1993; Wilkening, 1981, 1982) demonstrated in his experiments that already children at the age of five have functional understanding of all three concepts. Especially the time concept is available far earlier than expected by Piagetian theories. Grounding on those experimental observations and contesting the stage models, theories of continuous development and adaptive thinking have been developed. In the framework of *information integration theory*, Anderson and Wilkening (1991) introduced the *assemblage theory*, which they considered as an "antithesis" of stage models. Assemblage theory is based on the observation that children demonstrate a specific competence in a certain situation and that this competence might not be demonstrated in another situation. A central assumption is adaptive thinking, which is the ability to adapt cognitive strategies to the requirements of a problem by selecting one of a set of available strategies. Knowledge and cognition of different levels and qualities in interacting and operating memory is seen as an "assemblage" of different competences. Development is seen as a "continual expansion, ramification, and interlocking with other abilities and knowledge" (Anderson & Wilkening, 1991, p. 26).

Robert Siegler (1996) introduced another theory of adaptive thinking, the so-called *overlapping wave model*. This model criticizes stage models for ignoring children's abilities in using multiple strategies in one age or in one stage of development. He emphasizes that children can adapt the selection of problem-solving strategies in the same task to personal preferences and task characteristics. For example, in a simple addition task, children might remember the solution while they might calculate the solution in more difficult tasks. "Rather than stepping from Strategy 1 to Strategy 2 to Strategy 3, children would be expected to use several strategies at any one time, with the frequency of use of each strategy ebbing and flowing with increasing age and expertise" (Siegler, 1995, p. 410). His model represents different levels of cognition as probability distributions (waves) over age. With certain ages, certain strategies appear with a certain probability and these distributions overlap so that at a given age a child might be able to apply more than one strategy.

The development of competence and performance

The theories of child development in the DST domain represent development in a differently fine-graded manner, from rough discrete stages to continuous and adaptive processes. They are based on empirical findings resulting from different experimental paradigms. In the present work, we introduce a formal framework for modelling learning and development. On this basis, we establish a developmental model, which can be considered as a more general case that widely includes facets and properties of existing theories and models and that potentially is able to explain the empirical results obtained under different experimental paradigms.

The basic assumption of this formal framework is that development is characterized by a continuous acquisition of finely graded, elementary competences. Such competences may be task-independent or related to one or more tasks, and they might refer to explicit knowledge (e.g., knowing arithmetic operations) and also intuitive knowledge (e.g., estimating the solution of a problem). Some previous theories described child development as a strictly linear process; in our point of view, development proceeds along different individual developmental paths. For instance, one child might first acquire the understanding of the time concept and later the understanding of the speed concept while another child might acquire these competences in opposite order. However, the developmental paths are limited by certain structural assumptions of the model. Acknowledging an adaptive, variable use of different strategies to solve problems, the introduced framework allows separating latent, unobservable competences and observable performance (i.e., mastering a task or not).

The presented framework is based on *Competence-based Knowledge Space Theory* (CbKST; cf. Albert & Lukas, 1999), which is based on structural assumption about a domain of knowledge and which is characterized by a separation of latent competences and observable performance (Doignon, 1994; Düntsch & Gediga, 1995, 1998; Korossy, 1997, 1999).

A brief introduction to CbKST

CbKST originates from KST established by Doignon and Falmagne (1985, 1999), which is a well-elaborated set-theoretic framework for addressing the relations among problems (e.g., test items). It provides a basis for structuring a domain of knowledge and for representing the knowledge based on *prerequisite relations* (see Fig. 1 for an example). While KST focuses only on performance (the behaviuor; for example, solving a test item), CbKST introduces a separation of observable performance and latent, unobservable competences, which determine the performance.

An empirically well-validated approach to CbKST was introduced by Korossy (1997, 1999); basically, the idea of the *Competence–Performance Approach* (CPA) is to assume a finite set of competences

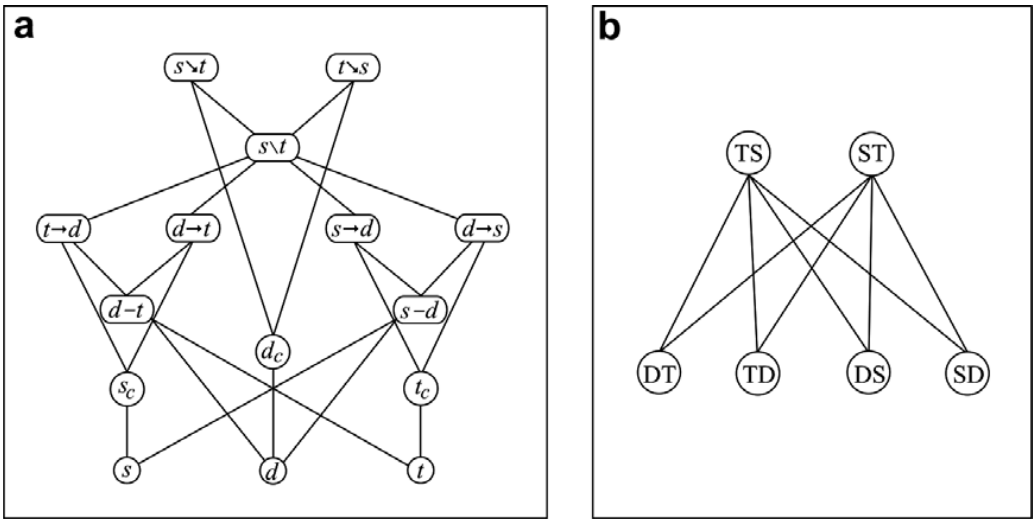

Fig. 1. Upward drawing displaying the prerequisite relations of the proposed competence (a) and performance model (b).

$C = \{a,b,c,d,\ldots\}$ and a prerequisite relation $\preceq$ between those competences. A prerequisite relation $a \preceq b$ states that competence a (e.g., to multiply two positive integers) is a prerequisite to acquire another competence b (e.g., to divide two positive integers). If a person possesses competence b, we can assume that the person also posses competence a. To account for the fact that more than one set of competences can be a prerequisite for another competence (e.g., competence a or competence b are a prerequisite for acquiring competence c), *prerequisite functions* have been introduced, relying on and/or-type relations. A person's *competence state* is described by a subset of competences of C, for example $\{a,b\}$. Due to the prerequisite relations between the competences, not all subsets of competences are possible competence states. To give an example, imagine four competences from the domain of basic algebra, the abilities to add, subtract, multiply, and divide numbers. Given four competences, the set of all possible knowledge states is 2^4. If we assume that the competences to add, subtract, and multiply numbers are prerequisites for the competence to divide numbers, not all of the 16 competence states are plausible. For example, it is highly unlikely that a child has the competence to divide numbers but not to add numbers. The collection of possible competence states corresponding to a prerequisite relation is called *competence structure*. Such competence structure also singles out different learning paths for moving from the naïve state {} (having no competences of a domain) to the state of possessing all of a domain's competences C.

So far, the structural model focuses on latent, unobservable competences. By utilizing *interpretation* and *representation functions* the latent competences are mapped to a set of tasks (or test items) $Q = \{p,q,r,s,\ldots\}$ relevant for a given domain. Through the aforementioned example, Q might include a number of addition, subtraction, multiplication, and division tasks. The interpretation function assigns a set of competences required to solve a task to each of the problems in Q. Vice versa, by utilizing a *representation function*, a set of problems is assigned to each competence state, which can be mastered in this state. This assignment induces a *performance structure*, which is the collection of all possible *performance states*. Due to these functions, latent competences and observable performance can be separated and no one-to-one mapping is required. Moreover, learning or development is not seen as a linear course; equal for all children, rather, development follows one of a set of individual learning or developmental paths.

This approach entails several advantages. The separation of performance and underlying latent competences enables a model of development independent from concrete tasks. The tasks of different test instruments (e.g., different school tests) or of different experimental paradigms are mapped to a single competence structure. This enables a continuous monitoring of development, a comparison of performance across different instruments and experimental paradigms, and it supports educational measures (e.g., it allows determining what competences a student must acquire to reach a certain educational goal).

A competence–performance model for the DST system

CbKST provides a general framework for modelling learning and developmental processes. The aim of the present work is to exemplify this approach on a model of the development of understanding the DST system.

We first subsume elementary competences required for understanding the DST system. Next, we derive a prerequisite relation (a prerequisite functions is not required in the present model) for those competences and an associated competence model. Finally, we assign competences to tasks, prototypically the task types of the experimental paradigm of Matsuda (1994), inducing a prerequisite relation between the task types and therefore inducing a performance structure.

Latent competences

On the basis of previous studies, we summarized 15 elementary competences (Table 1), which we hypothesize to be required to understand the DST concepts and the interrelations between them.

The basis of the developmental course in this domain is children's understanding of DST values as distinct entities (Table 1, competences s, d, t). Matsuda (1994) argued that the discrimination among these three concepts is rather difficult for 4-year olds. A number of studies showed that this ability is present in children at the age of five (e.g., Matsuda, 2001; Wilkening, 1981). We assume that mastering DST judgments requires the detection of the constant value of the third concept, as related to each two-by-two relation (Table 1, competences s_c, d_c, t_c). Even if most experimental paradigms cannot distinguish whether a constant value is detected or ignored (cf. Wilkening, 1982), the literature suggests that it is a fundamental competence to achieve full understanding of this domain. In a next step, we consider the competence of understanding the interrelations between DST concepts. As suggested by a variety of previous studies (e.g., Matsuda, 1994, 2001; Squire & Bryant, 2003; Wilkening, 1981, 1982), we have to distinguish between three possible two-by-two relations: (a) The direct relation between time and distance, (b) the direct relation between speed and distance, and (c) the inverse relation between time and speed (Table 1, competences d–t, s–d, $s\backslash t$). The direct relations are understood clearly before the inverse relations (Levin, 1979). Finally, we have to consider the ability to apply correctly the understanding of the interrelations between the concepts. As emphasized by Matsuda (2001), even if children are aware of the correct relationship in principle, the application of this knowledge might be unstable and the application of direct and inverse relations might be conflicting. Thus, we distinguish the competences to make correct inferences from one concept to another (Table 1, competences $d \to t$, $t \to d$, $d \to s$, $s \to d$, $t \searrow s$, $s \searrow t$).

In order to establish a prerequisite relation (Fig. 1a), we assume that the most basic competences refer to the understanding of DST values as distinct entities. For example, the detection of the direct

Table 1
Elementary competences for the understanding of DST concepts and their interrelations

Symbol	Elementary competences
t	Understanding of time values
d	Understanding of distance values
s	Understanding of speed values
t_c	Detection time as constant
d_c	Detecting distance as constant
s_c	Detecting speed as constant
d–t	Detecting the direct relation between distance and time
$d \to t$	Inference of longer distance from longer time
$t \to d$	Inference of longer time from longer distance
s–d	Detecting the direct relation between speed and distance
$s \to d$	Inference of higher speed from longer distance
$d \to s$	Inference of longer distance from higher speed
$s\backslash t$	Detecting the inverse relation between speed and time
$s \searrow t$	Inference of higher speed from shorter time
$t \searrow s$	Inference of longer time from lower speed

relation between speed and distance requires at least the understanding of speed and distance values. To infer correctly one concept from another, we propose that the detection of the constant value of the third variable is required. According to previous findings (e.g., Acredolo, Adams, & Schmid, 1984; Crépault, 1980; Levin, 1992; Matsuda, 2001) direct-relation tasks are understood before inverse-relation tasks. Studies regarding the inverse relation in division (Correa, Nunes, & Bryant, 1998; Squire & Bryant, 2003) found that 5-year olds could not master division problems; however, 7-year olds could master such problems. From mastering the inverse relation, we can surmise that a child also understands the direct relations.

The competence structure is determined by the family of subsets of competences that are consistent with the prerequisite relation (Fig. 1a); it consists of 129 competence states, which determine 241.440 individual developmental paths, following 16 competence-related developmental steps.

Task performance

A major advantage of CbKST is that no one-to-one mapping between competences and tasks is required and, therefore, a variety of task types can be assigned to the competence model. In the present work, we apply the task types introduced by the experimental paradigm of Matsuda (1994) to investigate the proposed competence model. In this paradigm, children are presented three alternative toy trains, each travelling at a different speed on a single track. Three toy stations are presented at different distances from the trains' starting point. Children are shown that a specific train reaches one of the three train stations within a specific time interval, indicated by a train whistle. Children are asked to make several judgments; for example, which of the trains can reach a specific station within a specific time interval. Matsuda realized six experimental task types (Table 2). Task type *DS* requires the inference of distance from speed at constant time, *SD* requires the inference of speed from distance at constant time, *DT* requires the inference of distance from time at constant speed, *TD* requires the inference of time from distance at constant speed, *ST* requires the inference of speed from time at constant distance, and *TS* requires the inference of time from speed at constant distance.

To establish a prerequisite relation for the six tasks types, we assigned sets of competences to each task type, which are necessary to master the corresponding tasks. This mapping, formalized as interpretation function in Table 3, was based on the analysis of the cognitive requirements for each task type. Through set inclusion, the structure on the competences induces a structure on the task types (Fig. 1b). For instance, when the set of competences assigned to one task is a subset of those assigned to another task, the first task is naturally a "prerequisite" for the second (cf. Table 3 and Fig. 1b). The corresponding performance structure P includes 19 performance states: $P = \{\{\}, \{DS\}, \{SD\}, \{DT\}, \{TD\}, \{DT, TD\}, \{DT, SD\}, \{DT, DS\}, \{TD, SD\}, \{TD, DS\}, \{SD, DS\}, \{DT, TD, SD\}, \{DT, TD, DS\}, \{DT, SD, DS\}, \{TD, SD, DS\}, \{DT, TD, SD, DS\}, \{DT, TD, SD, DS, ST\}, \{DT, TD, SD, DS, TS\}, \{DT, TD, SD, DS, ST, TS\}\}$.

Table 2
Task types of Investigations 1 and 2

Symbol	Task type principles
DT	(1) Inference of longer distance from longer time at constant speed (2) Inference of shorter distance from shorter time at constant speed
TD	(1) Inference of longer time from longer distance at constant speed (2) Inference of shorter time from shorter distance at constant speed
SD	(1) Inference of higher speed from longer distance at constant time (2) Inference of lower speed from shorter distance at constant time
DS	(1) Inference of more distance from more speed at constant time (2) Inference of less distance from less speed at constant time
ST	(1) Inference of more speed from less time at constant distance (2) Inference of less speed from more time at constant distance
TS	(1) Inference of more time from less speed at constant distance (2) Inference of less time from more speed at constant distance

Table 3
Interpretation function: mapping between task types and competence states

Task types	Required competences
DT	$\{s, d, t, s_c, d–t, d \rightarrow t\}$
TD	$\{s, d, t, s_c, d–t, t \rightarrow d\}$
SD	$\{s, d, t, t_c, s–d, s \rightarrow d\}$
DS	$\{s, d, t, t_c, s–d, d \rightarrow s\}$
ST	$\{s, d, t, s_c, d_c, t_c, d–t, d \rightarrow t, t \rightarrow d, s–d, d \rightarrow s, s \rightarrow d, s\backslash t, s \searrow t\}$
TS	$\{s, d, t, s_c, d_c, t_c, d–t, d \rightarrow t, t \rightarrow d, s–d, d \rightarrow s, s \rightarrow d, s\backslash t, t \searrow s\}$

Note: For a definition of competences, refer to Table 1.

It is very likely that a certain portion of empirical answer patterns deviates from the hypothesized ones. As an example, a child might master an inverse-relation task but fail in a direct-relation task. There are several options to explain such deviations, for example, careless errors or lucky guesses. Another explanation that is frequently referenced in literature is systematic misconceptions.

Overgeneralization

Misconceptions such as applying a certain rule correctly in a situation where it is not suitable, are a natural part of developmental processes. Overgeneralization is a prominent type of misconception, which refers to a generalization of concepts or rules to situations where an application is not suitable. Previous authors (e.g., Levin, 1979, 1992) proposed two common misconceptions: (a) if children are not already capable of understanding the inverse relation between time and speed, they might overgeneralize from direct to inverse relations. In contrast, (b) if children understand the inverse relation between speed and time but this understanding is not stable, these children might overgeneralize from inverse to direct relations. However, it is a persistent problem to differentiate between systematic misconceptions and careless errors or guessing. The presented framework offers a basis to define misconceptions formally and to investigate such assumptions empirically. We focus on overgeneralization from inverse to direct relations, since overgeneralization from direct to inverse relations is already captured by the introduced competence–performance model. The original performance model, subsequently, is complemented with the performance states arising from the six definitions of overgeneralization.

Complete overgeneralization. The most severe definition of overgeneralization states that when a child understands the inverse rule "more is less" (task types *ST* and *TS*), the child would apply this rules erroneously to all task types requiring the direct rule "more is more" (task types *DT*, *TD*, *SD*, and *DS*): $P_a = P \cup \{\{ST\},\{TS\}, \{ST,TS\}\}$.

Complete overgeneralization by factors. Overgeneralization within one factor may occur from the factor speed in inverse-relation tasks (*ST* and *TS*) to speed in direct-relation tasks, leading to wrong answers in task types *SD* and *DS* and, equivalently, from the factor time in the inverse-relation tasks (*ST* and *TS*) to the factor time in the direct-relation tasks, leading to wrong answers in task types *DT* and TD: $P_b = P_a \cup \{\{DS,SD,TS\},\{DT,TD,TS\},\{DT,ST,TD\},\{DS,SD,ST\}\}$.

Partial overgeneralization by factors. Extending complete overgeneralization by factors, this definition states that overgeneralization is given if only one task type is affected: $P_c = P_b \cup \{\{DS,TS\}, \{SD,TS\},\{DT,TS\},\{TD,TS\},\{DS,ST\},\{SD,ST\},\{DT,ST\},\{TD,ST\},\{SD,DT,TD,TS\},\{DS,DT,TD,TS\},\{SD,DS,DT, TS\}, \{SD,DS,TD,TS\},\{SD,DT,TD,ST\},\{DS,DT,TD,ST\},\{SD,DS,DT,ST\},\{SD,DS,TD,ST\},\{SD,DS,ST,TS\},\{DT,TD,ST,TS\}, \{SD,DT,TD,T,S\},\{DS,DT,TD,ST,TS\},\{SD,DS,DT,ST,TS\},\{SD,DS,TD,ST,TS\}\}$.

Partial overgeneralization by number of errors. Another type of overgeneralization occurs from inverse to direct relations when failing in at least four (P_{d4}), three (P_{d3}), two (P_{d2}), or one (P_{d1}) direct-relation tasks. Definition P_{d4} is identical to complete overgeneralization P_a: $P_{d4} = P_a = P \cup \{\{ST\},\{TS\},\{ST,TS\}\}$; $P_{d3} = P_{d4} \cup \{\{TD,TS\},\{DT,TS\},\{DS,TS\},\{SD,TS\},\{TD,ST\},\{DT,ST\},\{DS,ST\},\{SD,ST\},\{\{TD,ST,TS\},\{DT,ST,$

$TS\}, \{DS, ST, TS\}, \{SD, ST, TS\}\}$; $P_{d2} = P_{d3} \cup \{\{DT, TD, ST, TS\}, \{DS, TD, ST, TS\}, \{DS, DT, ST, TS\}, \{SD, TD, ST, TS\}, \{SD, DT, ST, S\}, \{SD, DS, ST, TS\}, \{DT, TD, ST\}, \{DS, TD, ST\}, \{DS, DT, ST\}, \{SD, TD, ST\}, \{SD, DT, ST\}, \{SD, DS, ST\}, \{DT, TD, TS\}, \{DS, TD, TS\}, \{DS, DT, TS\}, \{SD, TD, TS\}, \{SD, DT, TS\}, \{SD, DS, TS\}\}$; finally, definition P_{d1}, failing in least one of the direct-relations tasks, results in the power set, the set of all possible combinations of performance states $P_{d1} = 2^{|Q|}$, where Q denotes the set of all tasks types. This definition includes 64 performance states.

To investigate the occurrence of overgeneralization, the question is whether the proposed performance states arising from these definitions can contribute to the explanation of empirical answer patterns and consequently the goodness of fit of model and data. The definition P_{d1} is trivial; it includes all possible performance states and therefore includes all possible answer patterns. Definition P_{d4} is identical to definition P_a. Thus, in the empirical investigations we focus on definitions P_a, P_b, P_c, P_{d2}, and P_{d3}.

Empirical investigations

In a first step, we aim to evaluate the proposed, theoretically obtained competence–performance model and at investigating whether the model covers a significant portion of the empirical answer patterns as hypothesized. In a second step, we aim to analyze the occurrence of answer patterns predicted by the definitions of overgeneralization. These analyses were carried with the data of two empirical investigations, both based on the research paradigm introduced by Matsuda (1994), as described above. Investigation 1 was conducted with Japanese children and published by Matsuda (2001). The data of this investigation were re-analyzed according to the current hypotheses. Investigation 2 was conducted with Austrian children, applying a comparable method. The experimental design, the utilized task types, and the experimental procedure were identical to those in Investigation 1. Instead of real toy trains, participants were presented with computer animations.

Method

Participant

In total, 222 Japanese children participated in Investigation 1. The children's age ranged from four to eleven; 54% of the children were male and 46% female. The children attended a kindergarten or an elementary school. Five children at the age of ten and all of the 24 children at the age of eleven had already learned about speed in fifth-grade math. In Investigation 2, 42 Austrian children took part. They attended a kindergarten, an elementary school, and a grammar school in Graz, Austria. The ages ranged from 4 to 10 years, and half of the participants were male, the other half was female. In addition, 22 adults participated in Investigation 1 and six participated in Investigation 2 to control for final state of development. As expected, all adult participants mastered all tasks, these data were not analyzed in this context.

Apparatus

In Investigation 1, the participants were presented real toy trains on a 149 cm track. One of three 10 cm-long trains (a green, a red, and a blue one) could run on these rail tracks at the same time. The speed of each train was fixed to 25 cm/s for the green train, 10 cm/s for the red train, and 4 cm/s for the blue train. Alongside the rails tracks, three stations (10 cm wide and 4 cm high) were placed in distances of 20 cm (white station), 50 cm (gray station), and 125 cm (black station) from the trains' starting point. To start a train and to determine the time interval the train was running, the participants were presented a switch box. The train ran for 2 s when a red button was pushed, 5 s when a blue button was pushed, and 12.5 s when a yellow button was pushed. During each run, the participants heard a train whistle sound.

Unlike Investigation 1, Investigation 2 was presented onscreen using a software named *wzgANI*[1] (Kickmeier-Rust, 2002). As shown in Fig. 2, this tool presented three rail tracks with three different train stations and one train on each track. Initially, the trains were presented on the left end of the rail tracks.

[1] The software tool *wzgANI* (in German language) is freely available for Windows platforms from the authors.

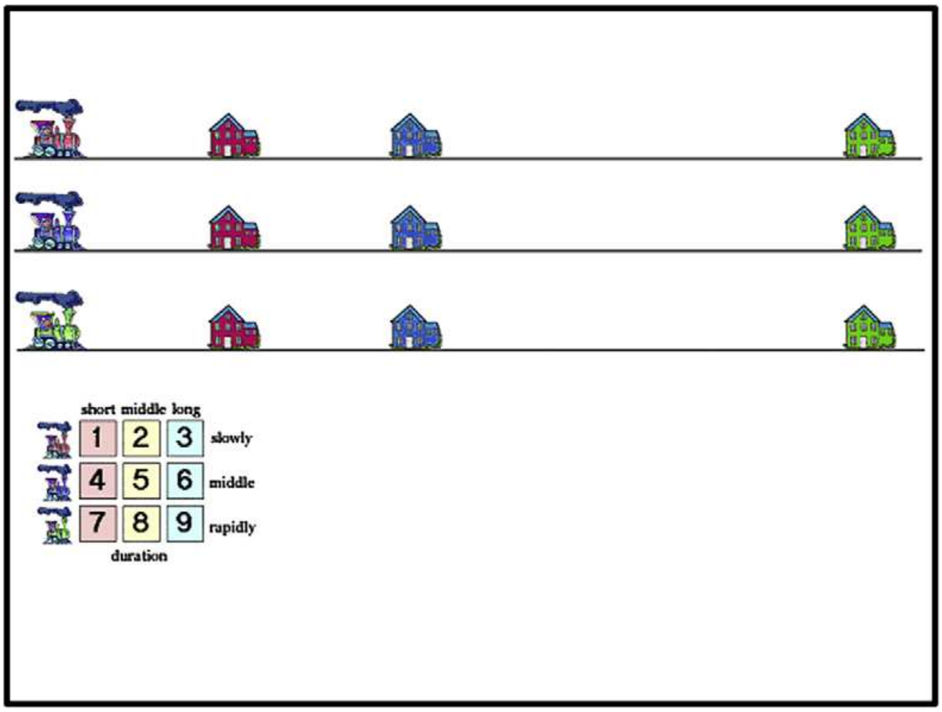

Fig. 2. Screenshot of the experimental setup of Investigation 2 using the software wzgANI.

The trains ran with a fixed speed of 27.2 pixels/s (red train on the top), 68 pixels/s (blue train in the middle location), and 17 pixels/s (green train at the bottom). Along the rails, three small stations were displayed in distances of 136 (red station), 340 (blue station), and 850 pixels (green station). Below the rail tracks, a switch panel with nine buttons was presented that allowed controlling the trains. The children could choose which of the three trains should run for one of three time intervals (2, 5, or 12.5 s). During each run, the participants heard a train sound. As an example, when the button labelled "6" was clicked, the blue train ran with a speed of 68 pixels/s for 12.5 s, reaching the third (green) station in a distance of 850 pixels.

Experimental design and tasks

The experimental design included six task types and was identical in both investigations. The task types are symbolized by capital letters *D* for distance, *S* for speed, and *T* for time. For instance, in task type *DT*, relative judgments about distance were required when time varied with constant speed of the train. Each task type included two similar tasks. The experimental tasks are summarized in Table 2. Task types *DT*, *TD*, *SD*, and *DS* referred to the direct relations between distance and time and distance and speed, respectively. Task types *ST* and *TS* referred to the inverse relation between speed and time. The tasks were presented in two sessions and in one of six orders to counterbalance order effects. The second session was performed about a week after the first session. Tasks containing the same combination of factors (i.e., *DT* and *TD*, *SD* and *DS*, and *ST* and *TS*) were carried out on separate days. As dependent variables, the participants' responses (the selected button, station, or train) were recorded.

Procedure

Investigation 1 was conducted at the Naruto University of Education, Japan. The participants were examined in individual sessions. In the first experimental session, the participants were required to discriminate between the three values of distance (20, 50, and 125 cm) and among the three values of time (2, 5, and 12.5 s), respectively. All participants were able to discriminate the concepts; consequently, none of them had to be excluded from data analysis. Subsequently, three task types were performed in each session with two tasks for each type.

First, the experimenter demonstrated the equipment (the three trains, the three stations, and the three buttons to trigger the three different time intervals) and the principles of a task type. By the example of task type *TD*, the experimenter set the red train (10 cm/s) on the rail tracks and said that this red train would always run with the same speed from the starting point to one of the stations, accompanied by a whistle sound. The experimenter demonstrated that, in this example, the red train reaches the second station by pushing the blue button (5 s). Then the participants were asked which button (the red or yellow) should be pushed to let the same train run to the farthest (black) station. In a next step, the participants were asked why they thought that their answer is correct. Finally, the experimenter demonstrated the result of the related answer with the train and provided the participants with verbal feedback. The second task of this type, equivalently, was to answer which button should be pushed to reach the nearest (white) station. The experimenter asked whether the participant could correctly recognize the constant variable (e.g., the variable speed in task type *TD*) throughout the three runs, which could be stated by all participants. Finally, children were asked for a justification of their selections. These justifications were primarily intended to prevent children from guessing and were not considered in data analyses. The other five task types were conducted equivalently by interchanging the attributes of SDT values. A major premise was to avoid influences of verbal demands—as far as possible—to ensure that observable behaviour was primarily determined by judgments on physical knowledge.

Investigation 2 was conducted equivalently in separate and quiet rooms in the kindergarten and schools, using a laptop computer with a screen resolution of 1024×768 pixels that displayed the computer animations.

Results

Performance

As aforementioned, Investigation 1 was a re-analyses of the data recorded by Matsuda (2001). In a first step, she analyzed the frequencies of correct responses in each task for each age group. A task was considered as mastered when responses for both of a type's tasks were correct. These results are summarized briefly in Fig. 3.

Already at the age of four, a considerable understanding of the direct relation "more is more" between distance and time (task types *DT* and *TD*) and between distance and speed (task types *SD* and *DS*) was found. The understanding of the inverse relation "more is less" between speed and time (task types *ST* and *TS*) is developed later (Fig. 3b). Between 5 and 6 years and between 7 and 8 years of age, Matsuda found a distinct increase of correct answers for these tasks (Fig. 3a). Generally, the understanding of the inverse relation develops about 3 years later than understanding the direct relations. This confirms previous research that found that children recognize the direct relations before they recognize the inverse relation. While Crépault (1980) argued that the direct relation between speed and distance has consolidated before the direct relation between time and distance throughout all age groups (6–11 years), Matsuda found such effect only at 4-year olds (Fig. 3a). In this age group, the mean rate of correct answers in *SD* tasks was clearly lower than other direct-relations tasks and also the *DS* tasks.

Investigation 2 revealed, as expected, similar results. The mean rates of correct answers are summarized in Fig. 3. An age × task (8×6) analysis of variance (ANOVA) was computed based on the mean rate of correct answers for each task type. Both the main effects as well as the interaction of independent variables were significant ($F(5,42) = 18.04$, $p < .01$; $F(7,40) = 5.082$, $p < .01$; $F(7,40) = 1.765$, $p < .01$).

As in Investigation 1, already at the age of 4 years, the understanding of the direct relation "more is more" between distance and time (task types *DT* and *TD*) and between distance and speed (task types *DS* and *SD*) are well developed. The understanding of the inverse relationship "more is less" between speed and time (tasks *ST* and *TS*) develops about 3 years later than understanding the direct relation (Fig. 3b). As in Investigation 1, in children between 5 and 6 years of age we found a clear increase of the rate of correct answers for these tasks. Task type *DS* is a distinct outlier. None of the 6-year olds of Investigation 2 mastered both tasks of this type. In contrast to Investigation 1, here we found some evidence for Crépault's (1980) finding that the direct relation between speed and distance has consolidated before the direct relation between time and distance (Fig. 3a). Finally, the results of both investigations did not confirm Montangero's (1979) argument that children understand the two-by-two

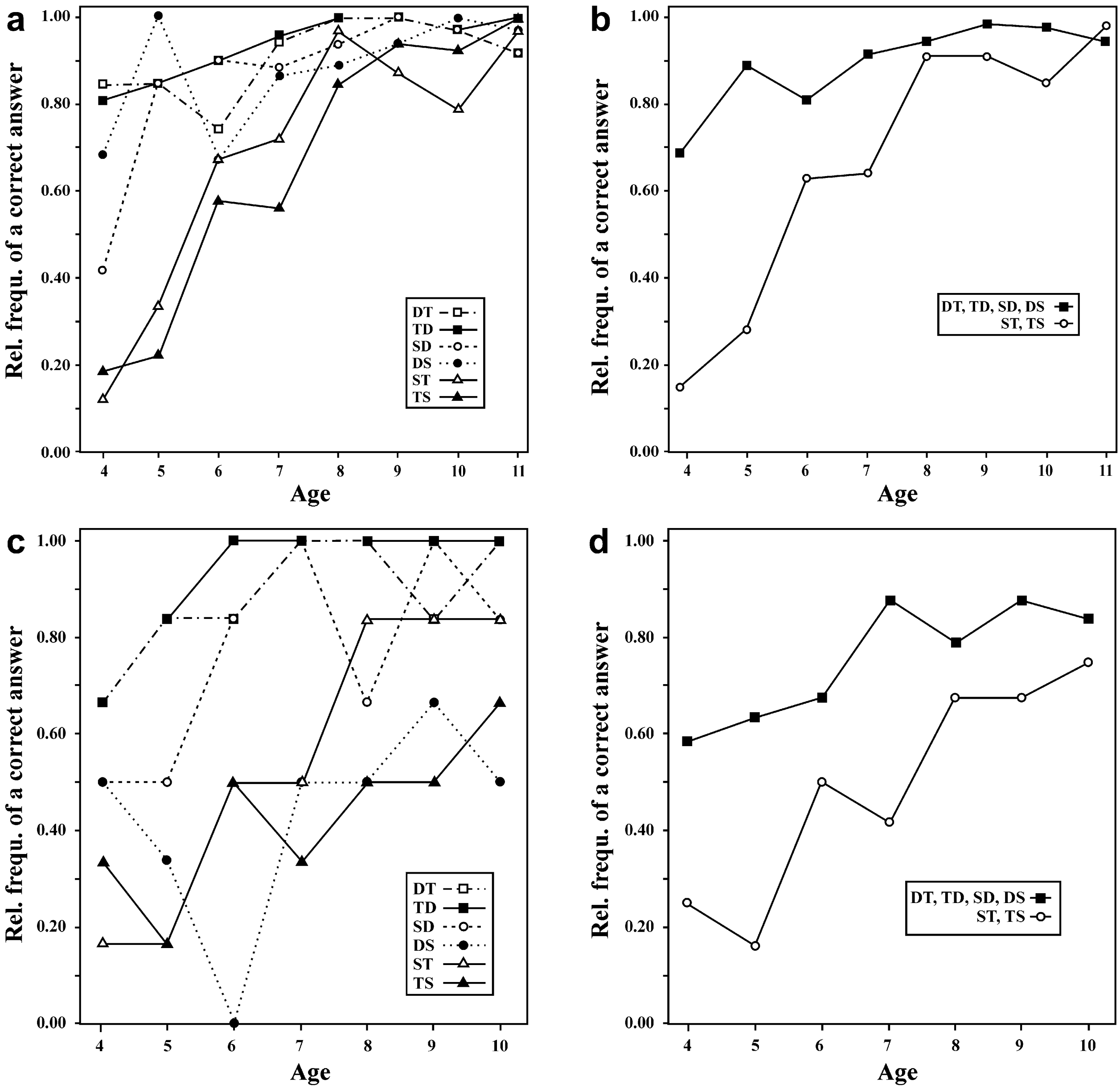

Fig. 3. (a) The relative frequencies of correct answers for each of the six experimental task types for each age group and (b) the relative frequencies of correct answers summarized for tasks covering direct (*DS*, *SD*, *DT*, and *TD*) and inverse relations (*ST* and *TS*) in Investigation 1. (c) The relative frequencies of correct answers for each of the six experimental task types for each age group and (d) the relative frequencies of correct answers summarized for tasks covering direct and inverse relations in Investigation 2.

relations first in one direction and then in the other. As in Acredolo et al. (1984), no evidence for such asymmetry was found (Table 4).

The competence–performance model

To analyze the proposed performance model *P* (Fig. 1b), we first excluded trivial answer patterns from the data.[2] In both investigations, each of the participants was able to master at least one of the task

[2] Excluding trivial answer patterns (i.e., children that mastered all tasks or none at all) had two reasons. First, these patterns do not offer any information about the domain structure. Second, because the empty set {} and the whole set {Q} are parts of a performance structure, adding corresponding answer patterns (i.e., utilizing floor and ceiling effects) to a set of empirical data could arbitrarily increase the fit of model and data.

Table 4
Answer patterns and their frequencies in each age group of both investigations

Tasks types						Age								
DS	*SD*	*DT*	*TD*	*ST*	*TS*	4	5	6	7	8	9	10	11	
Answer patterns covered by the proposed performance model														
–	–	–	–	–	–									–/–
–	–	X	–	–	–	2/–								2/–
–		–	X	–	–		–/1	1/–						1/1
–	X	–	–	–	–			1/–						1/–
X	–	–	–	–	–	1/1								1/1
–	–	X	X	–	–	2/–	–/1	–/1		–/1				2/3
–	X	X	–	–	–		–/1							–/1
X	–	X	–	–	–	2/–	2/–							4/–
–	X	–	X	–	–	1/–			1/–					2/–
X	–	–	X	–	–	1/1	1/–		1/–					3/–
X	X	–	–	–	–		1/–							1/–
–	X	X	X	–	–	1/1		2/1	–/2					3/4
X	–	X	X	–	–	2/–	–/1	1/–	2/–	1/–		–/1		6/2
X	X	X	–	–	–							–/1		–/1
X	X	–	X	–	–									–/–
X	X	X	X	–	–	8/–	12/1	4/–	6/–		1/–	1/–		32/1
X	X	X	X	X	–		3/–	3/–	8/2	2/–	–/2			16/4
X	X	X	X	–	X		1/–	1/–	3/–		1/1	2/–		8/1
X	X	X	X	X	X	1/–	3/–	8/–	17/1	20/3	13/1	18/2	23/–	103/7
						21/3	23/5	21/2	38/5	23/4	15/4	22/3	23/–	186/26
Deviating answer patterns														
–	–	–	X	–	X								1/–	1/–
–	–	X	X	–	X	2/1								2/1
X	–	–	X	–	X	1/–								1/–
X	–	X	X	–	X		1/–					1/–		2/–
X	X	X	–	–	X	–/1								–/1
–	X	X	X	–	X			–/1	–/1					–/2
–	–	X	X	X	–					–/1				–/1
X	–	X	X	X	–	1/–								1/–
–	X	X	X	X	–	–/1		1/1	2/–	1/1	–/1	–/1		4/5
X	X	X	–	X	–		1/–							1/–
X	X	–	X	X	–		1/–							1/–
X	X	–	–	X	X			1/–						1/–
X	–	X	–	X	X				1/–					1/–
–	X	X	–	X	X			1/–						1/–
X	–	–	X	X	X				1/–				1/–	2/–
–	X	–	X	X	X			3/1			–/1			3/2
X	–	X	X	X	X	1/–		1/–		1/–				3/–
–	X	X	X	X	X		–/1	1/1	3/–	2/–	1/–	–/2		7/4
X	X	X	–	X	X				1/–					1/–
X	X	–	X	X	X		1/–	2/–				1/–		4/–
						5/3	4/1	10/4	8/1	4/2	1/2	2/3	2/–	36/16
					Total sum	26/6	27/6	31/6	46/6	27/6	16/6	24/6	25/–	222/42

Note. The X denotes mastered tasks types. Numbers before slashes denote the results of Investigation 1, the numbers after the slashes the results of Investigation 2. As an example, two 4-year olds and two 5-year olds from Investigation 1 and no child of Investigation 2 could master both *DT* and *DS* task types but only one or none of the *TD*, *SD*, *ST*, and *TS* task types (8th row from above).

types. The data sets of 103 children in Investigation 1 and seven children in Investigation 2 who were able to master all six of the task types were rejected from data analysis.

The analysis of the remaining 119 answer patterns in Investigation 1 resulted in a reasonably good fit of answer patterns and performance models. Predicting only 19 (29.69%) of the possible 64 performance states, the model covered 83 (69.75%) of the empirical answer patterns (Table 4, upper part

Table 5
Goodness-of-fit of basic performance model and its extensions by definitions of overgeneralization

		Basic model	Models of overgeneralization					
		P[a]	P_a[e]	P_b	P_c	P_{d3}	P_{d2}	P_{d1}[f]
Size[b]		19 (29.69)	22 (34.38)	26 (40.63)	48 (75.00)	34 (53.13)	52 (81.25)	64 (100.00)
Investigation 1								
Avg. distance[c]		0.37	0.37	0.34	0.07	0.29	0.29	0.00
Distances[d]	0	83 (69.75)	83 (69.75)	85 (71.43)	111 (93.28)	84 (70.59)	95 (79.83)	119 (100.0)
	1	28 (23.53)	28 (23.53)	27 (22.69)	8 (6.72)	35 (29.41)	24 (20.17)	0
	2	8 (6.72)	8 (6.72)	7 (5.88)	0	0	0	0
	3	0	0	0	0	0	0	0
Investigation 2								
Avg. distance[c]		0.51	0.51	0.46	0.06	0.46	0.34	0.00
Distances[d]	0	19 (54.29)	19 (54.29)	21 (60.00)	33 (94.29)	19 (54.29)	23 (65.71)	35 (100.00)
	1	14 (40.00)	14 (40.00)	12 (34.29)	2 (5.71)	16 (45.71)	12 (34.29)	0
	2	2 (5.71)	2 (5.71)	2 (5.71)	0	0	0	0
	3	0	0	0	0	0	0	0

[a] Basic performance model without overgeneralization.
[b] Number of performance states in the performance structure. The numbers in parentheses denote the percentage of the number of performance states in relation to the maximum number of states, which is 64.
[c] Average minimal symmetric distance.
[d] Number of patterns with a minimal symmetric distance of 0–3 (the maximum distance for six items is 3). The numbers in parentheses denote the percentage of number of answer patterns in relation to the maximum number of answer patterns (excluding trivial patterns) in the data sets of Investigations 1 (in total 119 patterns) and 2 (in total 35 patterns).
[e] Please note that P_{d1} is equivalent to complete overgeneralization P_a.
[f] Power set of all possible states.

excluding the 103 trivial answer patterns; Table 5). This is a difference of 40.06%, which is clearly larger than expected. To analyze the deviations between answer patterns and the performance structure, we used the software tool *DI*, version 2.4.5 by Held, revised by Hockemeyer (2001) from the KST Tools[3] package. *DI* computes the minimal average distances between empirical answer patterns and performance states[4]. The average minimal distance for the performance structure was .37 (the maximum distance for six items is 3.00), 28 answer patterns deviated with a distance of 1, further 8 answer patterns with a distance of 2, and none with a distance of 3 (Table 5).

For more in-depth analyses, we used a graphical method. We plotted the relative frequencies of 0-distance patterns (answer patterns with a minimal symmetric distance of 0) and the relative size (number of performance states divided by the 64 possible states) of the performance model *P*. If the model covers the empirical answer patterns as hypothesized, it should be located clearly above the diagonal with a slope of 1. This diagonal indicates the increase of fit by consecutively accumulating performance states randomly selected from the set of in total 64 possible states. As shown in Fig. 4a, the proposed performance model *P* was clearly located above the diagonal. A binomial test yielded that the basic performance model *P* covered significantly more answer patterns than expected by chance ($p < .001$); by randomly accumulating 19 performance states (29.69%), one would expect to cover 35 (29.41%) instead of 83 (69.75%) of the 119 empirical answer patterns. Hence, the results provide some evidence for the validity of the proposed performance model *P*.

The analysis of the remaining 35 answer patterns in Investigation 2 revealed similar results. Predicting 19 (29.69%) of the possible 64 performance states, the model covered 19 (54.29%) of the empir-

[3] The KST tools software package is freely accessible via the software portal *ePsyt* (http://css.uni-graz.at/epsyt).
[4] The distance describes the degree of congruence between an answer pattern and a performance state. If an answer pattern and a performance state match, the distance is 0. The more the differences occur between the two sets, the larger the distance. The minimal distance describes the smallest possible difference between an answer pattern and the entire set of performance states. As an example, given the answer pattern {a, b, c} and two performance states {a, b, e} and {c, d, e}, the distance of answer pattern and performance states are 1 in the first case and 2 in the second case. Consequently, the minimal distance in this example is 1. The average minimal distance (the average minimal symmetric set difference; cf. Garnier & Taylor, 1992) is a descriptive value providing the mean of the minimal distances of all answer patterns to a given set of performance states (i.e., a performance structure).

ical answer patterns (Table 4, upper part excluding the seven trivial answer patterns; Table 5). This is a difference of 24.60%, which is clearly larger than expected. The average minimal symmetric distance, computed with *DI*, for the performance structure was .51 (the maximum distance for six items is 3.00), 14 answer patterns deviated with a distance of 1, further 2 answer patterns with a distance of 2, and none with a distance of 3 (Table 5).

As in Investigation 1, for more in-depth analyses we plotted the relative frequencies of 0-distance patterns and the relative size of the performance model *P*. As shown in Fig. 4b, the proposed performance model *P* was clearly located above the diagonal. A binomial test yielded that the basic performance model *P* covered significantly more answer patterns than expected chance ($p < .002$); by randomly accumulating 19 performance states (29.69%) one would expect to cover 10 (28.57%) instead of 19 (54.29%) of the 35 empirical answer patterns.

The results of both investigations provide some evidence for the validity of the proposed performance model *P*. Still, the question regarding the reasons for the 30% of deviating answers patterns in Investigation 1 and the 45% in Investigation 2 remains. A hypothesis that has to be taken into account and that frequently can be found in the literature (e.g., Levin, 1979, 1992; Matsuda, 1994, 2001) is overgeneralization.

Overgeneralization

If overgeneralization occurs, occurs, one would expect a decrease in the rate of correct answers in the direct-relation task types (*DS*, *SD*, *DT*, and *TD*) when the understanding of the inverse relation develops. As shown in Fig. 2a and b, the mean results of Investigation 1 slightly indicate such effect. We found a decrease of correct answers in *DS* and *DT* tasks between the ages of five and six, accompanied by a clear increase of correct answers in tasks *ST* and *TS*. Also, the results of Investigation 2 slightly indicate such an effect (Fig. 2c) with a distinct decrease of *DS* task performance from 5 to 6 years and of *SD* task performance from 7 to 8 years, both accompanied with an increase in *ST* and *TS* task performance.

To analyze whether one (or more) of the suggested definitions of overgeneralization from inverse to direct relations can cover a significantly larger number of answer patterns than expected by chance, we extended the basic performance model by adding the performance states that arise from these definitions. Starting from the basic model *P*, with each additional performance state, a corresponding percentage of the remaining answer patterns should be covered. In Fig. 4 this assumption is indicated by the dotted lines. If a performance structure offers a systematic explanation for the data, it should be

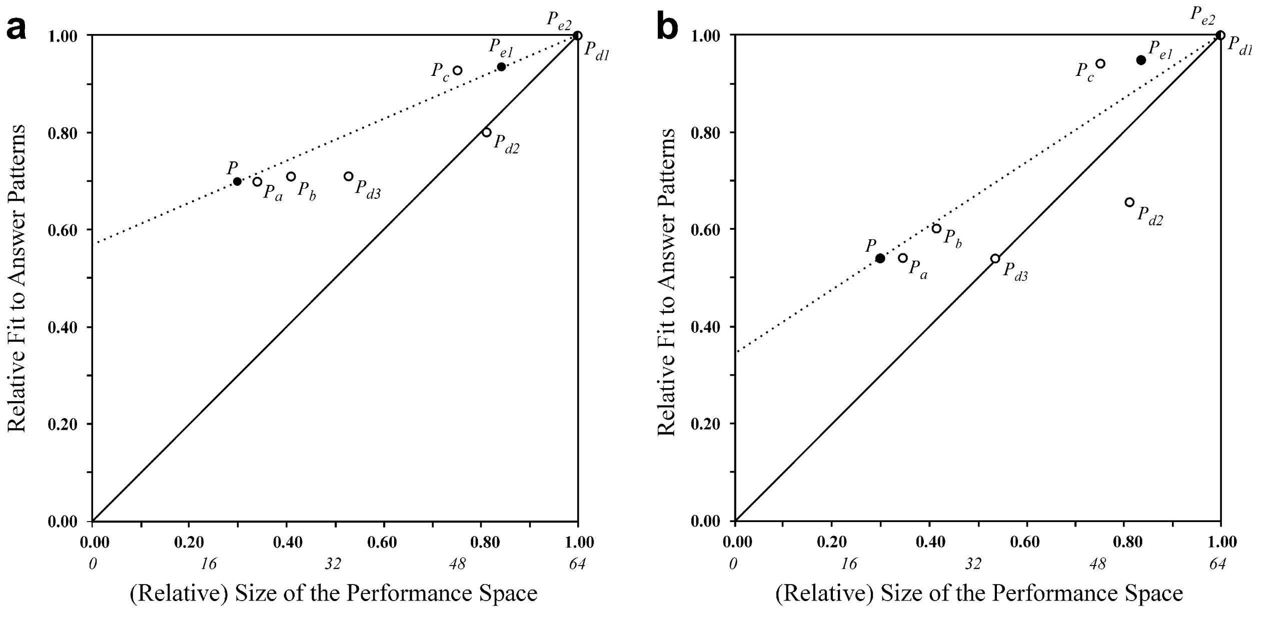

Fig. 4. (a) The relative frequencies of 0-distance patterns for the proposed performance model *P* and the six definitions of overgeneralization in Investigation 1 and (b) those of Investigation 2. P_{e1} and P_{e2} indicate the performance model, including states that might occur from making one or two careless errors or lucky guesses.

located significantly above this line. Because the definitions of overgeneralization extend the basic model *P*, the fit cannot decrease in comparison to *P*. Models below the dotted line cannot be considered explanatory; instead, only the models above the dotted line are of interest.

The results of Investigation 1 are shown in Fig. 4a, those of Investigation 2 in Fig. 4b. According to these results, none of the definitions of overgeneralization appeared to be a reasonable explanation for the empirical data, providing a clearly increasing fit for the empirical data accounting for the increase of additional performance states. Binomial tests for the five analyzed definitions of overgeneralization reported that none of the performance structures covered significantly more answer patterns than expected by chance. On the basis of the results of both investigations, we argue that an increase of fit obtained with any definition of overgeneralization is solely due to a random increase of fit with increasing size of the performance structure.

Discussion

Cognitive development in general and the development of understanding and integrating DST concepts in particular were subject of a large body of theoretical and empirical work. Traditional models suggest that development occurs as discrete shifts of cognitive abilities. Consequently, such models propose several distinct developmental stages. It often was assumed that each developmental stage determines a global level of cognition. More recent approaches contested these assumptions and suggested that development is a steady, continuous process that does not necessarily determine a global level of cognition.

In the present work, we introduced a more general theoretical framework of modelling cognitive development based on CbKST. We exemplified this framework through a model of the development in the DST domain, which not only includes facets and assumptions of previous models, but potentially bridges the gap between stage models and continuous models.

The competence–performance model

The main interest of the current study was to introduce a competence–performance model and demonstrated how to validate this model—at least partially—using empirical data.

The present model focuses on physical competences in this domain, outlined by previous authors. The competence model proposes sixteen levels of the developmental course including 129 different competence states. The basic idea of this approach is a separation of latent, unobservable competences and observable performance (i.e., whether a child masters a given task or not). The present competence–performance model is based on the assumption that development is determined by a continuous acquisition of elementary competences, which can be interpreted as small, distinct entities of knowledge or ability. Insofar, this model adopts the view of discrete developmental stages; however, on a much finer scale, it bridges the gap to models of continuous development. The developmental course is characterized by a sequence of acquiring competences, which follows a specific individual developmental path, proceeding from one individual competence state to another. However, the individual paths cannot occur arbitrarily, but are limited by the prerequisite relation between competences and follow a common direction of development. For example, a child cannot acquire the competence to detect that speed is constant in a given task without knowing what speed actually means. Additionally, certain bottlenecks may occur along the developmental paths, meaning that all children may have to pass a certain competence state and that the same final developmental state is reached. Consequently, development is seen as a process including tiny developmental steps that are different at an individual level and that are not linked to age, and that continuously increment knowledge and abilities in a domain.

This developmental process is a latent, cognitive process. Although previous authors suggested that competence and performance cannot be viewed separately (Anderson & Wilkening, 1991), in our point of view, this distinction is useful to integrate different and—at the first sight—maybe contradictory empirical findings. Moreover, it seems not sensible to link theories of cognitive development to an arbitrary number of experimental (and everyday) tasks. The introduced approach utilizes interpretation and representation functions to map a set of tasks to a set of competences and vice versa. An advantage of this formal mapping is that no one-to-one correspondence between competences and tasks is required and, therefore, the developmental course of competences can be analyzed independent from a specific experimental paradigm. More importantly, the competence model makes assumptions about the latent development of competence and induces predictions on children's behaviour, which can be assessed by the performance in a set of tasks.

As already discussed in previous sections, the proposed model convincingly covered a major part of empirical answer patterns for the six tasks types introduced by Matsuda (1994). These results were found for Japanese as well as Austrian participants. In Investigation 1, a minimal symmetric difference of .37 was found, and in Investigation 2, a minimal symmetric distance of .51 (the maximum distance is 3 in each case). This is a remarkable result since the performance model predicts only a comparably small number of answer patterns (i.e., 19 out of 64 possible patterns). The somewhat better results of the Japanese study in comparison to the Austrian children in terms of the model's fit as well as in terms of performance may be explained by two factors. First, the usage of abstract computer animations in Investigation 2 perhaps was more difficult than the real toy trains in Investigation 1. Second, while Investigation 1 was conducted in a laboratory, in Investigation 2 children were examined directly at their kindergarten and schools, which perhaps were more disruptive environment (although the experimental session were conducted in separate rooms).

In addition, we attempted to apply five formal definitions of overgeneralization from inverse to direct-relation tasks to explain empirical answer patterns deviating from the basic performance model *P*. The results of both investigations revealed that these definitions could not contribute to the explanation of answer patterns deviating from the basic model *P*.

The question concerning the reasons for answer patterns deviating from the proposed performance model remains. An explanation is careless errors and lucky guesses (as introduced by Doignon & Falmagne, 1999). As shown in Fig. 3, if we include answer patterns that might occur from making a single error (i.e., 1-distance patterns, P_{e1}) in Investigation 1, we obtain a fit of 89.92% (i.e., 107 of 119 answer patterns). If we further include answer patterns that might occur from making two errors (i.e., 2-distance patterns, P_{e2}) we obtain a fit of 100.00%. Similar results were found in Investigation 2. As shown in Fig. 4, if we include 1-distance patterns (P_{e1}), we obtain a fit of 88.57% (i.e., 31 of 35 answer patterns). If we further include 2-distance patterns (P_{e2}) we obtain a fit of 100.00%. Remarkably, in both investigations no answer patterns with a minimal symmetric distance larger than 2 were found. This indicates that careless errors and lucky guesses are an appropriate explanation for the deviating answer patterns.

While the present model primarily focuses on physical concepts and related competences, the presented formal framework is also able to include competences, which, in addition to the physical competences, are likely important to fully meet the tasks' requirements (e.g., verbal competences, mathematical competences, or the ability for abstraction). Future research will investigate whether the extension of the model with such competences can further reduce the rate of deviating answer patterns and, therefore, increase the model's fit. However, such extensions require more advanced experimental settings. Similar challenges for future research arise from constraints of the experimental paradigm, including only six task types. In the present work, not all of the 129 competence states could be mapped to the 19 performance states. However, depending on a child's answer pattern, we can determine a child's competence state. For example, a child must have acquired the competences s, d, t, s_c, d_c, t_c, d–t, s–d, and $d \rightarrow t$ to master task type DT. Future analyses may include data collected with different experimental paradigms (e.g., Wilkening, 1982).

The relationship to other models and theories

The presented approach is targeted to include facets and assumptions of existing developmental models in the DST domain as a more general case. Moreover, it might offer a theoretical basis to bridge the gap between stage models and continuous models of cognitive development.

For instance, the two-stage model of Levin (1979) can be covered by a competence cluster $\{s, d, t, s_c, d_c, t_c, d–t, d \rightarrow t, t \rightarrow d, s–d, d \rightarrow s, s \rightarrow d\}$ that includes direct-relation competences, as well as a cluster

$\{s, d, t, s_c, d_c, t_c, d–t, d \to t, t \to d, s–d, d \to s, s \to d, s\backslash t, t \searrow s, s \searrow t\}$ that includes inverse-relation competences (Fig. 1a). As another example, the more comprehensive six-stage model introduced by Matsuda (2001) can be interpreted in terms of the proposed competence–performance model. Matsuda's first stage corresponds to the competence cluster $\{s, d, t\}$; in this stage, children are able to understand distance, speed, and time concepts. In the second stage, children understand direct relations; however, they are limited in their ability to verbalize their reasoning processes. Since the competence model excludes verbal competences, this stage is covered by the competence cluster $\{s, d, t, s_c, d_c, t_c, d–t, d \to t, t \to d, s–d, d \to s, s \to d\}$. According to Matsuda, in the third stage children begin to understand inverse relations between time and speed and, somewhat later, between speed and time. A mapping of this stage is more challenging and must be made into more than one competence state. We have to mention three possible competence states for this stage, i.e., $\{s, d, t, s_c, d_c, t_c, d–t, d \to t, t \to d, s–d, d \to s, s \to d, s\backslash t\}$, $\{s, d, t, s_c, d_c, t_c, d–t, d \to t, t \to d, s–d, d \to s, s \to d, s\backslash t, t \searrow s\}$, or $\{s, d, t, s_c, d_c, t_c, d–t, d \to t, t \to d, s–d, d \to s, s \to d, s\backslash t, s \searrow t\}$. Finally, stages four to six can be mapped to the competence state $\{s, d, t, s_c, d_c, t_c, d–t, d \to t, t \to d, s–d, d \to s, s \to d, s\backslash t, t \searrow s, s \searrow t\}$. This equivocal mapping results from competences regarding the level of consciousness in the application, because they are not part of the present model.

Not only stage models can be formalized with this competence–performance model, but also continuous models. For instance, assemblage theory of Anderson and Wilkening (1991) states that already very young children may have functional understanding of all three concepts: distance, speed, and time. This assumption is in accordance with the present model and covered by competence states {s}, {d}, and {t} and their combinations. Additionally, the present model does not assume a global level of cognition, but relates competences to certain concepts. Thus, a child might have a competence state that includes the ability to perform a certain level of operations (e.g., infer longer distance from longer time), but this does not necessarily mean that this child is also able to use the same level of operation with other concepts (e.g., infer longer distance from higher speed). Moreover, the claims of adaptive thinking (Anderson & Wilkening, 1991; Siegler, 1995, 1996) can be covered by the present model. Basically, adaptive thinking is the ability to adapt cognitive strategies to the requirements of a problem. The competence–performance approach is able to acknowledge this variability. On the one hand, a child, which is assigned a certain competence state, may use any cognitive operation or strategy that is in accordance with the set of competences of the given competence state. On the other hand, the present model can be extended, devising the principle of prerequisite functions (i.e., and/or-type relations). With such an extension, a task can be mapped to different sets of competences, each sufficing to master the task. In Siegler's overlapping wave model, moreover, the probabilities of applying certain operations or strategies change over time. In terms of the presented framework, children are in different competence states and development occurs differently fast. Therefore, a probability distribution over the competence states is induced, which corresponds to Siegler's "waves."

Finally, the presented framework might complement connectionist models in this domain. Connectionist models and related computational algorithms emerged from brain-inspired computation. Buckingham and Shultz (2000) presented a connectionist framework for the development of understanding the DST system. These authors could successfully demonstrate that children's development in this domain can be realistically modelled and mapped by connectionist framework algorithms. While this approach is process-oriented and re-constructs the development on a neural-network basis, CbKST adds a structure-oriented component to the understanding of the cognitive development. A strength both approaches have is that they are able to predict developmental paths (cf. Shultz, Mysore, & Quartz, 2008). In addition, both approaches can be utilized to validate each other. For example, one option is to compare the developmental paths predicted by CbKST models with the progress modelled by connectionist algorithms. Another option is described by Schrepp (1999); he generated different developmental paths with variants of a connection model to cover specific inter-individual differences. In a next step, he mapped a set of tasks to these variants, which induced a knowledge structure and finally used this structure for empirical validation. It is an interesting question for future research to analyze the relationship between these theoretical frameworks and their properties more deeply in order to find synergies and in order to extend and substantiate them.

Conclusions

The introduced competence–performance model well reflects the empirical data, and it turned out to be a suitable approach for formally modelling cognitive development in this domain. The approach can be seen as a more general model, which was derived from the results of previous research that widely includes theoretical assumptions of other developmental theories. We also referenced the possibility to extend the present approach by prerequisite functions (and/or-type relations) to address the ambiguity of mapping competence and performance. Although such extension were not in the scope of the present model, the utilization of a prerequisite function (Fig. 1a) would result in further performance states including the task types *ST* and *TS*.

The presented methodological approach can be applied to other domains of cognitive development to model and analyze the developmental course, especially in domains in which clear prerequisite structures can be hypothesized. The separation of latent competences and observable performance, moreover, allows analyzing the developmental course independent from concrete tasks or experimental paradigms. Finally, this approach also has the advantage that it is selectively applicable to practice, e.g., in teaching or adaptive e-learning systems. The approach allows assessing the developmental state of a child individually by adaptive testing methods, and it allows planning and timing sequences of teaching these concepts. Thus, a teacher can determine the available and missing competences when teaching such concepts. Future research will address challenges due to the constraints of the current model and the experimental task types for assessing the competence model. Moreover, future work will extend the deterministic approach of CbKST by probabilistic models and will analyze the relationship to process-oriented models, which are naturally more variable to cover the variations (and deviations) in the empirical data.

Acknowledgments

Collaboration on the topic presented in this paper between Graz University, Austria and Hiroshima University, Japan started in 2001, when the first author was visiting a professor at Hiroshima University, Graduate School of Education, Department of Learning Science, and when the third author held a professorship at Hiroshima University.

We thank Nathalie Burkert, Melanie Draxler, and Barbara Hermann for collecting the data of Investigation 2. Moreover, we want to express our gratitude to Friedrich Wilkening for providing us with comments and insights regarding a former version of the manuscript.

References

Acredolo, C., Adams, A., & Schmid, J. (1984). On the understanding of the relationships between speed, duration, and distance. *Child Development, 55*, 2151–2159.

Albert, D., & Lukas, J. (Eds.). (1999). *Knowledge spaces: Theories, empirical research, and applications*. Mahwah, NJ: Lawrence Erlbaum Associates.

Anderson, N. H., & Wilkening, F. (1991). Adaptive thinking in intuitive physics. In *Contributions to information integration theory*. In N. H. Anderson (Ed.). *Developmental* (Vol. III, pp. 1–42). Hillsdale, NJ: Erlbaum.

Buckingham, D., & Shultz, T. R. (2000). The developmental course of distance, time, and velocity concepts: A generative connectionist model. *Journal of Cognition and Development, 1*(3), 305–345.

Case, R. (1985). *Intellectual development: Birth to adulthood*. Orlando, FL: Academic.

Case, R., & Sowder, J. (1990). The development of computational estimation: A neo-Piagetian analysis. *Cognition and Instruction, 7*, 79–104.

Correa, J., Nunes, T., & Bryant, P. (1998). Young children's understanding of division: The relationship between division terms in a noncomputational task. *Journal of Educational Psychology, 90*, 321–329.

Crépault, J. (1979). Organisation et genèse des relation temps, espace et vitesse [Organisation and genesis of the relationship between time, space, and speed]. In P. Fraisse (Ed.), *Du temps biologique au temps psychologique [From the biological time to the psychological time]* (pp. 227–254). Paris: Presses Universitaires France.

Crépault, J. (1980). Compatibilité et symétrie. Etude génétique des inférences cinématiques chez des sujets de 11 et 13 ans [Compatability and symmetry: Genetic study of kinematic inferences in 11 and 13 year old subjects]. *L'Année Psychologique, 80*, 81–97.

Demetriou, A. (Ed.). (1988). *The neo-Piagetian theories of cognitive development: Toward an integration*. Amsterdam: North Holland.

Doignon, J.-P. (1994). Probabilistic assessment of knowledge. In D. Albert (Ed.), *Knowledge structures* (pp. 1–56). Berlin: Springer.

Doignon, J.-P., & Falmagne, J.-C. (1985). Spaces for the assessment of knowledge. *International Journal of Man–Machine Studies, 23*, 175–196.

Doignon, J.-P., & Falmagne, J.-C. (1999). *Knowledge spaces*. Berlin: Springer.

Düntsch, I., & Gediga, G. (1995). Skills and knowledge spaces. *British Journal of Mathematical and Statistical Psychology, 48*, 9–27.

Düntsch, I., & Gediga, G. (1998). Knowledge spaces and their applications in CALL. In S. Jager, J. Nerbonne, & A. van Essen (Eds.), *Language teaching and language technology* (pp. 177–186). Lisse: Swets and Zeitlinger.

Fraisse, P., & Vautrey, P. (1952). Perception of space, speed and time in the 5-year-old children. I. Space and speed. *Enfance, 5*, 1–20.

Garnier, R., & Taylor, J. (1992). Discrete mathematics for new technology. Bristol: Hilger. Hockemeyer, C. (2001). KST tools user manual. Retrieved April 15, 2008 from: <http://css.uni-graz.at/publicdocs/publications/file1108637829.pdf/>.

Kickmeier-Rust, M. D. (2002). wzgANI: Distance–Duration–Speed Animation (Unpublished Technical Report). University of Graz.

Korossy, K. (1997). Extending the theory of knowledge spaces: A competence–performance-approach. *Zeitschrift für Psychologie, 205*, 53–82.

Korossy, K. (1999). Modelling knowledge as competence and performance. In D. Albert & J. Lukas (Eds.), *Knowledge spaces: Theories, empirical research, and applications* (pp. 103–132). Mahwah, NJ: Lawrence Erlbaum Associates.

Krist, H., Fieberg, E. L., & Wilkening, F. (1993). Intuitive physics in action and judgment: The development of knowledge about projectile motion. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 952–966.

Levin, I. (1979). Interference of time-related and unrelated cues with duration comparisons of young children: Analyses of Piaget's formulation of the relation of time and speed. *Child Development, 50*, 469–477.

Levin, I. (1983). The sequence in development of time, speed, and distance concepts: Task analysis. *Cahiers de Psychologie Cognitive, 3*, 371–374.

Levin, I. (1992). The development of the concept of time in children. An integrative model. In F. Macar, V. Pouthas, & W. J. Friedmann (Eds.), *Time, action, and cognition. Towards bridging the gap* (pp. 13–32). Dordrecht: Kluwer.

Matsuda, F. (1994). Concepts about interrelations among duration, distance, and speed in young children. *International Journal of Behavioral Development, 17*, 553–576.

Matsuda, F. (2001). Development of concepts of interrelationship among duration, distance, and speed. *International Journal of Behavioral Development, 25*, 466–480.

Montangero, J. (1977). *La notion de durée chez l'enfant de 5 à 9 ans [The understanding of duration in children from 5 to 9]*. Paris: Presses Universitaires France.

Montangero, J. (1979). Les relations du temps, de la vitesse et de l'espace parcouru chez le jeune enfent [The relationship between time, speed, and distance in young children]. *L'Année Psychologique, 79*, 23–42.

Piaget, J. (1969). *The child's conception of time (A. J. Pomerans, Trans.)*. London: Routledge & Kegan Paul.

Piaget, J. (1970). *The child's conception of movement and speed (G. E. T. Holloway & M. J. Mackenzie, Trans.)*. New York: Basic Books.

Piaget, J. (1971). *Psychology and epistemology (A. Rosin, Trans.)*. New York: Grossman.

Schrepp, M. (1999). An empirical test of a process model for letter series completion problems. In D. Albert & J. Lukas (Eds.), *Knowledge spaces: Theories, empirical research, and applications* (pp. 133–154). Mahwah, NJ: Lawrence Erlbaum Associates.

Shultz, T. R., Mysore, S. P., & Quartz, S. R. (2008). Why let networks grow? In D. Mareschal, S. Sirois, G. Westermann, & M. H. Johnson (Eds.). *Neuroconstructivism: Perspectives and prospects* (Vol. 2, pp. 65–98). Oxford: Oxford University Press.

Siegler, R. S. (1995). Children's thinking: How does change occur? In F. E. Weinert & W. Schneider (Eds.), *M emory performance and competences: Issues of growth and development* (pp. 405–430). Mahwah, NJ: Lawrence Erlbaum Associates.

Siegler, R. S. (1996). *Emerging minds: The process of change in children's thinking*. New York: Oxford University Press.

Squire, S., & Bryant, P. (2003). Children's understanding and misunderstanding of the inverse relation in division. *British Journal of Developmental Psychology, 21*, 507–526.

Strauss, S. (1998). Cognitive development and science education: Toward a middle level model. In *Handbook of child psychology* (5th ed.). In I. E. Sigel & K. A. Renninger (Eds.). *Child psychology in practice* (Vol. 4, pp. 357–399). New York: Wiley.

Wilkening, F. (1981). Integrating velocity, time, and distance information: A developmental study. *Cognitive Psychology, 13*, 231–247.

Wilkening, F. (1982). Children's knowledge about time, distance, and velocity interrelations. In W. J. Friedman (Ed.), *The developmental psychology of time* (pp. 87–112). New York: Academic Press.

Wilkening, F., & Anderson, N. H. (1982). Comparison of two rule-assessment methodologies for studying cognitive development and knowledge structure. *Psychological Bulletin, 92*, 215–237.