



# Borrowing Norms for the Human Figure Drawing Test: how to validate this practice?

PAUL VEDDER, KATRIEN VAN DE VIJFEIJKEN

*Leiden Universiteit, Center for the study of Education and Instruction, PO Box 9555, Leiden, The Netherlands*

HETTY KOOK

*School Inspectorate Rotterdam, Rotterdam, The Netherlands*

**ABSTRACT** *Too often tests are used with clients for whom the validity of the test has not been established. As a case in point we studied the use of the Human Figure Drawing (HFD) test with children living in Curaçao, a small island in the Caribbean. In this community no time and money are available for developing tests and establishing their validity and norms. We suggest that borrowing such information can be a relatively good, inexpensive alternative, provided that clinicians make the best of choices. This paper formulates three requirements, which should be met by the group of clients a clinician is working with. As an example we explored to what extent the requirements are being satisfied by 96 Curaçaoan Grade 4 school children. With regard to these children we conclude that clinicians using the HFD test can best use US representative frequency tables for scoring.*

## INTRODUCTION

The Human Figure Drawing (HFD) test is a popular instrument for the assessment of children's cognitive functioning (Piotrowski *et al.*, 1985). The present study is inspired by an all too common practice in which the HFD is being used without having established its validity or norms for a particular group of clients. In this article we focus on the use of the HFD in a small Caribbean community: Curaçao. It is a 444 km<sup>2</sup> island off the Venezuelan coast with some 145,000 inhabitants. No time and money is available to develop psychological tests or to establish norms and yet decisions have to be taken about, for instance, children's transfer to special education. In the assessment practices leading to these decisions the HFD is one of the instruments. Various studies support the idea of using the quality of the drawing of the human figure as an estimate of intelligence (Koppitz, 1968, 1984; Gayton *et al.*, 1974; Fabry & Berninetti, 1990; Bardos, 1993; Naglieri, 1993), however, not everybody approves of this notion (cf. Scott, 1981; Aikman *et al.*, 1992; Motta *et al.*, 1993).

The primary purpose of the HFD is a rough screening of cognitive functioning. If the scoring results in a classification of low cognitive functioning a more precise assessment is indicated.

In this paper the use of the HFD with Curaçaoan children is taken as an example. This example shows what can be done in a situation in which the best of practices, namely developing a culture-adequate test for cognitive functioning with adapted tasks, new scoring rules and standardization for the specific cultural group, is not a realistic option.

Although our case is restricted to just one instrument and how it is used in a particular region, the analyses and problem solutions presented in this paper have a broader scope. Using the HFD without having established its validity and norms is not only something typical of small and relatively poor countries. The Netherlands is a case in point. Although the HFD has been used frequently in the Netherlands, until 1994 little was known about its validity and no adapted norms were established (van de Vijfeijken, 1994). Furthermore, the problems we are dealing with are not only typical of the HFD. In the Netherlands, as in many other Western countries, clinicians are increasingly being invited or required to test persons with a non-Western cultural background. When testing these clients the clinicians either have to stick to their well-known tests for which the validity has been established, albeit not for this group of new clients, or they try less standardized procedures. These procedures may be experienced as less alienating to the clients, but they actually also lack validity and norms (Te Nijenhuis & Van der Flier, 1999; Van den Berg & Van Leest, 1999).

The popularity of the HFD test partly stems from Goodenough's notion (1926) of the drawing of the human figure as a culture-free product of the child's mind. Di Leo (1973) contends that probably no other test than the Draw a Man test developed by Goodenough comes so close to the ideal of a culture-fair, universally applicable intelligence test.

Di Leo forgot to say that what he contends is largely an assumption. As we shall clarify shortly, there is little reason to assume the cross-cultural validity of the HFD or of another test. Cross-cultural validity has to be established (see Harris, 1968). Once it is established it can be seen as the fulfillment of a requirement which is needed if borrowing norms is a necessary, non-expensive alternative to constructing ones own norms.

### *Drawing as a Culture Specific Activity*

A multitude of studies have shown that the meaning of human figure drawing differs between cultures (Schuster, 1990; Golomb, 1992; Cox, 1993). Drawing is not equally important in every culture. The extent to which it is practiced and drawing materials are available varies considerably. Pfeffer & Oluwu (1986) show that children's educational background has an important influence on their drawing. The conventions for drawing the human figure vary between cultures. In the Western world the head for instance is mostly a representation of mouth, nose and eyes enclosed in a circle or oval like contour. We see this as the best correspondence with

reality. Cox (1993), however, shows that other cultures may use other conventions. In some cultures the figures seem to have been constructed in a list or chain-like way. The parts of the face are drawn below each other without any attention being given to the proportions or the representation of the actual distance between the parts. Cox (1993) formulates the following conclusion: '... we cannot make evaluations of children's human figure drawings in different cultures using criteria—such as those which underpin the Draw-a-Man test—developed in one culture but not necessarily appropriate in another' (p. 108). We fully agree with Cox that rules for scoring a human figure drawing cannot be the same in all cultures. As Harris (1968, p. 133) suggests, the possibilities for using HFDs for measuring cognitive functioning have to be worked out empirically within each culture group. On the basis of Golomb's (1992) work we formulate three requirements which have to be fulfilled in order to be able to use human figure drawing as a measure for cognitive functioning.

1. Within a cultural group the graphic representation of objects and processes has to be alike for all members. If this were not be the case, then within that culture it would not make sense to use one common set of rules for scoring the drawings. Such comparability within a cultural group seems to depend on education, on everyday contacts with common two-dimensional representations of the human figure and on the availability of materials, room and time needed for drawing.
2. Corresponding to the children's age the quality of human figure drawing has to develop steadily towards a particular criterion, which can be a combination of a correspondence to reality and attention to detail.
3. It has to be shown that within a cultural group the quality of human figure drawing is related to levels of cognitive functioning. This requirement clearly departs from the claim made by Harris (1968, p. 133) that although the HFD should not be used for comparing children across cultures, it may still rank children within a culture according to relative intellectual maturity. We explicitly suggest that the correspondence has to be shown. This is a requirement that deals with the test's predictive validity. In this study we used the Raven Standard Progressive Matrices to explore the correspondence between children's HFD score and their level of cognitive functioning. Once the relation is confirmed it makes sense to establish or borrow norms for particular age groups which link the quality of human figure drawing to children's levels of cognitive functioning.

Whether or not the rules for scoring and the norms are interchangeable between cultures, and therefore can be borrowed, is a matter of comparability of the cultures as regards the manner in and the extent to which the two cultures meet the requirements. Indeed, two different cultures may meet all requirements and yet be so different in rules for drawing and qualities of drawing which are appreciated most that borrowing norms between the cultures would not be justified. That is why we suggest that one should not only compare the extent to which the requirements are met, but also the content of human figure drawing between the cultures. In terms

of validity, and more generally, it is not just a test's predictive validity that matters but, particularly for cross-cultural purposes, foremost its construct validity.

### *The Problem*

In the following we will explore whether or not the requirements are met for Curaçaoan school children and compare them primarily with children in the USA and The Netherlands, two countries to which clinicians in Curaçao are strongly oriented, because of their training and access to test information.

## METHODS

### *Subjects*

Forty-nine girls and 47 boys, all Grade 4 primary school children, participated in the study. On average the children were slightly older than 10.5 years ( $128.6 \pm 7.5$  months). Most of the children (53) were 10 years old, while 33 were 11 years old. These 11-year-old children already had repeated a grade in their relatively short school career. This percentage of grade repetition is normal for Curaçao (Vedder, 1995). About 90% of the children's parents had less than 11 years of formal education.

### *Instruments*

*The HFD.* Koppitz (1968) developed a procedure for assessing the presence or absence of various items in human figure drawing that could be expected to appear in HFDs of 5- to 12-year-old children. She developed a scoring system for the HFD as a test of mental maturity. Reliability studies (Koppitz, 1968; Snyder & Gaston, 1970; Fabry & Bertinetti, 1990; van de Vijfeijken, 1994; Abell *et al.*, 1996) indicate that the inter-rater reliability commonly exceeds 0.90. Several studies showed low to high correlations between children's HFD score and their intelligence score. Gayton *et al.* (1974) used the WISC and found a correlation of 0.68. Abell *et al.* (1996) compared the scores between HFD, the WISC and the Stanford-Binet test. Between the HFD and the other measures the correlations were maximally 0.29. The authors did not indicate in their article whether they used raw or age-corrected scores in their analyses. Normally the latter type scores would result in lower correlations than raw scores. Van de Vijfeijken (1994) did a study with 2085 primary school children in which she found a correlation of 0.75 between the (raw) HFD scores and (raw) scores on Raven's Colored Progressive Matrices.

The Raven Standard Progressive Matrices (SPM) was used to assess children's level of cognitive functioning. This test has been primarily designed as a measure of Spearman's *g* factor or general intelligence (Raven *et al.*, 1988). The scale consists of 60 problems divided into five sets of 12. Between and within sets the problems become progressively more difficult. Retest reliability and internal consistency are

generally good. Correlations with other intelligence tests range between 0.40 and 0.75 (Anastasi & Urbina, 1997).

### *Procedure*

The HFD test and Raven's Standard Progressive Matrices were administered as group tests. Some classes rearranged the positions of chairs and tables a little to make sure that children were working individually. Children first took the Raven and then the HFD test. While Dutch is the language of instruction in Curaçaoan primary schools, the tests were administered in Papiamentu, the children's home language and the language which they understand best.

For the HFD each child receives a blank sheet of A4 paper and a HB pencil with an eraser. The instruction is quite simple: 'On this piece of paper, I would like you to draw a whole person. It can be any kind of a person you want to draw, just make sure that it is a whole person and not a stick figure or a cartoon figure' (Koppitz, 1968, p. 6).

### *Scoring*

For scoring the HFD we used a Dutch translation of Koppitz's (1968) scoring manual. Koppitz distinguishes 27 characteristics, which are listed in Table I. We calculated two scores: a raw score and a standardized score. For each of 27 characteristics or details of the drawing children receive a point when it is present. The sum of these details is the total raw score. For the standardized score Koppitz (1968) suggests using country- or culture-specific age- and group-dependent representative frequency tables. We used frequency tables for the USA (Koppitz, 1968) and for The Netherlands (Van de Vijfeijken, 1994), which are available for age groups (nine, 10 and 11 years). These contain four broad frequency categories. If a detail is drawn in more than 85% of the drawings of a representative age group this detail or item is seen as an *expected* item. Items that are drawn in 51–85% of all drawings are *common* items. The third category refers to the *non-exceptional* items, which refer to details present in 16–50% of the drawings. The final category is for the *exceptional* items, which show up in < 16% of all drawings. Only the expected and exceptional items are used for calculating a score for cognitive functioning. If an HFD misses an expected detail the item is scored  $-1$ . The exceptional details which are present are scored  $+1$ . The resulting scores are summed and the outcome is subtracted from 5 to avoid negative scores. A HFD of a 10-year-old Dutch child which misses hair and a mouth (two expected items) but which represents elbows (an exceptional item) gets a score of  $5 - 2 + 1 = 4$ . For this cultural group and age groups this corresponds, according to Koppitz, to an average or slightly below average level of cognitive functioning.

With Raven's SPM each correctly answered item is one point. We transferred children's raw scores using different norm systems. We used the US national norms (percentile scores) from 1986 (Raven, 1990, RS3.SPM6), percentile scores for Puerto Rican children from 1977 (Raven & Court, 1989, RS4.14) and provisional

TABLE I. The items scored in Koppitz's Human Figure Drawing test

1. head	10. hair	19. fingers
2. eyes	11. neck	20. five fingers
3. pupils	12. body	21. legs
4. eyebrow	13. arms	22. legs two-dimensional
5. nose	14. arms two-dimensional	23. knee
6. nostrils	15. arm pointing down	24. feet
7. mouth	16. arm attached to shoulders	25. feet two-dimensional
8. two lips	17. elbows	26. profile
9. ear	18. hands	27. proportions

Dutch norms based on Van de Vijfeijken (1994). The children in this latter study took Raven's Colored Progressive Matrices. For the sake of comparability the scores on this latter version of the Raven were converted to SPM scores, which is possible because the two versions share items (Andrich and Daws, cited in Raven & Court, 1989). We used the percentile scores to convert the raw scores into scores representing broad categories of intellectual functioning, which were defined by Raven *et al.* (1988) as follows:

- I. 'Intellectually superior', > 95th percentile.
- II. 'Definitely above the average in intellectual capacity', 76–95th percentile.
- III. 'Intellectually average', 26–75th percentile.
- IV. 'Definitely below average in intellectual capacity', 6–25th percentile.
- V. 'Intellectually impaired', < 6th percentile.

How many Curaçaoan children are included in each category depends on the norms or percentile scores which were used.

In the introduction we have already explained why we chose to use Dutch and American norms. We used Puerto Rican norms because Puerto Rico is one of the few countries in the Caribbean for which norms have been established. The assumption is that, both being Caribbean countries, Curaçao and Puerto Rico have a lot in common in cultural respects.

All 96 HFDs included in this study were scored using the Dutch translation of Koppitz's (1968) scoring manual. A trained judge who was blind to any identifying details of the child, such as name, grade, date of birth or gender, rated the HFDs

RESULTS

*Requirement 1: Within a Cultural Group the Graphic Representation of Objects and Processes has to be Alike for all Members*

Curaçao is a former colony of The Netherlands. The Dutch captured the islands in the 17th century. It became an important trade centre specializing in trading slaves from the African West Coast. Nowadays the descendants of the slaves form the largest group of inhabitants (some 80%). Since about 1930 the Curaçaoan school system has been comparable with the Dutch school system. This means that most

TABLE II. Mean raw HFD scores of Curaçaoan children presented by age group and sex

Age	Boys ( $n = 47$ )	Girls ( $n = 49$ )	Boys + girls ( $n = 96$ )
9 years	$20.2 \pm 1.6$	$19.0 \pm 2.0$	$19.6 \pm 1.8$
10 years	$18.9 \pm 2.9$	$19.6 \pm 2.3$	$19.3 \pm 2.6$
11 years	$19.5 \pm 2.3$	$19.7 \pm 2.5$	$19.6 \pm 2.4$

parents of Curaçaoan children went to Western type primary schools. They were taught how to draw and what was proper drawing and now teach their own children how to draw and how to represent the human figure using paper and pencil. Of course, they are supported in this effort by the children's kindergarten or school. In short they learn to draw like most children in, for instance, the USA and The Netherlands do. Although, as compared with the USA or The Netherlands, Curaçao is a relatively poor country, the availability of drawing materials is not really a serious problem. As a consequence of this attention to drawing the Curaçaoan community, in particular Curaçaoan children, meet the first requirement. Of course this is also apparent from the children's HFDs.

The children drew an average of 19.4 details in their HFD. The standard deviation is rather low (2.4). Table II presents the results for boys and girls.

As can be seen, 9-year-old boys draw more details than do 10- and 11-year-old boys. For girls the number of details increases slightly with age. No significant differences were found either between age groups or between boys and girls.

The differences in drawn details were compared ( $\chi^2$  test or Fisher's exact test) for the three age groups and for boys and girls separately between Curaçaoan children and children who participated in the Dutch study of van de Vijfeijken (1994). For the 9-year-old children we found no differences. At the age of 10 significantly ( $p < 0.05$ ) fewer Curaçaoan children drew arms attached to the shoulders, elbows, hands, five fingers, knees and the correct proportions than Dutch children. At the age of 11 this was the case for hands, fingers and five fingers. The children from the two cultures also differed with respect to characteristics of the HFD which were not scored. Curaçaoan children more often drew bare feet, whereas Dutch children more often drew shoes. The hairstyle also differed between countries. Curaçaoan children drew persons with an afro hairstyle, whereas hair drawn by Dutch children was more often straight.

*Requirement 2: Parallel to an Age Increase Children have to Develop the Quality of Human Figure Drawing towards a better Correspondence to Reality and more Attention for Details*

The information used to determine that Curaçaoan children meet the first requirement can also be used as evidence that they meet the second requirement. Their education resembles the education of American and European children for whom we know that the quality of drawing progresses with age towards the criterion used.

TABLE III. Children’s mean score on Raven’s SPM by age and sex

Age	Boys ( <i>n</i> = 47)	Girls ( <i>n</i> = 49)	Boys + girls ( <i>n</i> = 96)
9 years	29.6 ± 15.1	16.4 ± 5.9	23.0 ± 12.9
10 years	28.9 ± 10.2	29.0 ± 10.5	29.0 ± 10.3
11 years	30.1 ± 9.0	28.9 ± 10.9	29.6 ± 9.7

The data presented in Table II, however, seem to contradict this conclusion. Nine-, 10- and 11-year-old children did not differ significantly in the number of details drawn. Our sample, however, does not allow us to conclude that there is no progress corresponding to increasing age. All children were in Grade 4. An important part of the age difference is due to grade retention. Assuming that the HFD is a measure of cognitive functioning one might expect that these children would not draw more details than their younger classmates. Therefore we maintain our assumption that the Curaçaoan children meet the second requirement.

*Requirement 3: It has to be Shown that within a Cultural Group the Quality of Human Figure Drawing is Related to Levels of Cognitive Functioning Raw scores*

We used the Raven SPM as a criterion for quality of the HFD. Table III presents the children’s raw scores by age and by sex.

No significant differences were found either between boys and girls or between age groups. The 9-year-old girls had remarkably low scores, as compared with the 9-year-old boys or the 10-year-old girls, but there were only five girls. The 11-year-old children were mostly children who had repeated a grade. For this reason the fact that they did not achieve better results than the 10-year-old children should not really surprise us.

The correlation for Curaçaoan children between the raw scores for the HFD and for Raven’s SPM was 0.45; for girls 0.44 and for boys 0.49. Given the slightly strange Raven scores of the 9-year-old children (10 children) and the fact that most children aged 11 (33) had repeated a grade, we decided to calculate the correlation between HFD and intelligence score only for the 10-year-old children. The resulting correlation was a little higher (0.53); 0.44 for boys and 0.65 for girls. These are modest values. Nevertheless, as compared with earlier reported correlations in other studies the correlations we found were certainly not low, given the fact that in this study we considered only one year group in one grade, whereas subjects in the other studies came from several age groups and grades.

*Standardized scores.* Using norms with the HFD lowers the correlations, because the norms reduce the range of scores. Moreover, using norms implies a correction for age. Variance due to age differences is dispensed with, in order to be able to compare age groups. However, norms are indispensable for giving meaning to individual children’s scores. This being said, we continued our exploration to find



TABLE IV. Mean standardized HFD scores calculated using US and Dutch frequency tables

Age	Mean standardized scores	
	US frequency tables	Dutch frequency tables
9 years	$4.9 \pm 1.0$	$4.4 \pm 1.3$
10 years	$4.3 \pm 1.0$	$4.0 \pm 1.4$
11 years	$4.6 \pm 0.7$	$3.0 \pm 1.8$
All	$4.5 \pm 0.9$	$3.7 \pm 1.6$

out what norms can best be used to optimize the correlation between HFD scores and scores on the Raven SPM.

As mentioned before, we have scored the Curaçaoan HFDs using American and Dutch frequency tables for expected and exceptional items. The resulting scores give, according to Koppitz, an indication of the level of children's cognitive functioning. Table IV presents the mean scores for 9-, 10- and 11-year-old children.

Table IV shows that for none of the age groups did the average score reach 5, the average score in the norm groups. This means that the Curaçaoan children generally had lower scores for their HFD than either the American or Dutch children. The difference with the Dutch children is larger than with the American children. Since we know too little about the meaning of the test task in different cultures, the differences in scores cannot be simply interpreted in terms of higher or lower levels of cognitive functioning (cf. Harris, 1968).

We compared the children's raw Raven scores with available norms from the USA, The Netherlands and Puerto Rico and sorted the scores into the five broad categories of intellectual functioning distinguished by Raven *et al.* (1988) (see Methods, subsection scoring). The outcomes are presented in Table V.

Table V shows that it makes quite a difference which norms are used. When using the American norms 43.8% of all children scored in the upper three categories, meaning that they had average to above average intelligence, while when using the Dutch norms no more than 19.8% of the children were evaluated as

TABLE V. Distribution of Raven SPM scores of Curaçaoan children using norm tables for the US, the Netherlands and Puerto Rico

Grade	US		Netherlands		Puerto Rico	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
I	3	3.1			7	7.3
II	4	4.2	2	2.1	22	22.9
III	35	36.5	17	17.7	48	50.0
IV	35	36.5	29	30.2	12	12.5
V	19	19.8	48	50.0	7	7.3

TABLE VI. Correlations between HFD and Raven SPM using three norm tables and HFD Dutch and US frequency tables with 96 primary school children

Raven norms	HFD	
	US frequency tables	Dutch frequency tables
US	0.44 <sup>a</sup>	0.35 <sup>a</sup>
The Netherlands	0.35 <sup>a</sup>	0.23
Puerto Rico	0.42 <sup>a</sup>	0.37 <sup>a</sup>

<sup>a</sup>  $p < 0.001$ .

having average to above average intelligence. With these norms 50% of all participating Curaçaoan primary school children would be seen as intellectually impaired, while these children were enrolled in normal primary schools. Clearly the most balanced distribution resulted from using the Puerto Rican norms.

Table VI presents Pearson p.m. correlations between HFD scores and Raven SPM scores using a variety of frequency and norm tables.

Table VI shows that a combination of the US frequency tables for the HFD and either US or Puerto Rican norms for the Raven SPM yields the highest correlation.

If we restrict our analyses to the 10-year-old Curaçaoan children the correlations were a little higher, as can be seen in Table VII.

Again, the US frequency tables for the HFD lead to a slightly higher correlation with the Raven SPM than the Dutch tables.

DISCUSSION

In this article an example is given of what can be done if diagnosticians want to use a test or another instrument for which no culture- or group-adequate norms are available and for which no group-specific validity information is available. It is pointed out that borrowing this information only makes sense if it has been shown that a group for which no validity and norms have been established sufficiently resembles a group for which the test validity has been explored and

TABLE VII. Correlations between HFD and Raven SPM using three norm tables and HFD Dutch and US frequency tables with 53 10-year-old primary school children

Raven norms	HFD	
	US frequency tables	Dutch frequency tables
US	0.52 <sup>a</sup>	0.44 <sup>b</sup>
The Netherlands	0.40 <sup>b</sup>	0.32
Puerto Rico	0.50 <sup>a</sup>	0.48 <sup>a</sup>

<sup>a</sup>  $p < 0.001$ ; <sup>b</sup>  $p < 0.01$ .

norms have been established. The groups should resemble each other in terms of skills and knowledge of rules important for task accomplishment. Of course these depend on the assessment instrument under study. As regards the HFD we present evidence that Curaçaoan children share a common way of graphically representing objects and processes in drawings (the first requirement). For the second requirement it was more difficult to show that Curaçaoan children meet it. Parallel to an age increase children have to develop the quality of human figure drawing towards a better correspondence with reality and more attention to detail. The human figures drawn by the children participating in our study did not show an increase in number of drawn details from 9 to 11 years. Golomb (1992) has described a study by Dennis to clarify that in some cultures children's drawing stops developing by the age of 7 years. If we assume that these children's cognitive development proceeds beyond that age, then it is clear that human figure drawing cannot be used as a measure of children's cognitive functioning in these cultures. We argue that our findings should not be taken as an indication that in Curaçao the development of drawing stops when children are 9 years old. Our findings can probably be attributed to the fact that the study was restricted to Grade 4 children.

We have also shown that the quality of children's HFD is positively correlated with their scores on Raven's SPM. Our conclusion is that a combination of the original US frequency tables (Koppitz, 1986) for the HFD and either US norms or Puerto Rican percentile scores for the Raven SPM lead to the highest correlations. This means that the US frequency tables are the best for scoring the HFDs of Curaçaoan children.

The limitations of the approach followed should be clear. The approach does not change the cultural adequacy of a test; neither of the explored test nor of the criterion test. For Curaçao it is still unclear whether it is possible to develop a test which predicts cognitive functioning or success in or out of school better than either the HFD test or Raven's SPM (cf. Sternberg *et al.*, 1995). To paraphrase Triandis (1994, p. 58), one should not be so interested in knowing how well they do *our* tricks as in knowing whether or not they do *their* tricks well. What is important to them in terms of cognitive functioning should be the important question.

## REFERENCES

- ABELL, S.C. VON BRIESEN, P. & WATZ, L.S. (1996). Intellectual evaluations of children using human figure drawings: an empirical investigation of two methods. *Journal of Clinical Psychology*, 52, 67-75.
- AIKMAN, K.G., BELTER, R.W., FINCH, A.J. (1992). Human figure drawings: validity in assessing intellectual level and academic achievement. *Journal of Clinical Psychology*, 48, 114-120.
- ANASTASI, A. & URBINA, S. (1997). *Psychological testing*. Englewood Cliffs, NJ, Prentice Hall.
- BARDOS, A.N. (1993). Human figure drawings: abusing the abused. *School Psychology Quarterly*, 8, 177-181.
- COX, M. (1993). *Children's Drawings of the Human Figure*. Hillsdale: LEA.
- DI LEO, J.H. (1973). *Children's Drawings as Diagnostic Aids*. New York, NY: Brunner/Mazel.
- FABRY, J.L. & BERTINETTI, J.F. (1990). A construct validation study of the human figure drawing test. *Perceptual and Motor Skills*, 70, 465-466.
- GAYTON, W.F., TAVORMINA, J.S. & EVANS, H.E. (1974). Comparative validity of Harris and Koppitz scoring systems for human figure drawings. *Perceptual and Motor Skills*, 39, 369-370.

- GOLOMB, C. (1992). *The Child's Creation of the Pictorial World*. Berkeley, CA: University of California Press.
- GOODENOUGH, F.L. (1926). *Measurement of Intelligence by Drawings*. New York, NY: Harcourt Brace & World.
- HARRIS, D.B. (1968). *Children's Drawings as Measures of Intellectual Maturity*. New York, NY: Harcourt, Brace & World.
- KOPPITZ, E.M. (1968). *Psychological Evaluation of Children's Human Figure Drawings*. New York, NY: Grune & Stratton.
- KOPPITZ, E.M. (1984). *Psychological Evaluation of Human Figure Drawings by Middle School Pupils*. New York: Grune & Stratton.
- MOTTA, R.W., LITTLE, S.G. & TOBIN, M.I. (1993). The use and abuse of human figure drawings. *School Psychology Quarterly*, 8, 162-169.
- NAGLIERI, J.A. (1993). Human figure drawings in perspective. *School Psychology Quarterly*, 8, 170-176.
- PFEFFER, K. & OLOWU, A. (1986). Effects of socioeconomic differences on the sophistication of Nigerian children's human figure drawings. *Perceptual and Motor Skills*, 62, 771-774.
- PIOTROWSKI, C., SHERRY, D. & KELLER, J.W. (1985). Psychodiagnostic test usage: a survey of the society of personality assessment. *Journal of Personality Assessment*, 49, 115-120.
- RAVEN, J. (1990). *American and International Norms*, Raven Manual Research Supplement 3. Oxford: Oxford Psychologists Press.
- RAVEN, J. & COURT, J.H. (1989). *Normative, Reliability and Validity Studies. References*. London, H.K. Lewis & Co.
- RAVEN, J.C., COURT, J.H. & RAVEN, J. (1988). *Standard Progressive Matrices*. London, H.K. Lewis & Co.
- SCOTT, L.H. (1981). Measuring intelligence with the Goodenough-Harris drawing test. *Psychological Bulletin*, 89, 483-505.
- SCHUSTER, M. (1990). *Die Psychologie der Kinderzeichnung* [The psychology of children's drawings]. Berlin: Springer Verlag.
- SNYDER, R.T. & GASTON, D.S. (1970). The figure of the first grade child: item analyses and comparison with Koppitz norms. *Journal of Clinical Psychology*, 26, 377-383.
- STERNBERG, R.J., WAGNER, R.K., WILLIAMS, W. & HORVATH, J. (1995). Testing common sense. *American Psychologist*, 50, 912-927.
- TE NIJENHUIS, J. & VAN DER FLIER, H. (1999). Bias research in the Netherlands. *European Journal of Psychological Assessment*, 15, 165-175.
- TRIANDIS, H. (1994). *Culture and Social Behavior*. New York, NY: McGraw-Hill.
- VAN DEN BERG, R.H. & VAN LEEST, P.F. (1999). Praktisch testgebruik bij allochtonen [Practical test use with immigrants]. *De Psycholoog*, 34, 256-260.
- VAN DE VIJFEIJKEN, K. (1994). *Normering menstekening voor 4-12 jarige kinderen* [HFD norms for 4-12 year old children]. Leiden: Leiden University Press.
- VEDDER, P. (1995). *Antilliaanse kinderen. Taal, opvoeding en onderwijs op de Antillen en in Nederland* [Antillean children. Language, child rearing, and education]. Utrecht: Jan van Arkel).