



Considerations when assessing urban South African children with the Developmental Test of Visual Perception 2nd edition (DTVP-2)

Mariska Smith, B OT (UFS), M OT (UFS)

Occupational Therapist, private practice

Marieta M. Visser, B OT (UFS), MSc (OT) (Wits)

Lecturer, Department of Occupational Therapy, School for Allied Health Professions, Faculty of Health Sciences, University of the Free State

Rita van Heerden, BOT (UFS), MOT (UFS), PhD (HPE) (UFS)

Occupational Therapist, private practice

Jacques Raubenheimer, PhD

Lecturer, Department of Biostatistics, Faculty of Health Sciences, University of the Free State

ABSTRACT

Introduction: *The Developmental Test of Visual Perception second Edition (DTVP-2) is frequently used by South African (SA) occupational therapists, despite the suitability of its US-based norms for SA children being questioned, and the 2014 release of the updated DTVP-3. This study investigated the suitability of the DTVP-2 norms for SA English-speaking children aged 5y 6mo–5y 11mo. SA sample scores were compared to American norms. Motor-reduced subtest scores were compared with and without the stop rule. Gender differences were tested.*

Methods: *A quantitative, descriptive study used 134 English speaking children. Motor-enhanced subtests were administered as prescribed, but motor-reduced subtests were administered without the prescribed stop rule, allowing prescribed and adapted scores to be computed.*

Results: *Scores of SA children varied from American norms, especially in visual closure, visual-motor speed and form constancy. Better visual closure scores were obtained when the stop rule was excluded. Boys and girls differed only on figure-ground, where girls scored higher.*

Conclusion: *The DTVP-2 is valuable, but caution is recommended when measuring visual perceptual skills for this age band. Alternative instruments should be used, or local norms, or even more contextually relevant instruments, must be developed locally. At least, occupational therapists should apply US norms with care.*

Key words: Developmental Test of Visual Perception 2nd edition (DTVP-2); visual perceptual skills; 5-year-old children; urban South Africa; suitability

INTRODUCTION

Occupational therapists are professionally trained to assess various developmental occupational performance components and skills of children¹. One aspect often assessed and addressed within the scope of occupational therapy practice is visual perceptual skills. The assessment of visual perceptual skills is useful for determining school readiness and identifying underlying difficulties experienced in occupational performance areas such as school activities².

Visual perceptual skills can be assessed during clinical observations, informal assessment and by means of standardised measurement instruments³. The use of standardised instruments provides occupational therapists with trustworthy evidence for their assessment findings, by means of which the child being assessed is compared with norms obtained from a normative sample. However, when tests are standardised in other countries, South African occupational therapists often rely on those standardised measurement instruments without questioning the suitability (for example with regard to culture, environmental affordance, socio-economic status, schooling system) of the instrument for the specific context⁴.

The lack of context-specific standardised assessment instruments is a universal problem in all fields of occupational therapy practice in South Africa (SA). In the field of paediatric practice, no comprehensive visual perceptual skills measuring instruments that have been standardised on a representative SA population, are cur-

rently available⁵⁻⁸. Therefore, occupational therapists in SA make use of measurement instruments standardised in other countries to measure children's visual perceptual skills⁵⁻⁸.

Internationally standardised measurement instruments most often used by paediatric occupational therapists in SA are, in order of preference⁹, (in a study conducted prior to the latest editions) the Developmental Test of Visual-Motor Integration (Beery VMI5th or 6th edition)¹⁰; the Developmental Test of Visual Perception 2nd edition (DTVP-2)¹¹; the Test of Visual Perceptual Skills (Non-Motor)—Revised (TVPS-R)¹²; the Southern California Sensory Integration Tests (SCSIT) / Sensory Integration and Praxis Tests (SIPT)¹³; and the Goodenough-Harris Draw-A-Person Test (DAP)¹⁴.

According to Van der Merwe, Smit and Vlok⁹, occupational therapists in SA prefer to use the DTVP-2 as a measurement instrument more frequently than occupational therapists in other countries. The DTVP-2 is regarded as well designed, easy to conduct and administer¹⁵ and is considered a valuable clinical measurement instrument to descriptively measure visual perceptual and visual-motor integration (VMI) skills¹⁶. The DTVP-2 is considered advantageous as the instrument allows occupational therapists to make additional observations regarding children's motor dexterity¹⁷ while completing the DTVP-2's motor-enhanced subtests, contrary to other visual perceptual tests that consist of motor-reduced items only. Another "hidden" advantage of this instrument (although not explicitly re-



ferred to in the manual but observed from clinical experience) is that the administration of the DTVP-2 is structured in such a way that young children's level of concentration is taken into account by alternating the motor-reduced and motor-enhanced subtests, providing variation for children participating in this standardised test.

Despite claims that the subtests of the DTVP-2 are unbiased for culture¹¹, SA children score differently on this standardised measurement instruments in relation to the norms for American children, on which it was standardised^{8,18,19}. Consequently, the norms of the DTVP-2 do not translate well to the SA context and the suitability of the DTVP-2 as a measurement instrument to assess visual perceptual skills in SA children is questioned by researchers^{8,18,19,20}, as well as occupational therapists working in clinical settings. Aspects questioned are, for example, item linearity and level of difficulty and therefore adapted administration and scoring methods have been investigated^{8,20,21}. However, in order to analyse item linearity of any subtest of the DTVP-2 for research purposes, results of all the items are required⁸. This requires the stop rule on the motor-reduced subtests to be abrogated in order for all the items to be scored and individually analysed. For the purposes of this study, this completion of all items will be referred to as the adapted administration and/or scoring.

With regard to gender performance, SA researchers^{8,18,19,20} have not differentiated and/or reported gender-specific findings on the DTVP-2 in SA. Visser et al.⁸ recommended further investigation in view of the fact that Cheung et al.²¹ had noticed a difference in gender performance in the Copying (CO) and Figure-Ground (FG) subtests of the DTVP-2 in Hong Kong children, with girls performing significantly better than boys. It would therefore be of value to establish whether gender-specific differences in DTVP-2 performance can be observed in SA, since differences in gender performance have been reported in studies using other standardised measurement instruments in the South African context.

Although a revised version of the DTVP-2, the Developmental Test of Visual Perception 3rd edition (DTVP-3)¹⁷, was published during the course of the research reported in this article, the DTVP-3 was still relatively unknown and/or minimally used at the time of the study. Furthermore, according to the authors' knowledge, very little research has been conducted on the DTVP-3 other than that reported in the DTVP-3 test manual.

This article reports on a study that aimed to investigate the suitability of the DTVP-2 norms for SA by firstly describing the differences in DTVP-2 scores among SA boys and girls, and secondly, relating how results of a sample of SA children compares to the norms derived from the American normative sample. Thirdly, the results of the prescribed scoring method and an adapted scoring method for the four motor-reduced subtests and composite scores are compared.

LITERATURE REVIEW

According to the Occupational Therapy Practice Framework¹, visual perception is an occupational performance skill needed to support engagement and participation in daily life occupations. Visual perception and visual-motor integration are considered functionally important for childhood occupations. It is imperative for paediatric occupational therapists to accurately assess the incidence, degree and impact of visual perceptual skill difficulties in children². To ensure best practice and accurate results, it is crucial to make use of culturally suitable measurement instruments⁴, since cultural diversity influences the interpretation of children's perceptual development^{15,21,22}. If a child's culture is inadequately represented in the normative sample, "test results should be interpreted with caution"^{11:50-51}. Under- or over-identification of visual perceptual difficulties should be avoided since under-identification of visual perceptual skill difficulties can denote misdiagnosis or the child not receiving therapy^{6,7}, while over-identification can lead to excessive recommendations, resulting in misuse of educational and medical resources⁷.

DTVP-2

The DTVP-2 was developed and standardised in 1992 on a sample of 1 972 English-speaking children from different cultures inherent to the American population, residing in 12 American states mainly in urban settings¹¹. Health professionals make use of the DTVP-2 to establish the incidence and extent of children's visual perceptual or VMI skill difficulties, plan interventions, make appropriate referrals, validate the efficacy of intervention programmes, and for research purposes¹¹.

The DTVP-2's psychometric properties display the test as a valid and reliable measurement instrument to assess visual perceptual and VMI skills¹¹. The DTVP-2 has adequate levels of validity and reliability, although it relates to the American normative population. The DTVP-2 measures skills related to eye-hand coordination (EH), position in space (PS), copying (CO), figure-ground (FG), spatial relations (SR), visual closure (VC), visual-motor speed (VMS) and form constancy (FC) of children in the 4–10 years age group.

Different stop rules apply for the subtests of the DTVP-2. With regard to the motor-enhanced subtests, EH has no stop rule and all items are administered; in CO and SR, testing is continued until a ceiling of 3 consecutive incorrect items is reached; VMS has no stop rule, but a 1-minute time constraint applies. The motor-reduced subtests (PS, FG, VC and FC) are administered until the child achieves a ceiling point of 3 incorrect responses out of 5 consecutive items¹¹.

Subtests are individually calculated in order to generate a total raw score (RS) for each subtest. Raw scores obtained can be converted to an age equivalent (not recommended), percentile and/or standard score (SS) by means of normative tables in the Examiner's Manual of the DTVP-2. Since no South African norms exist, South African occupational therapists use the US-based norms when testing South African children. Interpretation should not only rest on subtest scores, but also on the basis of the composite quotient scores¹⁷. Subtest standard scores (SSs) are used to generate composite scores either allocated to the motor-reduced or motor-enhanced component to provide occupational therapists with a motor-reduced visual perceptual quotient (MRPQ) and a visual-motor integration quotient (VMIQ), respectively¹¹. A General Visual Perception Quotient (GVPQ) is generated by means of combining the SSs of all the subtests¹¹. Composite scores (quotients) are the most reliable DTVP-2 scores¹⁷. These scores are the best measure to use for interpreting the results, and identifying general visual perceptual strengths and weaknesses. Children scoring high on the composite scores are likely to perform well on activities such as puzzles, copying and pegboards that require visual perceptual abilities. The MRPQ is the "purest" and most direct measure of visual perception as compared to the VMIQ that provides a more integrated score between visual perception and certain fine motor abilities, respectively¹¹. A General Visual Perception Quotient (GVPQ) is generated by means of combining the SSs of all the subtests¹¹.

To date, limited research has been conducted in SA to investigate the impact of culture, language and socio-economic status on the DTVP-2's psychometric properties, and norms have not been established specifically for the SA population. Despite the lack of research, the use of the DTVP-2 in SA as a psychometric instrument was condoned by the Health Professions Council of South Africa (HPCSA)²³. The validity of the DTVP-2 in the SA context has been questioned by several SA researchers^{8,18,19,20} and some important concerns were raised.

Visser et al.⁸ established a statistically significant difference between the DTVP-2's below average VC scores and the other seven DTVP-2 subtests. Moreover, 97.5% of forty Grade 1 children experienced VC as their greatest visual perceptual difficulty¹⁸. While 40% of 5-year-old English-speaking children scored below the normal score for their age in the VC subtest of the DTVP-2, 57.5% obtained a normal score for their age, and only 2.5% performed higher than their age-appropriate score. In addition, when

the prescribed motor-reduced stop rule (as set out in the DTVP-2 Examiner's Manual) was abrogated, results for VC were obtained which better corresponded to the other subtests⁸.

This problem with VC scoring when using the stop rule may be related to the item linearity of the VC subtest. Richmond and Holland²⁰ noted that the DTVP-2 and TVPS-R's VC subtest scores did not compare with each other and that the DTVP-2 scored significantly lower. They queried the item linearity of the VC subtest and proposed that the DTVP-2 might possibly over-identify VC difficulties²⁰.

Even though VC has been identified as the biggest concern emanating from the DTVP-2, it is also noteworthy that the majority of the Grade 1 children scored below average on the PS and SR subtests of the DTVP-2¹⁸. Furthermore, the majority of these children scored average and above average on the CO, FC and VMS subtests of the DTVP-2¹⁸. Visser et al.⁸ further established that all the 5-year-old English-speaking children tested in their study scored average and above average on the SR and FC subtests, while the majority of the children scored average and above average on the EH, CO and FG subtests of the DTVP-2⁸.

Acknowledging the value and the popular use of the DTVP-2, the suitability of the DTVP-2 norms is, however, questioned in SA and other countries. Cheung et al.²¹ described distorted item difficulty in the DTVP-2's EH, PS and SR subtests for Hong Kong children. Additionally, Lai and Leung²² established that 5-year-old Chinese children obtained high scores on the CO and SR subtests of the DTVP-2, while 5-year-old English-speaking Australian children obtained the lowest mean score for the VC subtest. Guntayoung et al.¹⁵ found that Thai children scored similarly to their American counterparts with the exception of VMS, in which Thai children aged 4–5 years and 8–11 years scored higher. Guntayoung et al.¹⁵ reasoned that a difference in scores was the result of cultural and environmental experiences relating to, amongst other factors, learning the complex Chinese letter formation from an early age.²²

The rationale of this study is therefore supported by South African and international literature, aiming to build on previous research to further investigate the DTVP-2 in a South African context, so as to translate evidence into practice.

METHOD

Study design

A quantitative, descriptive, observational study was conducted.

Population and sampling

The study sample was recruited using stratified random sampling from urban schools located in the City of Tshwane metropolitan area. The Gauteng Department of Education divides the geographical area that falls within the City of Tshwane into four educational districts, namely Gauteng North, Tshwane North, Tshwane South and Tshwane West²⁴. Diverse ethnic, culture and language groups from different socioeconomic backgrounds reside across the four districts, which represent the SA population in the study sample. However, it should be noted that this urban SA sample could differ from the DTVP-2 American norm population with regard to ethnicity, culture, language and socioeconomic status. The American sample consisted of children from rural and urban areas in 12 American states with approximately one third from non-American ethnic groups, and 3% children with disabilities¹⁷.

This study population was limited to exclusively English-language speakers since the DTVP-2 was standardised in English and is not available in any other SA language. English-language speakers were defined as children who had attended an English Language of Learning and Teaching (LOLT) school for at least two years and spoke English as their first, second or third language. The time frame of two years was supported by Du Plessis and Louw²⁵ who stated that it requires approximately two years of constantly speaking English

on a daily basis to become proficient in basic English interpersonal communication skills²⁵. Furthermore, the translation of the DTVP-2 subtest instructions could lead to different perceptions and/or misinterpretation of test items and instructions, consequently influencing the child's results^{6,26,27}.

Although the population comprised a heterogeneous group of children, ethnicity, culture, language and socio-economic status were not included and controlled in the study since these variables would have necessitated a larger study sample, which was beyond the scope of the study. However, gender and age were controlled based on the findings reported by Cheung et al.²¹, who identified a gender difference in the performance of Hong Kong boys and girls, and literature²⁸ differentiates gender in the acquisition of visual perceptual skills.

Furthermore literature indicates that visual perceptual skills demonstrate an increase in development from 5 years of age⁶, but which could also fluctuate with the introduction of more formal schooling. The choice of this particular age band also made possible comparisons with existing SA studies and allowed the researchers to build on previous research on the DTVP-2 on five-year-olds in SA, since studies conducted by Van Romburgh¹⁸, Visser¹⁹ and Visser et al.⁸ provide evidence of a norm discrepancy in DTVP-2 SSs in 5-year-old children.

Although previous studies looked at the whole 5 year age group, this study delineated the population to only one of the DTVP-2 age intervals: the 5 years and 6 months to 5 years and 11 months interval, in order to perform a more in-depth analysis. Because this study was temporally and financially limited, substantial numbers of study participants (a minimum of 10 boys and 10 girls per interval) were included equally within each of the one-month intervals of 5 years and 6 months to 5 years and 11 months.

Children were excluded from the study if their parents/caregivers/teachers suspected developmental delays, or they had been diagnosed with sensory, physical, emotional or intellectual impairment, or any other condition not mentioned that could influence participation and/or their results on the DTVP-2. These conditions included (i) autism spectrum disorder; (ii) attention deficit disorder/attention deficit hyperactivity disorder (ADD/ADHD); (iii) cerebral palsy; (iv) Down's syndrome; (v) learning disability; and (vi) psychiatric conditions such as conduct disorder or oppositional defiance disorder. Children having been tested with the DTVP-2 within the previous 6 months, or those who had previously received or were currently receiving occupational therapy or any other kind of therapy that could influence a child's visual perception, were also excluded. Children whose parents or caregivers had not given informed consent for their child to participate in the study, or children who had not given assent to participate, were excluded.

A list of 108 eligible nursery/pre-primary and 43 primary schools, spread across the four educational districts was compiled from various sources (Department of Basic Education (Gauteng), the Yellow Pages, Internet searches, and community-information-search results, obtained from the Tshwane Community Library). A total of 92 schools were contacted randomly by telephone and in relatively equal proportions per district, to request permission to conduct the study at the school. From these 92 schools, 17 schools refused participation, a number of nursery schools were unreachable, some schools did not have children who met the inclusion criteria and other schools had to be excluded since they used dual language teaching. Information was emailed to the remaining 69 schools, of which 29 schools provided permission. These 29 schools comprised two private, two primary and 25 nursery/pre-primary schools.

School principals and teachers at the 29 schools identified 391 potential children that fell within the correct age bracket of 5 years 6 months and 0 days to 5 years 11 month and 29 days during the data collection period. Parents/caregivers of these children received research information packs containing a research information pam-

Table I: School selection in relative equal proportion to the size of the four educational districts

	District				Total
	Gauteng North	Tshwane North	Tshwane South	Tshwane West	
Number of schools in district	4	19	118	10	151
Schools providing permission	1	5	22	1	29

phlet, a self-administered screening questionnaire and an envelope. One hundred and eighty-three (183) children's parents/caregivers provided consent by completing and returning the questionnaire, resulting in a low-average response rate of 46.8%, where around 55% is considered as a norm for paper-based surveys according to Nulty²⁹. However, 41 of these children did not meet the inclusion criteria, five children were unavailable during data collection and three did not complete the testing, leaving a final sample of 134 children.

Data collection

The DTVP-2 was administered individually, in English, with all the children (134), by the researcher. The following method was used: Instructions and administration for the motor-enhanced subtests followed the prescribed methods exactly. However, for the DTVP-2 motor reduced subtests, the prescribed instructions were followed, but the administration was modified by abrogating the stop rules, thus allowing all children to complete all the test items. Scores were calculated after testing in two ways: one imitating the scoring had the stop rule been used, and the other providing a score on all the items. Even though, if the subtest items were truly linear (with a monotonic increase), this should not have changed the children's scores materially, it should be noted that this use of the adapted method was purely for research purposes and was not aimed at changing the DTVP-2, nor to undermine the prescribed test administration and scoring, but rather to have scores on all the test items for each subtest for more extensive data analysis in item linearity (which may form the focus of a later article).

Use of a single administrator eliminated the problem of inter-rater reliability, and administrator bias was controlled for by using instruction cue cards during administration of the DTVP-2 and by strictly following administration procedure as set out in the Examiner Manual of the DTVP-2¹⁷ for each child. Scoring of the subtests, and conversion to standard scores and composite quotients, was done by the researcher after the data collection period to ensure confidential recording of results. Two third-party Occupational therapists with at least 10 years' clinical experience in paediatric Occupational therapy practice verified the scores before the data were transferred to the data score sheet. The researchers minimised possible clerical errors by verifying and correlating all manually calculated scores with automatically computed scores with the Microsoft Excel programme. Coding of the data score sheets was verified by an independent third-party person.

Data analysis

For this article, categorical data are described using frequencies and percentages. Continuous data are reported using means and/or medians with inter-quartile ranges (IQRs).

Gender differences were examined by means of Student's t-tests, paired t-tests were conducted to establish differences in prescribed and adapted test scores, and single-sample t-tests were used to test the difference between the mean scores for this sample, and the means reported in the DTVP-2 manual for the normative sample. The data analysis was done by the Department

of Biostatistics, University of the Free State (UFS) using SAS/STAT software, Version 12.3 of the SAS System for Windows (©2012 SAS Institute Inc.).

Ethics

Approval for this study was obtained from the Ethics Committee of the Faculty of Health Sciences at the University of the Free State. To ensure ethical conduct, written permission was received from Pro-Ed Inc., publishers of the DTVP-2, the Gauteng Department of Education and the participating schools' principals. Consent was obtained from all parents/caregivers of participating children, and assent from children prior to the study.

RESULTS AND DISCUSSION

Of the sample of 134 typically developing children, 65 (48.5%) were boys. The children were distributed reasonably evenly across the ages 5 years 6 months and 0 days through 5 years 11 months and 29 days. *Table II* shows the distribution across gender and age. Thirty-five (26.1%) of the children were enrolled in grade RR and 99 (73.9%) in grade R.

Table II: Age and gender of study participants

	Age per one-month interval					
	5 y 6 m	5 y 7 m	5 y 8 m	5 y 9 m	5 y 10m	5 y 11 m
Boys	11	11	11	12	11	11
Girls	9	11	11	14	12	10
Total	20	22	22	26	23	21

The study sample was heterogeneous, with the majority of study participants indicating English as their first language (25.4%), followed by Setswana (17.9%), Sepedi (9.0%), Northern Sesotho (6.7%), isiZulu (6.7%), Afrikaans (6.0%), five other South African languages totalling 14.9%, and foreign languages (e.g., French, Italian, Shona, Chinese) being the first language of 13.4% of the children. This distribution was slightly different to official statistics³⁰ for the region, although it should be noted that the official figures were from the 2011 census, and represented the entire population, not only five-year-olds.

Differences in the DTVP-2 scores among boys and girls of the SA sample

Owing to differential developmental milestones in children, the best gender comparison is considered to be based on standard scores (SSs), which would allow for developmental differences. Since the American norms might not be entirely suitable for a sample of the SA population, *Table III* on page 55 presents a comparison of both raw scores (RSs) and SSs for the boys and girls. Comparisons were calculated by means of Student's t-tests.

Boys and girls obtained similar SSs on all the subtests of the DTVP-2 when using the prescribed scoring method, with the exception of FG where a statistically significant difference ($p=0.03$) in both RSs and SSs was evident, with girls obtaining on average one SS higher than boys. When the stop rule was not implemented in the adapted method, scores for both groups increased, but boys' scores increased slightly more, so that boys and girls obtained similar RSs, SSs and composite scores for all subtests, which showed no statistically significant differences.

Although literature is available to describe the DTVP-2 test performance for SA children^{8,18,19,20}, no gender comparisons have been reported that could be compared with the results from this study. Even though the DTVP-2 was designed to minimise gender bias¹⁷ findings in our study are at least partially consistent with a Hong Kong study in which girls perform significantly better than boys in the FG subtest of the DTVP-2²¹. It could therefore be concluded that the DTVP-2 might be relatively free of gender bias in

Table III: Comparison of gender differences in the DTVP-2's raw scores (RS) and standard scores (SS), scored according to the prescribed and adapted scoring methods

DTVP-2 subtests and composites	Score	Prescribed method				Adapted method			
		Means		Comparison		Means		Comparison	
		Boys	Girls	t	p	Boys	Girls	t	p
Position in space	RS	12.06	11.33	0.94	0.35	15.96	14.81	1.45	0.15
	SS	9.17	8.94	0.62	0.54	10.91	10.41	1.39	0.17
Figure-ground	RS	10.19	11.36	-2.17	0.03*	11.59	12.30	-1.70	0.09
	SS	10.19	11.36	-2.17	0.03*	11.59	12.30	-1.70	0.09
Visual closure	RS	4.92	4.71	0.52	0.60	8.14	7.74	0.96	0.34
	SS	7.79	7.64	0.39	0.69	10.57	10.23	0.96	0.34
Form constancy	RS	11.23	10.80	1.30	0.20	12.39	11.86	1.73	0.09
	SS	12.00	11.70	1.45	0.15	12.74	12.38	1.63	0.10
MRP	–	39.14	39.64	-0.48	0.64	45.80	45.32	0.48	0.63
MRPQ	–	98.54	99.38	-0.48	0.63	109.66	108.90	0.46	0.65
Eye-hand coordination	RS	138.35	140.32	-0.57	0.57	–	–	–	–
	SS	10.43	10.68	-0.66	0.51	–	–	–	–
Copying	RS	14.20	15.17	-1.34	0.18	–	–	–	–
	SS	9.22	9.51	-1.25	0.21	–	–	–	–
Spatial relation	RS	29.55	29.22	0.19	0.85	–	–	–	–
	SS	11.90	11.90	-0.38	0.71	–	–	–	–
Visual-motor speed	RS	4.94	5.64	-1.06	0.29	–	–	–	–
	SS	9.94	10.33	-1.09	0.28	–	–	–	–
VMI	–	41.28	42.42	-1.11	0.27	–	–	–	–
VMIQ	–	102.11	103.97	-1.09	0.28	–	–	–	–
GVP	–	80.42	82.06	-0.92	0.36	87.08	87.74	-0.37	0.71
GVPQ	–	100.42	101.81	-0.90	0.37	106.20	106.77	-0.36	0.72

MRP = motor-reduced perception; MRPQ = motor-reduced perceptual quotient; VMI = visual-motor integration; VMIQ = visual-motor integration quotient; GVP = general visual perception; GVPQ = general visual perceptual quotient; *p≤0.05 statistically significant

this specific age group in a SA population, with the exception of FG (under the assumption that Occupational therapists will use the DTVP-2 according to the prescribed scoring method).

The DTVP-2 raw scores of the SA sample in comparison to the American normative sample

The RS of each subtest is calculated as the sum of the items answered correctly¹¹. Table IV on page 56 presents the DTVP-2 subtests' mean RSs of the study sample in comparison to the mean RSs of the 5-year-old American normative sample, tested with a single sample t-test for the difference in the sample mean obtained in this study, using the American normative mean as the hypothesised mean. The final columns of the table show paired t-test values for the comparison between the prescribed and adapted scoring methods. The RS were used for this table, as they give the best indication of how the SA sample compared to the normative US sample, without having the interpretive filter of normalisation applied. It should be noted, though, that RSs are not to be used in the interpretation of test results. Because of the different number of items and different scoring for each subtest, a RS on one subtest cannot be compared to a RS on another subtest for the same individual, but comparisons amongst RSs for different samples on the same subtest remain valid comparisons, if only for illustrative purposes.

Although all the t-test results were statistically significant (bearing in mind that for VMS, the significance was less than 5% but considerably more than 1%), this underscores the difference between statistical significance and clinical meaningfulness. For ex-

ample, with reference to PS, it has been noted that a RS difference of one less for the SA sample still equates to the same equivalent SS for that age interval. The difference is thus statistically significant, but not clinically meaningful. Nonetheless, it is evident that the SA sample also obtained statistically significantly higher scores than the American norm sample for EH, FG, SR, and FC, all of which also appeared to be clinically meaningful. Although the single point difference in RS for VMS does not appear considerable, the VMS subtest consists of only eight items, which equates to one SS less for the SA sample when compared to the American norm sample. We observed that the SA sample performed worse on CO and even worse on VC than the American norm.

The SA sample's DTVP-2 subtest' standard score distribution scored according to the prescribed method along with the American norm distribution

Raw scores are converted into SSs (with a mean of 10 and an SD of 3) for clinical interpretation using tables in the DTVP-2 Examiner's Manual. Accordingly, SSs are descriptively rated as average for SSs between 8 and 12, below average for SSs of 7 and less, and above average for SSs of 13 and more¹¹. The SS distributions of the SA sample are presented in Figure 1.

We found notably large numbers of participants with similar scores, which was evident in the third quartile (3Q) scores that have been highlighted in the chart. If, for example, 35% of the respondents obtained a SS of 10 for EH, the second quartile (2Q) collapses to a single point [the median is 10, and the entire inter-quartile range



Table IV: Comparison of the American normative sample and the SA study sample's mean raw scores scored according to the prescribed and adapted scoring methods

DTVP-2 subtests	US norm scores		SA sample		Comparison of RSs		Prescribed-adapted RS difference		
	Mean RS ^a	Corresponding SS ^b	Mean RS ^c (95%CI)	Corresponding SS	t	p	X _{diff}	t	p
Motor-enhanced subtests									
Eye-hand coordination	126	9	139 (135.9–142.8)	10	7.71	<0.0001	–	–	–
Copying	16	10	15 (14.0–15.4)	9	-3.56	0.0005	–	–	–
Spatial relations	21	9	29 (27.6–31.2)	10	9.25	<0.0001	–	–	–
Visual-motor speed	6	11	5 (4.6–6.0)	10	-2.12	0.0355	–	–	–
Motor-reduced subtests									
Prescribed ^d			12 (10.9–12.5)	9	-3.40	0.0009			
Position in space	13	9					3.55	20.89	<0.0001
Adapted ^d			15 (14.6–15.8)	10	7.35	<0.0001			
Prescribed			11 (10.2–11.3)	11	6.51	<0.0001			
Figure-ground	9	9					1.16	10.20	<0.0001
Adapted			12 (11.5–12.4)	12	13.88	<0.0001			
Prescribed	7	10	5 (4.4–5.2)	8	-10.73	<0.0001	3.12	22.99	<0.0001
Visual closure				10	4.48	<0.0001			
Adapted			8 (7.5–8.3)						
Prescribed			11 (10.7–11.3)	12	18.01	<0.0001			
Form Constancy	8	10					1.10	11.94	<0.0001
Adapted			12 (11.8–12.4)	12	26.68	<0.0001			

^aHammill, Pearson and Voress¹¹ only provide mean raw scores for the entire 5-year old interval (p. 42)
^bStandard scores are reflect specifically for the 5y6m–5y11m values, as per p. 55 of Hammill, Pearson and Voress¹¹
^cValues rounded to the nearest integer since the DTVP-2 reports rounded values;
^dprescribed and adapted methods applicable only to the SA sample

(IQR) is 10–11]. Similar effects were seen with CO, FC adapted score (FCa), and FC prescribed score (FCp), where the 3Q is represented by a single point, which in itself was not extremely problematic. What was more of an issue were those subtests where the 2Qs and 3Qs lay outside the average range and showed major deviations from the American norm sample. The SA sample obtained much higher than normative scores for SR, FG when using the prescribed scoring method, and even more so when using the adapted scoring method. Form constancy skewed high, but within the norm, for the prescribed scoring method, and well above the norm for the adapted scoring method, with 59.7% of the sample scoring ≥ 12 . Only PS and VC appeared to fare worse than the American norms, although this might again be attributed to problems with the item linearity which render the stop rule inappropriate, as both of these subtests only appeared so when using the prescribed scoring method, but when the adapted scoring method was used, the scores were well within the norms. (See Figure 1 on page 57).

It should be noted at this point that these differences need to be put into context. Firstly, in the chapter in interpreting the DTVP-2

scores, the DTVP-2 authors¹¹ note that “tests don’t diagnose”^{11:27}. They then refer to the fifth chapter of the manual about the test’s reliability and then state that the reliability results show that the test’s “results may be interpreted with confidence.” However, in that chapter on the DTVP-2 reliability, the authors then provide SEMs for each 1-year age band which “can be used to estimate a confidence interval that surrounds a particular test score” (p. 33), providing a concrete example to demonstrate the principle. It should be noted that the SEMs for all subtests of the DTVP-2 for the 5-year age band are 1, and thus each child being tested will have a true score that can be assumed, with 95% certainty, to be within 2 SSs of their test score (i.e., test score ± 2). However, it is the experience of the authors that occupational therapists in practice would seldom read this additional information in the chapter on the test’s reliability, as this guideline is not stated in the chapter on the interpretation of the test scores, and even if they did, occupational therapists in practice would seldom be inclined to allow such a large margin on the interpretation of a child’s scores, given the interpretive bands supplied by the test authors on p. 24

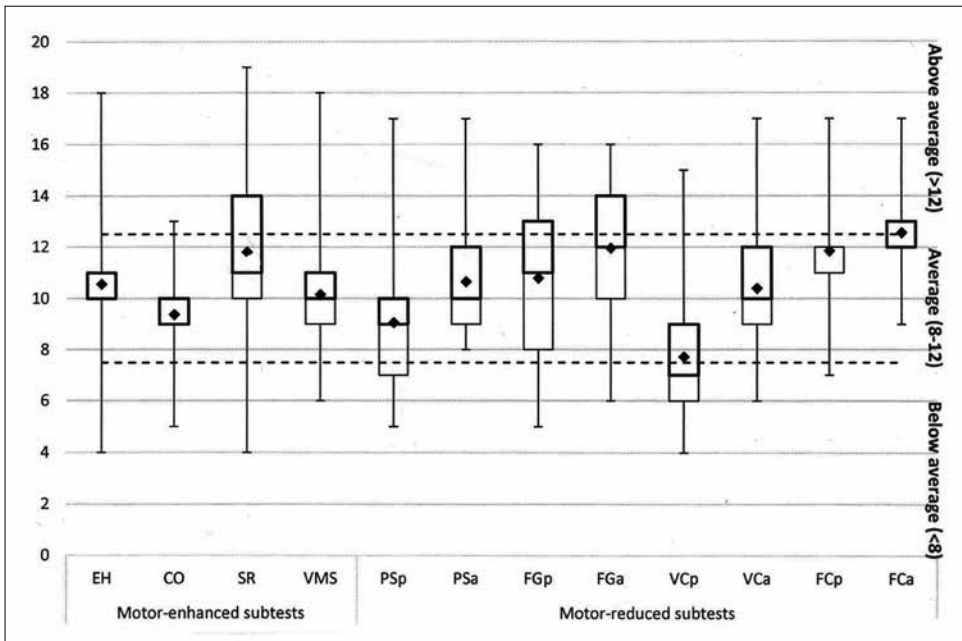


Figure 1: DTVP-2 South African SSs plotted on American norms (3Q box highlighted)

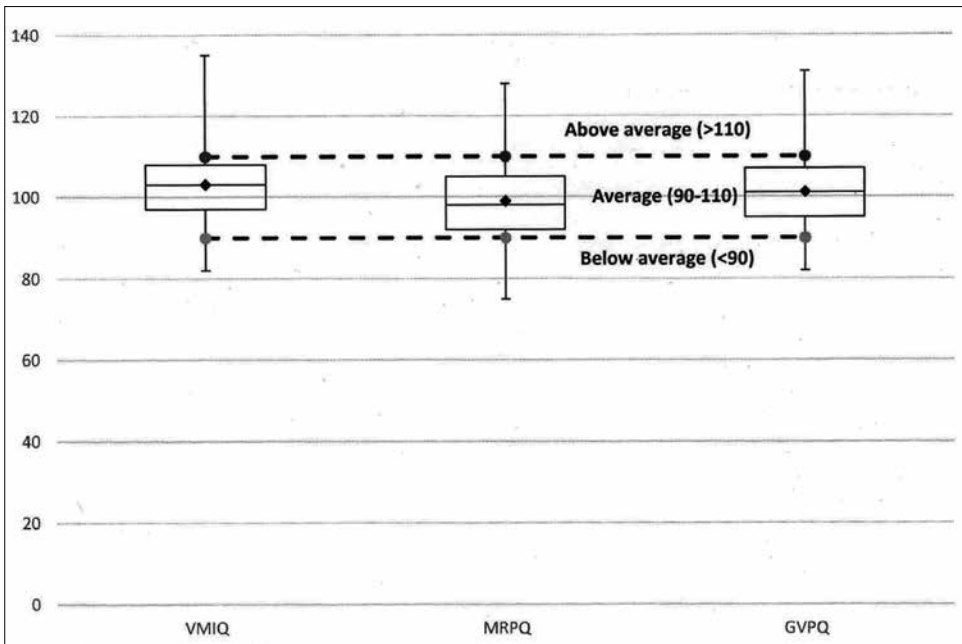


Figure 2: DTVP-2 South African Quotient SSs

(for example, a child in this age group scoring 6, which is indicated on p. 24 as being the bottom score of the “below average” range, should be considered to have a 95% probability of also having a score as high as 8, which is the bottom of the “average” range. Few occupational therapists would be willing to be so lenient with the interpretation of test scores in today’s climate, although it can be argued that they should. This, together with the demonstrated deviation from the norm in the sample, gives caution to the rigid interpretation of DTVP-2 test scores.

However, it should further be noted that occupational therapists should not be over-reliant on individual subtest scores, as the DTVP-2 authors^{11,24} indicated that the “composite quotients are the most reliable DTVP-2 scores. Here, occupational therapists are advised to consider scores in the range 90–110 as average. Furthermore, the SEMs for the quotient scores (p. 34) indicate that a child’s true score may be seen, with 95% certainty, to be within 4 SSs of their quotient scores for MRPQ and VMIQ, and within 2 SSs for

GVPQ. Furthermore, Figure 2 shows that, for the quotient scores, the South African sample showed a distribution that is well within the bounds set by the authors.

The findings of the study thus appear to be consistent with findings in both South African^{8,18,19,20} and international literature^{2,15,21,22,31}, where it was speculated that children from other cultures might perform differently to American children. Differences could occur either because of cultural artefacts, or because of different sociocultural contexts where children are exposed to certain tasks at different ages in different cultures, or simply because the items have been practiced and learnt through similar task exposure.

Some specific recommendations for some subtests could be made, although it should be noted that many of these would need to be confirmed through further investigation and may have been addressed in the DTVP-3. However, the authors of this study find themselves walking a delicate tightrope across the conundrum presented by the two versions of the DTVP: While it must be conceded that the DTVP-3 should be a superior test to the DTVP-2, it cannot be recommended that clinicians switch to the DTVP-3 as a means to avoiding these problems in particular, until the DTVP-3 has been subjected to similar testing and has proved to eliminate these problems on a South African sample of children. The authors would thus not want to create the opinion that they are opposed to clinicians “upgrading” to the DTVP-3, but want to caution clinicians to note that it has not yet been demonstrated that the DTVP-3 is not also prone to these biases in a South African sample. Concerns raised here should thus at least be given consideration when children are tested with the DTVP, regardless of which version. However, the antipode to this is that the researchers do not

want to make clinicians unnecessarily suspicious of the DTVP-3, because neither has it been demonstrated to be prone to these biases. In the end, more research in this area, specifically now with the DTVP-3, is urgently needed. Furthermore, because the DTVP-3 no longer includes the PS, SR and VMS subtests, occupational therapists in practice may still prefer to use the DTVP-2, which would make these recommendations still valid.

Regarding the individual subtests: Copying could also benefit from more specific administration instructions, detailed scoring guidelines and detailed scoring practice. Currently a 2, 1, or 0 score is awarded depending on the quality of the drawing¹¹. Looking at the quality of examples provided by the authors, there seems to be a slight difference between awarding a 2 and a 1, but a larger difference between awarding a 1 and a 0. As a result, clinicians in practice could easily award a 2 score for a 1-score-quality drawing. This tendency was observed when the scores of the CO subtest items were validated by third-party occupational therapists, as therapists differed regarding the allocation of marks - some thera-

pists would score more-, and others less, stringently. The authors of the DTVP-2, however, did improve this shortcoming with the DTVP-3, by publishing the CO scoring template¹⁷. Therefore, the simplest recommendation would be to use the latest version of the DTVP, the DTVP-3.

The scoring criteria for SR are also a point of concern. The subtest is scored by awarding one point per dot that the pencil mark touches for reasonable attempts to copy the design. No points are subtracted for incorrect dots touched. It is therefore recommended that research be conducted to examine whether points should be deducted for incorrect dots touched, since spatial relations are defined as the ability to apply direction to objects, as well as to see the relation between objects²⁸. Therefore, touching an incorrect dot influences the direction of the line.

The authors concur with Cheung et al²¹ who questioned the scoring instructions for VMS, and who recommend that the administration instruction should be revised.

The motor-reduced subtests of the DTVP-2 are discussed with reference to both *Figure 1* and *Table IV*, which show the DTVP-2's motor-reduced subtests' mean prescribed and adapted RSs of the SA sample in comparison to the mean RSs of the 5-year-old American normative sample and the SSs of the American normative sample from the same age band (5y6m–5y11m). It should be remembered that the *purpose* of a stop rule is to reduce test fatigue for the testee and also reduce the administrative burden for the testor. The *assumption* underlying the stop rule is that, given a monotonic increase in the difficulty of items, when a testee has not succeeded in a certain number of items, then that testee should also not succeed in performing the remaining items, and the test can be stopped without materially impacting on the testee's score. Conversely, having the testee complete all items from that point will delay the completion of the test, but should not change the outcome of the test. The rationale behind the adaptation to the scoring of the motor-reduced subtests examined in this study arose from previous studies on the DTVP-2. Richmond and Holland²⁰ proposed that the DTVP-2 might possibly over-identify VC deficits and queried the item linearity of the VC subtest. Visser et al⁸ concurred, and recommended exclusion of the prescribed stop rule improve the consistency between subtests and the validity of the motor-reduced subtests results in the South African population. In all the motor-reduced subtests, when scored according to the adapted scoring method of allowing the children to complete all the items, the SA sample obtained higher mean RSs, indicating that the SA sample was able to complete numerous additional items as the subtests progressed to the last item. The participants consequently scored significantly higher than what their score would have been had the prescribed stop rule been implemented. In the case of PS, FG and VC, the adapted method resulted in an average of one or two SSs higher for the sample, which in turn might have resulted in under- or over-diagnoses of visual perceptual difficulties. When children are eligible for therapy, these scores guide occupational therapists in clinical decision making, intervention planning and tracking of the child's progress. It is therefore imperative that the scores provide a true reflection of the child's capabilities.

These results reported here serve as a further impetus (in addition to earlier findings^{8,18,19,20}) to question the item difficulty and subsequent item linearity of the motor-reduced subtests. In summary, the motor-reduced subtests of the DTVP-2 vary significantly from the norms reported in the DTVP-2 Examiner's Manual. Similarities, differences and correlations are apparent in reports from Thailand¹⁵, Hong Kong²¹, Australia²² and Canada³¹. The scores display an abnormal distribution. It also became evident that the FG, VC and FC test items appear to have distorted difficulty levels. An item analysis of the motor-reduced items was also conducted as part of the larger study, but these results are beyond the scope of this article. Occupational therapists in practice are thus recommended to give far more weight to the quotient scores in their interpretation than to individual subtest scores.

Limitation, implications for practice and recommendations

The limitations associated with this study were, firstly, the study population that was relatively small, on one age band of DTVP-2, and not representative of the whole of SA. Secondly, the population was selected from an urban area, geographically restricted to the City of Tshwane, consisting of children who were proficient in English and attending formal educational settings. Thirdly, no differentiation other than language has been made between cultural/ethnic groups within this relatively small English-speaking sample. Lastly, although the DTVP-3 was published in 2014 this study was conducted when the DTVP-2 was still widely used by SA Occupational therapists. It must also be noted that not all Occupational therapists might have the funding for regular upgrading of measuring instruments, and/or might not choose the 2014 DTVP-3 version since certain subtests were omitted, and may continue to use the DTVP-2. The findings of this study thus still remain germane.

Considering these limitations, implications for practice are that these results may not be generalised to similar SA populations. Care must be taken when interpreting and conveying scores to parents/caregivers and other healthcare professionals.

Based on the results and conclusions drawn from this study, the authors make the following recommendations with regard to future research, whether initiated by occupational therapists in the clinical field, at universities or by the OT association of SA (OTASA):

- ❖ investigation of the suitability of the DTVP-3 norms on a representative sample of SA children;
- ❖ investigation of the suitability of other available and commonly used visual perceptual instruments such as the Beery VMI and the TVPS-R;
- ❖ concurrent validity studies of the DTVP-2 with the DTVP-3 in order to compare norms, as well as the measurement properties of these two respective tests in SA; and
- ❖ development of visual perceptual measurement instruments standardised on the SA population.

Recommendations in terms of practice include the following:

- ❖ to train undergraduate students in the use of a variety of instruments and the limitations of the non-context-specific instruments such as the DTVP-2; and
- ❖ to do appropriate translation (such as conceptual equivalent translation) of the prescribed instructions of measuring instruments used, into the SA official languages, rather than developing SA specific versions.

CONCLUSIONS

Results of this study provide evidence of differences in DTVP-2 scores between boys and girls, how the scores of a small SA study sample within a specific age band (children aged 5 years and 6 months to 5 years and 11 months), compared to those of the American normative sample, and how the prescribed scoring method compares to the adapted scoring method for the motor-reduced subtest cluster. Although the DTVP-2 is considered a comprehensive measurement instrument, certain subtests for this specific age band of the test seems to be problematic in their current form (norms and/or item linearity).

While no context-specific visual perceptual measuring instrument is currently available in SA, occupational therapists should thoughtfully consider the methods in which they use these tests, and perhaps ask themselves the following questions: Do I interpret and use the results of the DTVP-2 with care? Do children need to be assessed with an appropriately translated version of the instructions? Can I contribute to the development of a context-specific, culturally suitable visual perceptual measuring instrument for the assessment of children's important occupational performance components and skills such as visual perception in SA?

ACKNOWLEDGEMENTS

Dr. Daleen Struwig, Faculty of Health Sciences, UFS, is acknowledged for editorial and technical preparation of the manuscript. The authors would like to thank a reviewer for pointing out the use of the SEM in interpreting individual scores, as recommended in the DTVP2 manual.

REFERENCES

1. American Occupational Therapy Association. Framework: Domain and Process, 3rd ed. American Journal of Occupational Therapy. 2014; 68(Suppl1): S1-S48.
2. Brown T, Rodger S, Davis A. Factor structure of the four motor-free scales of the Development Test of Visual Perception, second edition (DTVP-2). American Journal of Occupational Therapy. 2008; 62(5):502-513.
3. Richardson PK. Use of standardised tests in pediatric practice. In Case-Smith J (ed). Occupational Therapy for Children. 6th ed. St. Louis, MO: Elsevier/Mosby, 2010: 216-243.
4. Thorley M, Lim SM. Considerations for occupational therapy assessment for indigenous children in Australia. Australian Occupational Therapy Journal. 2011; 58(1): 3-10.
5. Clutten SC. The development of a visual perception test for learners in the foundation phase. MEd Research Report. Pretoria: University of South Africa; 2009.
6. Eksteen T. The evaluation of the reliability of the Motor-free Visual Perceptual Test (third edition) when translated into Afrikaans, on an Afrikaans first language urban population (east of Pretoria, South Africa) aged 8 years 0 months to 8 years 11 months. MSTSH Research Report. Johannesburg: University of Witwatersrand; 2007.
7. Rens Z. The standardisation of the Beery-Buktenica Developmental Test of Visual-Motor Integration with Supplemental Developmental Tests of Visual Perception and Motor Coordination (4th Edition, Revised, 1997) on an Eastern Cape population aged 7 years 0 months to 7 years 3 months. MSTSH Research Report. Johannesburg: University of Witwatersrand; 2008.
8. Visser MM, Cronje M, Kemp B, Scholtz M, Van Rooyen W, Nel M. The DTVP-2 visual closure subtest: a closer look. South African Journal of Occupational Therapy. 2012; 42(2): 21-25.
9. Van der Merwe J, Smit N, Vlok B. A survey to investigate how South African Occupational Therapists in private practice are assessing and treating poor handwriting in foundation phase learners: Part 1 demographics and assessment practices. South African Journal of Occupational Therapy. 2011; 41(3): 3-11.
10. Beery KE, Buktenica NA, Beery NA. Developmental Test of Visual-Motor Integration. 5th ed. Minneapolis, MN: Pearson Assessments; 2004.
11. Hammill DD, Pearson NA, Voress JK. Developmental Test of Visual Perception. 2nd ed. Texas: Pro-Ed Inc.; 1993.
12. Gardner MF. Test of Visual Perceptual Skills (Non-Motor) - Revised. Hydesville, CA: Psychological and Educational Publications; 1996.
13. Ayres AJ. Sensory Integration and Praxis Tests. Los Angeles, CA: Western Psychological Services; 1989.
14. Harris DB. Children's Drawings as Measures of Intellectual Maturity: A Revision and Extension of The Goodenough Draw-A-Man Test. Harcourt, NY: Brace & World Inc.; 1963.
15. Guntayong C, Chinchai S, Pongsaksri M, Vittayakorn S. Determination of normative values of the Developmental Test of Visual Perception (DTVP-2) in Thai children. International Journal of Medicine and Pharmaceutical Sciences. 2013; 3(2): 113-126.
16. Burtner PA, Bordegaray J, Moedl D, Roe RJ, Savage AR, Wilhite C. Critical review of visual perceptual tests frequently administered by pediatric therapists. Physical and Occupational Therapy in Pediatrics. 1997; 17(3): 39-61.
17. Hammill DD, Pearson NA, Voress JK. Developmental Test of Visual Perception. 3rd ed. Texas: Pro-Ed Inc.; 2013.
18. Van Romburgh JA. Die voorkoms van visuele-persepsieprobleme en die effektiwiteit van Arbeidsterapie groepbehandeling onder Gr. 1 kleurlingkleuters. MOT Research Report. Bloemfontein: University of the Free State; 2006.
19. Visser MM. The association of an omitted crawling milestone on pencil grasp and control in a 5 and 6 year old population. MSTSH Research Report. Johannesburg: University of Witwatersrand; 2004.
20. Richmond J, Holland K. Correlating the Developmental Test of Visual Perception-2 (DTVP-2) and the Test of Visual Perceptual Skills-Revised (TVPS-R) as assessment tools for learners with difficulties. South African Journal of Occupational Therapy. 2011; 41(1): 33-37.
21. Cheung PP, Poon M, Leung M, Wong R. The development test of visual perception-2 normative study on the visual perception function for children in Hong Kong. Physical and Occupational Therapy in Pediatrics. 2005; 25(4): 29-43.
22. Lai MY, Leung FKS. Visual perceptual abilities of Chinese-speaking and English-speaking children. Perceptual and Motor Skills. 2012; 114(2): 433-445.
23. HPCSA (Health Professions Council of South Africa). List of classified and certified psychological tests. (Board Notice 93 OF 15) 2010. <http://www.hpcs.co.za/PBPsychology/Guidelines> (accessed 26 June 2017).
24. Department of Basic Education. Educational Districts: Gauteng 2017. <http://www.education.gov.za/Informationfor/EducationDistricts.aspx> (accessed 30 May 2017).
25. Du Plessis S, Louw B. Challenges to preschool teachers in learners' acquisition of English as language of learning and teaching. South African Journal of Education. 2008; 28: 53-75.
26. Josman N, Abdallah TM, Engel-Yeger B. A comparison of visual-perceptual and visual-motor skills between Palestinian and Israeli children. American Journal of Occupational Therapy. 2006; 60(2): 215-225.
27. Maree K. First Steps in Research. Pretoria: Van Schaik Publishers; 2007.
28. Schneck, CM. Visual perception. In Case-Smith J (ed). Occupational Therapy for Children. 6th ed. St. Louis, MO: Elsevier/Mosby. 2005; 373-403.
29. Nulty DD. The adequacy of response rates to online and paper surveys: what can be done? Assessment & Evaluation in Higher Education. 2008; 33 (3): 301-314.
30. Stats SA. 2012. Census 2011. www.statssa.gov.za. Date of access: 05/01/2015. http://www.statssa.gov.za/census/census_2011/census_products/Census_2011_Census_in_brief.pdf.
31. Brown T, Hockey SC. The validity and reliability of the Developmental Test of Visual Perception – 2nd edition (DTVP-2). Physical and Occupational Therapy in Pediatrics. 2013; 33(4): 426-439.

Corresponding Author

Marieta M Visser

Email: vissermm@ufs.ac.za