

Reliability and Validity of a New ICF-CY-Based Mobility Screening for Children From 8 to 10 Years

How to cite:

Dincher 2023. "Reliability and Validity of a New ICF-CY-Based Mobility Screening for Children From 8 To 10 Years". *Journal of Applied Health Sciences and Medicine* 3(5): 1 – 11. <https://doi.org/10.58614/jahsm351>

Andrea Dincher^{1*}

¹Institute for Sports Sciences RPTU Kaiserslautern-Landau Campus Landau Fortstr. 7, building H 76829 Landau, Germany.

Abstract:

Background: In this study, a new mobility screening for children from eight to ten years is evaluated for its reliability and validity.

Methods: 324 children (197 boys and 127 girls, mean age of $9.30 \pm .89$ years) participated. All underwent the MobiScreen 8-10 and a further motor test. Item split times including penalty seconds of MobiScreen 8-10 and raw values and percentile ranks of the additional test are used to evaluate reliability and validity.

Significance level: $p < .05$. **Results:** Coefficients for inter-rater: $r = .99$ to $r = 1.00$, for test-retest: $r = .35$ to $r = .77$, for paralleltest: $r = -.02$ to $r = .62$, and Cronbach's $\alpha = .62$. Explorative factor analysis shows a single factor with an eigenvalue >1 (variance explanation 44 %). Criterion validity: $r = -.53$ to $r = -.71$. Discriminant analysis: significant differences between healthy and disabled children in all tasks excepting transporting. Diagnostic accuracy: The Area Under the Curve reaches a value of .89. **Conclusions:** Mostly all psychometric properties could be reached in a medium to high degree. Retest and Paralleltest Reliability will be repeated. Next step is the determination of a cutoff value to classify into "normal" and "conspicuous".

Keywords: assessment, motor development, primary school, validation study.

1. Introduction

Children with a motor development disorder can only manage their daily lives with problems. Their movements are poorly timed, unrhythmic and inefficient: they also have poorer postural control, which can interfere with the acquisition of other gross motor skills [1]. Affected children's movements are less smooth, their sense of balance is lacking, and they learn to ride a bike or swim later compared to healthy peers [2]. Catching the body is less successful, the gait shows high variability, and movement awareness is poorly developed [3]. They show problems with teetering and climbing, walking on uneven ground, and running with sudden stops or changes of direction and throws. They are usually conspicuous by being clumsy, falling frequently, or dropping objects [4]. Other deficits are evident in activities of daily living, such as eating with a knife and fork, dressing and undressing, tying shoes, or doing puzzles. Even in kindergarten, affected children are conspicuous for their clumsiness. However, these disorders only become relevant when higher gross and fine motor requirements are placed on the children, which is usually the case when they start school. Therefore, an accurate diagnosis as early as possible is of high practical importance [5]. However,

Received:

April 21, 2023

Accepted:

April 27, 2023

Published:

May 1, 2023

Corresponding author:

Andrea Dincher

dincher@uni-landau.de

© 2023

by the authors. The terms and conditions of the Creative Commons Attribution (CC BY) licence apply to this open access article

motor development disorders seem to be frequently overlooked at kindergarten age: At the school entry examination, gross and fine motor disorders are noticeable in 6% of children, but they were detected in only 40% of these affected children at the mandatory U9 screening examination at the pediatrician's office [6]. Here, the question would arise whether the available test procedures that are consistently used in practice (milestone concept, German U examinations) are either not sensitive enough or perhaps not suitable at all to detect such disorders, or whether there is perhaps too much subjective leeway in the assessment. Another problem with such examinations is that affected children often use a great deal of imagination and skill to hide their weaknesses, avoid appropriate situations, or simply refuse to perform a task. Therefore, even their own parents may not notice such a disorder or notice it only late [7].

For all children, participation in social life is of great importance, which is specified by the International Classification of Functioning, Disability and Health of Children and Adolescents ICF-CY [8]. In particular, the domain of mobility is therefore of great importance for this child to participate [9]. Here, the ICF-CY provides a good basis to systematically classify the impairment of children with a motor development disorder [4]. The aim of the ICF-CY is to provide a common language for describing health and health-related conditions in a uniform and standardized way to enable data comparisons between different countries, health disciplines and services [8]. Mobility as defined by the ICF-CY describes moving oneself or moving and handling objects, getting around in different ways, and using transportation. Without it, participation does not take place. Since children's life situations are constantly changing during their development, participation plays an overriding role [8]. This urge to participate is innate and has been described as the strongest developmental motor that is an unmistakable part of early childhood development [7].

A review shows that there is not yet a meaningful motor screening for primary school children that meets all psychometric properties to a high degree, excepting the MobiScreen 6-8 [10]. Therefore, this procedure is to be adapted and validated for children from the age of eight using various modifications.

The known mobility screening for children six four to eight years MobiScreen 6-8 is based on the mobility domain of the ICF-CY and is a so-called filter test that reliably identifies the conspicuous children as such. These children need to be further screened with a detailed motor test. The screening consists of the tasks "Getting up from a lying position", "Slalom", "Climbing", "Crawling", "Maneuvering" and "Transporting". The tasks are designed as a course that must be traversed [11].

Now, a new modified version of this screening instrument, MobiScreen 8-10, will be validated. Here, the slalom remains the same as in MobiScreen 6-8, the obstacle to crawl through is supplemented by a second one. The maneuvering task is made more difficult by the fact that the medicine ball must not only be stopped on a marked cross, but first must be maneuvered along a marked corridor. The transport task is now performed with two balls that must be transported simultaneously. The following figure 1 shows the setup of the new MobiScreen 8-10.

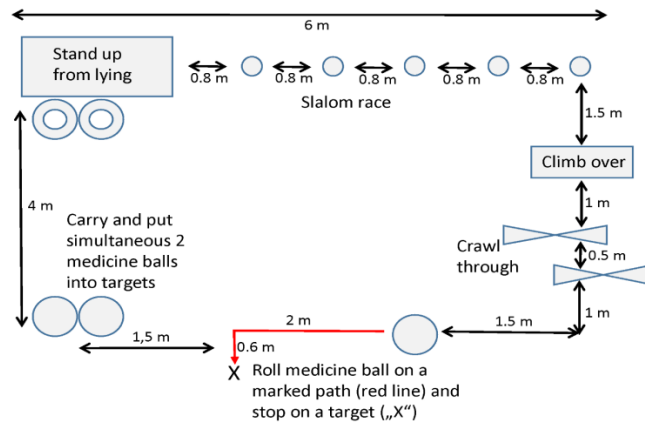


Figure 1. Plan of the MobiScreen 9-10 (adapted from 11)

MobiScreen 8-10 is a new procedure designed to be used in the first step in the diagnostic process.

To validate a test, the psychometric properties reliability and validity are mainly used as instruments for quality assessment and scientificity [12]. Testing and compliance with these criteria is considered essential [13] for this new screening instrument:

Reliability, the degree of accuracy of a test in measuring a particular characteristic [14], should be tested as early as possible in test development, using a sample of at least 100 subjects [15]. Internal consistency as an aspect of reliability can be determined by Cronbach's α from the correlation of all items with each other. Here, a value of $> .80$ is considered acceptable [16], and a value of $> .90$ is considered good [12].

Validity indicates the degree of accuracy with which the test measures the characteristic it is intended to measure [17]. Mangold [18] describes a validity coefficient $> .40$ as medium and $> .60$ as high. Criterion validity is regarded as the most important measure for assessing the practical relevance of a test. For this, test results are compared with the results of an existing test by calculating a correlation coefficient between both test results [19]. Construct validity checks whether the construct measured by the test is related to a similar construct [12]. It refers to the correlation between the test and a latent dimension [13]. The differentiation ability is a further criterion of validity. Thus, in a developmental test designed to separate motor abnormal children (with a given diagnosis such as DCD) from healthy children, it is precisely these two populations that should be compared by a discriminant analysis [20]. Marx and Lenhard [21] state that in addition to the main quality criteria, diagnostic validity is an important criterion of screenings. It can be determined using different concepts. For example, sensitivity, specificity, among others, best describe the ability of a test to classify a person as "impaired"/"conspicuous" or "normal" [22]. For evaluation, in addition to the newly developed procedure, a gold standard test is performed [23]. A predefined criterion, e.g., the sixth percentile or 1.5 standard deviations from mean [3], is used to check which individuals should also be rated as conspicuous in the new procedure. For each potential cutoff value, the corresponding key values of diagnostic quality are then determined: Sensitivity (SN) is used to determine the proportion of individuals correctly identified as having the condition by the test, Specificity (SP) is used to determine the proportion of individuals correctly identified as not having the disease by the test [22]. Meisels

[24] describes for a value of $> .80$ a high sensitivity. The Receiver Operating Characteristic (ROC) analysis aims to find the cutoff value that represents an optimal balance between sensitivity and specificity. The ROC curve is created by plotting specificity and sensitivity against each potential cutoff value. If sensitivity decreases, specificity increases [12]. The AUC (Area Under the ROC Curve) provides a sensitivity index independent of the cutoff value. Hosmer and Lemeshow [25] describe an AUC value of $> .80$ as excellent.

For this, whether reliability and validity are fulfilled for the MobiScreen 8-10 is the aim of this study.

Hypothesis

MobiScreen 8-10 shows good values for reliability and validity.

2. Methodology

Sample of persons

In total, 324 children from the third and fourth grades of randomly selected German primary schools, of which are 197 boys and 127 girls with a mean age of $9.30 \pm .89$ years, a mean height of $1.40 \pm .07$ m and a mean weight of 33.20 ± 8.54 kg participated. The school's management and the children's parents gave written informed consent to participate. Table 1 gives an overview of the sample.

Table 1. Anthropometric characteristics of the sample of persons (number of boys and girls, $M \pm SD$ of height and weight, number of children with a developmental disorder, member in a sports club and migration background)

Boys / Girls	Age	Height	Weight	Development	Sport	Migration
197 / 127	$9.30 \pm .89$	$1.40 \pm .07$	33.20 ± 8.54	57	192	75

Variable sample

The new version MobiScreen 8-10 is used. The procedure has already been described above. For the present study, the item split times including penalty seconds as well as total time needed are used. The measuring points for the split times are determined as follows:

- Slalom: Leaving mat until first contact with gymnastics box
- Climbing: First contact with gymnastics box until last contact with it
- Crawling: Last contact with gymnastics box to first contact with first medicine ball
- Maneuvering: First contact with first medicine ball to first contact with second/third medicine ball
- Transporting: First contact with second/ third medicine ball until safe placement of these balls on the tennis rings [26].

The individual test items are each based on a scale of 0 to 5 penalty seconds based on the error patterns that can be observed. A child gets no penalty second for a task if it has solved it without errors and a maximum of five penalty seconds if it has skipped or

not mastered the task. The following Table 2 shows the exact classification from one to four penalty seconds.

Table 2. Classification of performance errors using penalty seconds

	4 seconds	3 seconds	2 seconds	1 second
Slalom	2 or 3 poles missed or with help	1 pole missed	2 or 3 poles touched	1 pole touched
Climbing	with help	from wrong side	3rd try	2nd try
Crawling	overturned or with help	2nd try	2 or more times obstacle touched	1 time obstacle touched
Maneuvering	not moved on path AND not stopped on cross	hands used AND moved on path OR stopped on cross	hands used AND moved on path AND stopped on cross	moved on path OR stopped on cross
Transporting	ball dropped down, was rolled or with help	consecutive transport	simultaneous transport AND lifted in 2nd try AND placed in 2nd try	simultaneous transport AND lifted in 2nd try OR placed in 2nd try

DMT 6-18: The Deutscher Motorik-Test 6-18 is based on the concept of motor abilities. It consists of the tasks 20-m sprint, balancing backwards, jumping side to side, stand and reach, situp, pushup, standing long jump and 6-min run. Excepting the 6-min run, children have two trials for each task. The best trial counts. The raw values of the tasks are transformed into age and gender specific Z values, so all task Z values can be summarized to a total Z score. These Z scores classify children into strongly above average, above average, average, below average and strongly below average. Here, a Z score < 92 is the border on conspicuousness [27].

MOBAK 1-4: The MOBAK 1-4 is based on the basic motor competencies. It consists of the tasks rolling, balancing, jumping, running (moving oneself) and catching, throwing, dribbling, bouncing (moving objects). Excepting catching and throwing (six trials), children have two trials for each task. For each successful trial, they get one point, in maximum two per task. For catching and throwing, they get one point for three or four successful trials and two for five or six successful trials. All points are summarized to a total score [28].

Procedure: The study took place in November 2022. The children were given an explanation and demonstration of the MobiScreen 8-10 course, then had a practice run. Two test administrators then independently took times on the practice and second run and assigned penalty seconds for the tasks using the protocol sheet. DMT 6-18 and MOBAK 1-4 were conducted by two further test administrators. The order of the test procedures was randomized. All children underwent the MobiScreen 8-10, a part of them underwent the DMT 6-18, the rest the MOBAK 1-4.

Statistics

For the statistical analyses, the program SPSS version 29 is used. Inter-Rater Reliability for the task split times including penalty seconds is determined via Pearson correlation coefficients between the two assessors, Test-Retest Reliability is determined via Pearson correlation coefficients between two trials, and for internal consistency using Cronbach's α . Construct Validity is evaluated by an Explorative Factor Analysis EFA using the task split times including penalty seconds, Criterion Validity is determined via Pearson correlation coefficients between MobiScreen 8-10 total time and (raw-)scores of the tasks, total score and percentile rank of total score from DMT 6-18 and MOBAK 1-4. Differentiation ability is evaluated via discriminant analysis, using the task split times, comparing healthy children to children with a medical diagnosis. Diagnostic accuracy is determined via Receiver Operating Characteristic Curve ROC using the Area Under the Curve AUC. For this, MobiScreen 8-10 total time and the respective percentile ranks for DMT 6-18 (20th percentile) and MOBAK 1-4 (16th percentile) are used as criterion of conspicuousness. The significance level is set at $p < .05$.

3. Results

Inter-Rater, Test-Retest and Paralleltest Reliability

Table 3 presents the inter-rater, test-retest and paralleltest reliability results for the task split times including penalty seconds.

Table 3. Correlation coefficients for Inter-Rater, Test-Retest and Paralleltest Reliability of the MobiScreen 8-10 tasks (* = significant $p < .05$, ** = very significant $p < .01$, *** = highly significant $p < .001$)

	Slalom	Climbing	Crawling	Maneuvering	Transporting
Inter-Rater (n=263)	.99***	.99***	.99***	1.00***	1.00***
Test-Retest (n=90)	.46***	.35**	.56***	.77***	.47***
Paralleltest (n=71)	-.02	.46***	.42***	.62***	.29*

Inter-rater correlation coefficients are highly significant between .99 and 1.00. Test-retest correlation coefficients are most highly significant between .35 and .77. Paralleltest correlation coefficients are most highly significant between .42 and .62, excepting Slalom and Transporting.

Internal Consistency

The value for Cronbach's α is .62.

Construct Validity

Kaiser-Meyer-Olkin (KMO) reaches a value of .71 with a highly significant χ^2 of 169.47*** with ten degrees of freedom. For this, the data are suitable for a main factor analysis. The following figure 2 shows a screeplot with number and eigenvalues of the factors with their variance explanation.

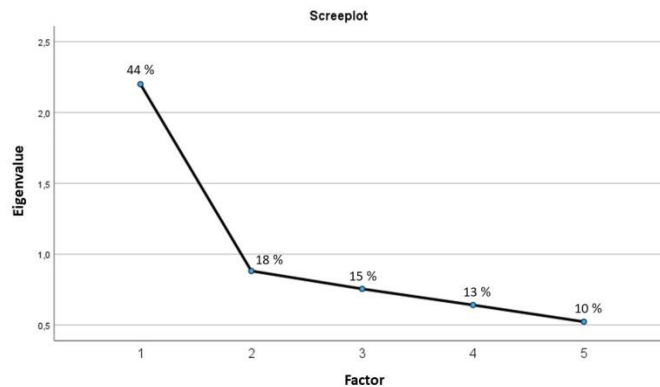


Figure 2. Screeplot of the MobiScreen 8-10 factors with their eigenvalues and variance explanation

A single factor with an eigenvalue of >1 is shown with a variance explanation of 44 %. Adding a second factor would show a variance explanation of 62 %.

Criterion Validity

The following table 4 shows Pearson correlation coefficients between MobiScreen 8-10 and MOBAK 1-4 tasks. The correlation between MobiScreen 8-10 and MOBAK 1-4 total score is $r = -.71^{***}$.

Table 4. Pearson correlation coefficients between MobiScreen 8-10 and MOBAK 1-4 tasks (* = significant $p < .05$, ** = very significant $p < .01$, *** highly significant $p < .001$), $n = 77$

	Throw	Catch	Bounce	Dribble	Balance	Roll	Jump	Run
Mobi-Screen 8-10	-.30	-.48**	-.62***	-.64***	-.31*	-.50***	-.33*	-.57***

Correlation coefficients show significant values between $-.30$ and $-.64$.

Table 5 shows the correlation coefficients between MobiScreen 8-10 and DMT 6-18 tasks. The coefficient between MobiScreen 8-10 and DMT 6-18 total score is $r = -.53^{***}$

Table 5. Pearson correlation coefficients between MobiScreen 8-10 and DMT 6-18 tasks (*** = highly significant $p < .001$)

	20-m Sprint	Balance	Jump	Stand & Reach	Pushup	Situp	Standing long jump	6-minRun
Mobi-Screen 8-10	.44***	-.32***	-.42***	-.27***	-.44***	-.32***	-.46***	-.37***

Correlations are highly significant and range from $-.27$ to $-.46$.

Differentiation Ability

The following table 6 shows the results of the discriminant analysis, comparing healthy children to children with a medical diagnosis.

Table 6. Results of the discriminant analysis, comparing healthy children to children with a medical diagnosis (** = very significant $p < .01$, *** = highly significant $p < .001$)

	Healthy M $\pm$ SD (n=196)	Medical diagnosis M $\pm$ SD (n=51)	Chi ²	Wilks λ	Correct
Slalom	5.69 $\pm$ 1.44	6.34 $\pm$ 1.82	7.26**	.97	79 %
Climbing	1.98 $\pm$ 1.01	3.49 $\pm$ 5.64	12.27***	.95	80 %
Crawling	4.65 $\pm$ 1.81	7.03 $\pm$ 3.67	27.73***	.88	80 %
Maneuvering	8.81 $\pm$ 2.98	12.54 $\pm$ 6.46	33.61***	.87	82 %
Transporting	8.90 $\pm$ 5.58	10.20 $\pm$ 6.86	1.97	.99	79 %

Excepting transporting, all tasks show highly significant differences between both groups. A correct classification is at mean at about 80 %.

Diagnostic Accuracy

The following figure 3 shows the result of the Receiver Operating Characteristic ROC analysis, showing the ROC curve and the Area Under the Curve AUC.

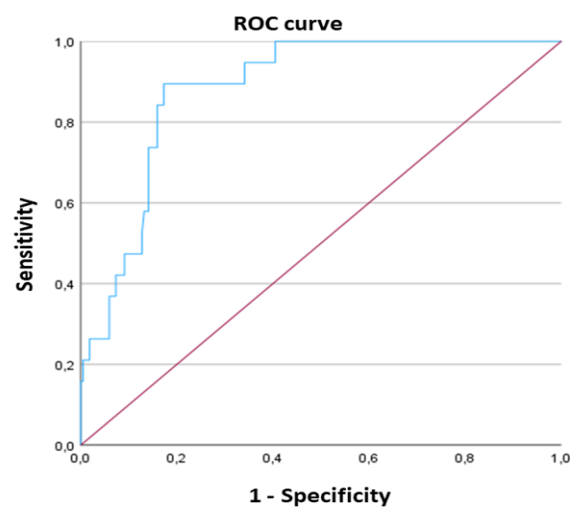


Figure 3. ROC curve, calculated from sensitivity and 1-specificity to each potential cutoff value

The ROC curve shows an AUC of .89.

4. Discussion

In the present study, reliability and validity for the new mobility screening for 8- to 10-year-olds were to be tested.

For inter-rater reliability, coefficients with values between .99 and 1.00 can be considered as excellent for all tasks [29].

Cronbach's α reaches a value of .62 and is at the moment not good in the sense of Pospeschill [12]. So it is possible that all tasks are independent from each other, because each task represents another facet of mobility. According to the ICF-CY, the tasks can be conjugated to d4503 bypassing obstacles (slalom), d4551 climbing (climbing), d4550 crawling (crawling), d4350 pushing with lower extremities (maneuvering), d4300 lifting, d4301 carrying with hands/ d4302 carrying with arms and d4305 set down (transporting) [8].

For retest reliability, coefficients range from .35 to .77 and can be considered as acceptable [13]. The time interval between both measurements should not be too short in order to exclude memory effects [15]. However, in this case there was only one day in between and the children knew exactly about their performance of the previous day. Because of this, the course was often run through faster, but with more errors, respective penalty seconds. Thus, an interval of two weeks could prove to be appropriate.

Coefficients for paralleltest reliability range from -.02 to .62 and can be considered between low and acceptable [13]. Perhaps the coefficients are not so high here because of the test design. For the sake of simplicity, the children ran through the original course in reverse order. There is only one meter of space between crawling and climbing, which makes it difficult for larger children in particular to crawl quickly and then stand up without touching this obstacle. The slalom was also problematic here, with many children running off on the right side who had started on the wrong side during the original parcours. This coefficient turns out to be the lowest in comparison with the other methods, but the method is supposed to meet the needs of practice in the athletic field most closely [15].

For construct validity, it shows very clearly a single factor with an eigenvalue > 1 and a variance explanation of 44 %. Thus, the assumed single factor model of mobility can be confirmed.

For criterion validity, coefficients reach -.71 for MOBAK 1-4 and -.53 for DMT 6-18. They can be considered as high [18]. Looking at the correlations of the MobiScreen 8-10 to the tasks of the MOBAK 1-4, it can be seen that bouncing and dribbling a ball in a slalom course and running have the highest coefficients. It can be assumed that these basic competencies are important for developing appropriate mobility and thus being able to participate in everyday life. Looking at the correlations to DMT 6-18, it can be seen that 20-meter sprint, push-ups and standing long jump have the highest coefficients. It can be assumed here that precisely the motor abilities measured here, strength and velocity, could be decisive in mastering the course.

The discriminant analysis shows highly significant differences between healthy children and children with a medical diagnosis excepting transporting. Here, a difference of about 1.5 seconds is visible. Healthy children mastered this task faster, but made a similar number of errors (e. g. ball dropped down). Overall, the MobiScreen 8-10 is clearly able to differentiate precisely between these two populations. For each task, an average of 80% of all children can be correctly assigned to their group via their performance.

The ROC analysis shows that the value for the AUC reaches .89 that can be considered as excellent [25].

Prospects

In the present study, the reliability and validity of the MobiScreen 8-10 were tested. Almost all psychometric properties were fulfilled to a high degree. Only the test-retest method and the parallel test method should be improved. For this purpose, the retests should take place after two weeks and the parallel test should not be run in the reverse order of the original course, but should be set up as a variation. Since diagnostic accuracy is also excellent, it is possible at this point to review potential cutoff values that could classify into normal and conspicuous.

Conflicts of interest: There are no conflicts.

Acknowledgment: This work was not funded. There is no financial interest or benefit in this work.

References

- [1] Schott, N. (2010). Motorische Ungeschicklichkeit. In N. Schott, & J. Munzert (eds.), *Motorische Entwicklung* (pp. 169-185). Göttingen: Hogrefe.
- [2] Esser, G., & Petermann, F. (2010). *Entwicklungsdiagnostik (Kompendien Psychologische Diagnostik, 13)*. Göttingen: Hogrefe.
- [3] Blank, R., & Vincon, S. (2020). Deutsch-österreichisch-schweizerische (DACH) Versorgungsleitlinie zu Definition, Diagnostik, Behandlung und psychosozialen Aspekten bei Umschriebenen Entwicklungsstörungen motorischer Funktionen (UEMF) (Kurzfassung Mai 2020). Available at 022-017k_S3_Umschriebene-Entwicklungsstörungen-motorischer-Funktionen-UEMF_2020-08_01.pdf (awmf.org).
- [4] Jaščenoka, J., & Petermann, F. (2018). LoMo 3-6. Leistungsinventar zur objektiven Überprüfung der Motorik von 3- bis 6-Jährigen. Manual. Göttingen: Hogrefe.
- [5] Suchodoletz von, W. (2005). Frühe Identifikation motorischer Entwicklungsstörungen. In W. von Suchodoletz (ed.), *Früherkennung von Entwicklungsstörungen. Frühdiagnostik bei motorischen, kognitiven, sensorischen, emotionalen und sozialen Entwicklungsauffälligkeiten* (pp. 45-74). Göttingen: Hogrefe.
- [6] Erb, J., & Werner, M. (2003). Prävalenz von Entwicklungsauffälligkeiten bei Vorschulkindern. *Kinderärztliche Praxis*, 74, 368-375.
- [7] Michaelis, R., & Niemann, G. (2017). *Entwicklungsneurologie und Neuropädiatrie. Grundlagen, diagnostische Strategien, Entwicklungstherapien und Entwicklungsförderungen* (5. rev. and ext. ed.). Stuttgart: Thieme.
- [8] Hollenweger, J., & Kraus de Camargo, O. (2013). ICF-CY. Internationale Klassifikation der Funktionsfähigkeit, Behinderung und Gesundheit bei Kindern und Jugendlichen (2. reprint of 1. ed.). Bern: Hans Huber.
- [9] Straßburg, H.-M., Dacheneder, W., & Kreß, W. (2008). *Entwicklungsstörungen bei Kindern. Praxisleitfaden für die interdisziplinäre Betreuung* (4. ed.). München: Urban & Fischer.
- [10] Dincher, A., & Dincher, L. M. (2023). Motor screening for primary school children - A Review. *Archives of Clinical Trials and Case Reports*, 2(1), 123-132. DOI: 10.47191/ijcsrr/V5-i8-41.
- [11] Dincher, A., & Dincher, L. M. (2022). First Reliability Aspects of a new ICF-CY-based Mobility Screening for 7- and 8-year old Children. *International Journal of Medical Studies*, 7(8), 1-9. <https://ijmsonline.in/index.php/ijms/article/view/229>.
- [12] Pospeschill, M. (2010). *Testtheorie, Testkonstruktion, Testevaluation*. München: Reinhardt.
- [13] Bös, K. (2017). *Handbuch Motorische Tests. Sportmotorische Tests, Motorische Funktionstests, Fragebögen zur körperlich-sportlichen Aktivität und sportpsychologische Diagnoseverfahren* (3., rev. and ext. ed.). Göttingen: Hogrefe.
- [14] Lienert, G. A., & Raatz, U. (1998). *Testaufbau und Testanalyse* (6. ed.). Weinheim: Beltz.

- [15] Brunner, G., & Thieß, G. (1970). Zu Fragen der Reliabilität sportmotorischer Tests. *Theorie und Praxis der Körperkultur*, 19(5), 417-423.
- [16] Schnell, R., Hill, P. B., & Esser, E. (2011). *Methoden der empirischen Sozialforschung* (9., actual. ed.). München: Oldenbourg.
- [17] Lienert, G. A. (1969). *Testaufbau und Testanalyse* (3., by an annex on factor analysis suppl. ed.). Weinheim: Beltz.
- [18] Mangold, S. (2013). *Evidenzbasiertes Arbeiten in der Physio- und Ergotherapie: Reflektiert – systematisch – wissenschaftlich fundiert* (2. actual. ed.). Berlin: Springer.
- [19] Gulliksen, H. (1950). *Theory of mental tests*. New York: Wiley.
- [20] Leonhart, R. (2013). *Lehrbuch Statistik: Einstieg und Vertiefung* (3. rev. and ext. ed.). Bern: Huber.
- [21] Marx, P., & Lenhard, W. (2011). Diagnostische Merkmale von Screening-Verfahren zur Früherkennung möglicher Probleme beim Schriftspracherwerb. In M. Hasselhorn, & W. Schneider (Hrsg.), *Frühprognose schulischer Kompetenzen* (pp. 68-84). Göttingen: Hogrefe.
- [22] Ivnik, R. J., Smith, G. E., Cerhan, J. H., Boeve, B. F., Tangalos, E. G., & Petersen, R. C. (2001). Understanding the Diagnostic Capabilities of Cognitive Tests. *The Clinical Neuropsychologist*, 15(1), 114-124.
- [23] Schwarzer, G., Türp, J. C., & Antes, G. (2002). Die Vierfeldertafel (in Diagnosestudien) – Sensitivität und Spezifität. *Deutsche Zahnärztliche Zeitschrift*, 57(6), 333-334.
- [24] Meisels, S. J. (1989). Can developmental screening tests identify children who are developmentally at risk? *Pediatrics*, 83(4), 578-585.
- [25] Hosmer, D. W., & Lemeshow, S. (2000). *Applied Logistic Regression* (2. ed.). New York: Wiley.
- [26] Dincher, A. (2019). Screeningverfahren in der Bewegungsförderung (Schriften der Deutschen Vereinigung für Sportwissenschaft, 277; Forum Sportwissenschaft, 34). Hamburg: Czwalina.
- [27] Bös, K., Schlenker, L., Büsch, D., Lämmle, L., Müller, H., Oberger, J., Seidel, I., & Tittlbach, S. (2009). *Deutscher Motorik-Test 6-18 (DMT 6-18)*. Hamburg: Czwalina.
- [28] Herrmann, C., Seelig, H., Heim, C., Kehne, M., & Gerlach, E. (2018). *MOBAK 1-4. Test zur Erfassung Motorischer Basiskompetenzen für die Klassen 1-4*. Göttingen: Hogrefe.
- [29] Clarke, H. H. (1976). *Application of measurement*. New York: Prentice Hall.