

Deterioration and Recovery of ~~Draw-A-Person~~ IQ Scores in the Repeated Assessment of the Naglieri Draw-A-Person Test in 6- to 12-Year-Old Children

AQ: au

Christiane Lange-Küttner
London Metropolitan University

Enno Küttner
Black Country ~~NHS~~ Partnership Foundation Trust, West
Bromwich, England

Marta Chromekova
Care+ Ltd., London, England

The study investigated whether mental age in children, as assessed by the IQ in the Draw-A-Person (DAP) test (Naglieri, 1988), can be improved by practice. In addition, it was tested whether children needed novel content to keep up their performance level during test repetition. The DAP test was given to 6-, 8-, 10-, and 12-year-old children ($N = 80$) 3 times. In addition, they drew a police figure 3 times, with task sequence counterbalanced. Repeated drawings resulted in significant omission of detail and deterioration of scores, but the novel task instruction temporarily recovered the IQ scores. This did not occur in the reverse sequence of the tasks, with the less specified DAP instruction given in the 2nd half of the drawing series. Furthermore, structural regressions in the human figure drawing could be observed in individual cases. However, 12-year-olds did not need the external introduction of novel content to maintain their IQ score. Correlations showed that this age group redefined their drawing plan on each repetition, and often created unique figures on each occasion, even if the instruction had stayed the same.

AQ: 1

Keywords: Naglieri DAP ~~test~~-IQ scores, mental age, updating, deterioration, regression

The distinction between mental age and chronological age is an important differentiation in developmental psychology that allows clinicians and school psychologists alike to assess whether a child is achieving below or above his or her age according to his or her date of birth. This study investigates whether mental age can be improved by practice. Although retest reliability of intelligence tests typically assesses whether a score is stable across repeated sessions over days, weeks, or months, we were interested in whether children can improve their mental age within a single session. We used the IQ score of the Naglieri (1988) Draw-A-Person Test (DAP) as measurement, as it is one of the most widely and internationally used screening test of children's intelligence. Beyond the mere repetition, we also introduced novel content into the test instruction in order to make the task more interesting to children.

The Naglieri DAP Test

The drawing of the human figure is an early, if not the earliest, activity of children when they begin to use a pencil (Campbell, Duncan, Harrison, & Mathewson, 2008; Chan & Louie, 1992; Cox, 1993; Cox & Parkin, 1986; Freeman, 1975; Goodnow, 1977; Lange-Küttner, 2009; Silk & Thomas, 1988). Because drawing the human figure is and always has been such a widespread activity in children, Goodenough (Goodenough, 1926; Goodenough & Harris, 1950) pioneered a human figure drawing test where a score was calculated based on the amount of detail that is specified by the child. The human figure drawing test was found to correlate moderately to highly with IQ and cognitive development until adolescence (Abell, Horkheimer, & Nguyen, 1998; Abell, Von Briesen, & Watz, 1996; Abell, Wood, & Liebman, 2001; Holtzman, 1993; Naglieri, 1993; Sisto, 2000). The human figure drawing test is still worldwide in use in order to diagnose intellectual development and retardation in children (e.g., Bacchini, Amodeo, Vajro, & Licenziati, 2003; Bombi, 1995; Cannoni, 1993; Chappell & Steitz, 1993; Cox & Cotgreave, 1996; Dandii, 2002; Dennis, 1960; Goldman & Gilbert, 1992; Golomb, 1977; Jing, Yuan, & Liu, 1999; Kay, 1989; Kifune & Tachibana, 1991; La Voy et al., 2001; Loxton, Mostert, & Moffatt, 2006; Martlew & Connolly, 1996; Naglieri & Pfeiffer, 1992; Richter, Griesel, & Wortley, 1989; Taborda de Velasco, 1993; Vedder, van de Vijfeijken, & Kook, 2000). Norms were updated by Harris (1963) and in the Naglieri (1988) DAP test. Socioeconomic differences in children's background were found to be not relevant for the human figure drawing test (Golomb, 1977; Willcock, Imuta, & Hayne, 2011).

Christiane Lange-Küttner, Faculty of Life Sciences and Computing, School of Psychology, London Metropolitan University, London, England; Enno Küttner, Black Country ~~NHS~~ Partnership Foundation Trust, West Bromwich, England; Marta Chromekova, Neuropsychiatric Rehabilitation Unit, Care+ Ltd., London, England.

We are grateful to children and staff of the independent Hall Primary and Secondary School in Putney Vale and Wimbledon, London.

Correspondence concerning this article should be addressed to Christiane Lange-Küttner, Faculty of Life Sciences and Computing, School of Psychology, London Metropolitan University, 166-220 Holloway Road, London N7 8DB, England. E-mail: c.langekuettner@londonmet.ac.uk

AQ: 2

The validity of human figure drawing test scores is debated (Gresham, 1993; Motta, Little, & Tobin, 1993a, 1993b), but only with respect to behavioral risk assessment using particular features of the human figure as indicators of disturbed emotions (Aldridge et al., 2004; Bekhit, Thomas, & Jolley, 2005; Lange-Küttner, 1989; ter Laak, de Goede, Aleva, & van Rijswijk, 2005; Thomas & Jolley, 1998). In young children, drawing recognizable human figures requires the ability to follow simple response rules (Riggs, Jolley, & Simpson, 2013). Accordingly, in older children, correlations of children's human figure drawing test scores with the Stanford-Binet IQ were above .60 (Terman & Merrill, 1960), and with the Wechsler Intelligence Scale for Children between .60 and .80 (Wechsler, 1967). The DAP showed higher correlations with the Performance IQ than with the Verbal IQ (Fabry & Bertinetti, 1990). Hence, the DAP test was evaluated as "a useful addition to the test battery" (Wisniewski & Naglieri, 1989, p. 346) and as an appropriate screening test for intellectual ability in children (Prewett, Bardos, & Naglieri, 1989).

There can be a relatively high number of false positives (children with an IQ above 110 with poor figure drawings) and a relatively high number of false negatives (children with an IQ beyond 90 with good figure drawings; Willcock et al., 2011). However, especially the existence of the combination of a low IQ test result with high DAP IQ in children was a reason to use the test to identify intellectual potential in culturally different children's populations, such as the Inuit Eskimos in Canada (Wilgosh, 1991).

Retest reliability of the repeated human figure drawing test scores was shown to be relatively high, with .94 for two drawings on two successive days in a sample of 197 seven-year-olds (Goodenough, 1926). Similar high reliability was found in a follow-up study with .94 for two drawings when repeated after 1 week and scored by the same evaluator, and .90 when scored by two evaluators, in a sample of 386 seven- to eight-year-old children (McCarthy, 1944). However, in a larger sample reliability, this high was found only in about two thirds of 2,500 children, whereas in one third of the sample only .50 retest reliability occurred (Kelllogg, 1969). Thus, some researchers do not recommend the human figure drawing test (Aikman, Belter, & Finch, 1992). We investigated whether low reliability could have occurred due to test repetition. Drawing is a productive higher order activity that requires action plans for sequencing the graphic schemata (Freeman, 1980; Thomas, 1995; Thomas & Tsalimi, 1988). Thus in this study it was hypothesized that multiple repetitions of the DAP test within one session could either help (practice) or hinder (load) children to improve on their score, as fatigue can be an important factor in testing (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999, Standard 10.6).

The Effect of Content on the DAP

In addition, we used a modified instruction insofar as children were asked to draw a police person (DAPP). It was expected that the instruction to draw a police person would cue children to draw more specific detail that would result in an improved score. In the DAP scoring manual, there are features such as ears, nose, and mouth, but also more specific features of the human figure drawing such as clothing (Pfeffer & Olowu, 1986). The specification of

novel content such as clothing (e.g., "Draw a person in a swimsuit") can raise children's performance (Lange-Küttner, Kerzmann, & Heckhausen, 2002), and thus it was expected that the novel instruction to draw a person in a uniform should help to maintain interest during repetition.

As in the Wisconsin Card Sorting Test, a change in instruction was introduced halfway through the task (Zelazo, 2006). However, whereas the switch in the Wisconsin Card Sorting Test is implicit (Kirkham, Cruess, & Diamond, 2003; Mack, 2007; Towse, Redbond, Houston-Price, & Cook, 2000), in the current study, the change of the figure drawing instruction was introduced explicitly halfway during the session. Children were required to draw a sequence of three human figures and then three police figures, with half the sample carrying out the two tasks in the reversed sequence. It was predicted that the sequence of the tasks would matter. In the police-then-human figure (DAPP-DAP) sequence, a more elaborate drawing plan is required first. Because the DAP human figure instruction is comparably underspecified and would only constitute a subtotal of the more highly specified DAPP drawing instruction, it was expected that the DAP instruction in the second half of the task would not help children to maintain interest: The more specific police figure drawing plan is set up first and can later easily accommodate a sparser task version of the drawing plan. In contrast, in the human-then-police figure (DAP-DAPP) sequence, a simple drawing plan for the DAP human figure version is set up first, and hence this plan would need updating halfway through the task because the DAPP drawing of a police figure requires more specific detail. Thus, it was predicted that only the DAPP instruction in the second half should refresh interest and activate additional effort.

Method

A sample of 80 children took part in the drawing experiment. Children were tested in an independent (fee-charging) primary and secondary school in Wimbledon, London. The school is nonselective and admits children with all abilities. Children with learning disabilities were schooled separately but shared the school breaks and wore the same uniform. Only the mainstream children were tested. Parents were all from at least upper middle class background, as they could afford to pay the £7,683–£9,780 (approximately \$12,396–\$15,670) annual school fees. School results for mainstream children were about 20% above the national average in the year of the data collection (Walker, 2007). Each age group consisted of 20 children—6-year-olds (range: 5 years 5 months to 6 years 6 months), 8-year-olds (range: 7 years 5 months to 8 years 6 months), 10-year-olds (range: 9 years 4 months to 10 years 6 months), and 12-year-olds (range: 11 years 6 months to 12 years 6 months)—with equal numbers per condition and gender.

Children were tested in their classrooms. There was no time restriction; however, pupils were expected to finish their drawings within one teaching hour. Two school classes were tested per age group. One class of each age group was allocated to one sequence condition, the DAP-DAPP sequence, and the other class to the DAPP-DAP sequence.

Children were drawing on white paper 21 cm × 29.5 cm in size. Use of erasers was not allowed to prevent lengthy corrections. Children received loose blank drawing sheets, not a drawing booklet. Children were drawing only one figure per sheet. They wrote their

name, date of birth, and gender on the back of each drawing sheet. On completion of each human figure, they proceeded to the next drawing without delay. Dates of birth were controlled with the school records and added on the DAP scoring sheet later on.

The long instruction of the DAP (Naglieri, 1988)—“I’d like you to draw some pictures for me. On this page, I’d like you to draw a picture of a man. Make the very best picture you can. Take your time and work very carefully, and I’ll tell you when to stop. Remember, be sure to draw the whole man. Please begin”—was shortened to “Draw a man” for the boys and “Draw a woman” for the girls, without any further explanation, in order to keep instructions clear and precise during the drawing series, which consisted of six drawings in total. Likewise, in the DAPP, the instruction was “Draw a police man” for the boys and “Draw a police woman” for the girls, without any further explanation. The DAP instruction was also modified because we were not interested in the absolutely best product and performance level of the children, but in the process that occurred during repetitions, that is, in children’s improvement, stagnation, or deterioration of the score during the repetition of the test in comparison to their first trial. In the DAP-DAPP sequence, children received first the instruction “Draw a man”/“Draw a woman,” followed by the new instruction after the third drawing, “Draw a police man”/“Draw a police woman,” halfway through the task. These instructions were reversed in the DAPP-DAP sequence.

The scoring system of the Draw A Person: A Quantitative Scoring System (Naglieri, 1988) was used. Scored body parts were arms, attachment, clothing, ears, eyes, feet, fingers, hair, head, legs, mouth, neck, nose, and trunk. Scored were presence, kind of

attachment, detail, and proportion. The maximum raw score was 65. The second author, a consultant clinical child psychologist and National Health Service clinical director, scored all 960 drawings. Scorings of the third author, who had collected the drawings, of a subsample randomly drawn from all age groups amounting to 50% drawings, were used for comparison. Interrater reliability was 94%. Raw scores were transferred into normative IQ scores with the DAP manual (Naglieri, 1988).

Results

The DAP-DAPP and DAPP-DAP sequence condition was a between-subjects factor because one half of the children were drawing the two tasks in one sequence and the other group in the other sequence. Thus, although the task sequence was a between-subjects variable (condition) with two levels, the drawing series was a within-subjects factor with six levels. Data were analyzed with a 4 (age groups) × 2 (condition) × 6 (repetition) analysis of variance with repeated measurements on the last factor. Because the raw scores produced the same results as the Naglieri IQ scores (see plotted means in Figure 1), only the results of the IQ scores are reported. Correlational analysis was used for assessment of common variance between the test scores of the drawing series. When the Mauchley’s test of sphericity was significant, degrees of freedom were adjusted according to Huynh-Feldt.

Group means and standard deviations are listed in Table 1. A main effect of gender showed that girls ($M = 108$) showed a higher DAP IQ than boys ($M = 97$), $F(1, 80) = 21.29, p < .001$,

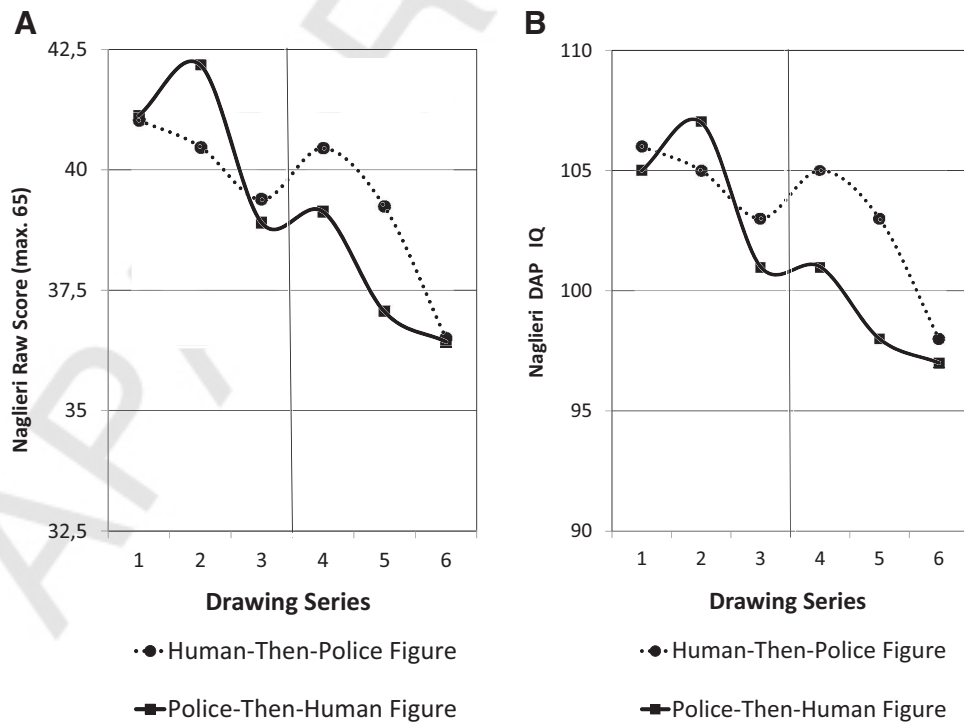


Figure 1. Draw-A-Person (DAP) raw score (A) and DAP IQ score (B) deterioration and recovery. In the police-then-human figure sequence condition (solid line), scores showed only a linear downward trend. In the human-then-police figure sequence condition (broken line), a cubic trend demonstrated that the drawing score temporarily recovered when drawing the more highly specified police figure.

Table 1
Draw-A-Person IQ Scores per Age Group per Condition (N = 80)

Age group (years)	1		2		3		4		5		6	
	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
Human-then-police figure sequence (n = 40)												
Boys												
6	104	18	95	13	95	14	97	11	98	12	97	13
8	98	11	94	10	98	10	97	21	94	13	88	10
10	108	9	105	13	101	16	99	16	98	13	95	7
12	101	25	101	24	92	13	99	9	93	16	83	26
Girls												
6	126	19	125	21	120	21	127	19	119	23	118	12
8	98	20	104	11	104	7	106	15	102	13	104	13
10	112	12	114	5	114	5	116	6	116	13	111	12
12	101	11	101	12	98	17	99	17	100	16	87	16
Police-then-human figure sequence (n = 40)												
Boys												
6	98	13	99	15	97	12	92	10	89	13	88	11
8	101	23	98	21	91	21	89	16	84	14	89	19
10	109	11	105	15	107	13	102	10	94	11	89	5
12	96	13	96	11	95	7	99	10	99	20	98	9
Girls												
6	118	6	122	5	115	4	115	15	111	15	106	15
8	113	10	118	8	99	17	106	7	104	7	101	9
10	110	10	112	7	110	11	107	10	108	10	109	11
12	93	16	104	11	95	14	101	12	97	15	95	8

$\eta^2 = .25$. There was a significant age effect, $F(3, 80) = 4.17, p < .01, \eta^2 = .16$ (6 years, $M = 107$; 8 years, $M = 99$; 10 years, $M = 106$; 12 years, $M = 97$), that interacted two-way with gender, $F(1, 80) = 3.00, p < .05, \eta^2 = .12$. Post hoc tests showed that girls had significantly higher DAP IQ scores than boys at all ages, but not at age 12. Gender was not significant in interactions with repetition ($ps > .11$).

A within-subjects effect of repetition, $F(5, 80) = 10.75, p < .001, \eta^2 = .14$, showed that the DAP IQ deteriorated during the session (Drawing 1, $M = 105$; Drawing 2, $M = 106$; Drawing 3, $M = 102$; Drawing 4, $M = 103$; Drawing 5, $M = 100$; Drawing 6, $M = 97$); however, the extent of the DAP IQ deterioration was dependent on age and sequence, $F(15, 80) = 2.02, p < .05, \eta^2 = .09$. In order to explain this three-way interaction, the same multivariate analysis of variance was run separately for each sequence.

There was a significant repetition effect in the DAP-DAPP sequence, $F(4.29, 40) = 4.10, p < .01, \eta^2 = .11$, with no significant interaction with age, $F(15, 40) = 0.52, ns$. Polynomial contrasts of the repetition effect showed a significant linear, $F(1, 40) = 8.19, p < .01, \eta^2 = .20$, and a significant cubic trend, $F(1, 40) = 4.14, p < .05, \eta^2 = .11$. Hence, the DAP IQ score deteriorated, then significantly recovered halfway through the experiment with the new, more specific instruction, but then deteriorated again (see Figure 1B, broken line).

Also in the DAPP-DAP sequence, the repetition effect was significant, $F(5, 40) = 8.14, p < .001, \eta^2 = .20$. Polynomial contrasts showed that only the linear trend was significant, $F(1, 40) = 21.40, p < .001, \eta^2 = .40$ (see Figure 1B, solid line). Thus, the introduction of a simpler task halfway in the series could not elicit a recovery of the DAP IQ, and a linear deterioration of the scores continued throughout the drawing series. The interaction of

repetition with age was just a trend, $F(15, 40) = 1.63, p = .071, \eta^2 = .13$. The continuous deterioration of drawing scores (solid lines) occurred in the three younger age groups, but not in the 12-year-old children.

Table 2 shows the correlations for the repeated drawing scores throughout the drawing series. Correlations in the DAP-DAPP sequence were high and continuously significant in the younger age groups. This showed that they must have perceived the DAPP as a quasinalatural continuation of the DAP. However, correlations in the 12-year-olds were lower in the drawing series and sometimes did not reach significance. In comparison, the DAPP-DAP

Table 2
Correlations Between Draw-A-Person IQ Scores per Age Group per Condition (N = 80)

Age group (years)	1-2	2-3	3-4	4-5	5-6
Human-then-police figure sequence					
6	.96***	.95***	.81**	.93***	.75**
8	.60	.44	.89***	.95***	.92***
10	.77**	.95***	.86***	.90***	.93***
12	.98***	.66*	.54	.68*	.45
Police-then-human figure sequence					
6	.81**	.89***	.63	.88***	.97***
8	.88***	.54	.62	.95***	.76*
10	.79**	.86**	.56	.52	.87**
12	.39	.16	.45	.36	.37

Note. Correlations between scores when the change in test instruction occurred are set in bold.
 * $p < .05$. ** $p < .01$. *** $p < .001$.

sequence correlations were also relatively high, but in contrast, they were mostly not significant between Drawings 3 and 4, when the new instruction for the less specified task was given. The correlations of the 12-year-old children in this sequence condition were consistently nonsignificant (see Figure 2).

F2,
AQ:3
F3-6

What was the change in the execution of the DAP and DAPP in the 12-year-olds in comparison to a younger age group? Figures 3–6 give

examples of the drawing series, of a 7-year-old girl (Figure 3) and boy (Figure 4), and of a 12-year-old girl (Figure 5) and boy (Figure 6). The drawings of the two 7-year-old children illustrated the correlations in the younger age groups, as figures within each drawing task were very much alike. This was less so the case in the two 12-year-old children (where correlations were low and not significant). They drew different individuals, even if task instructions were identical.

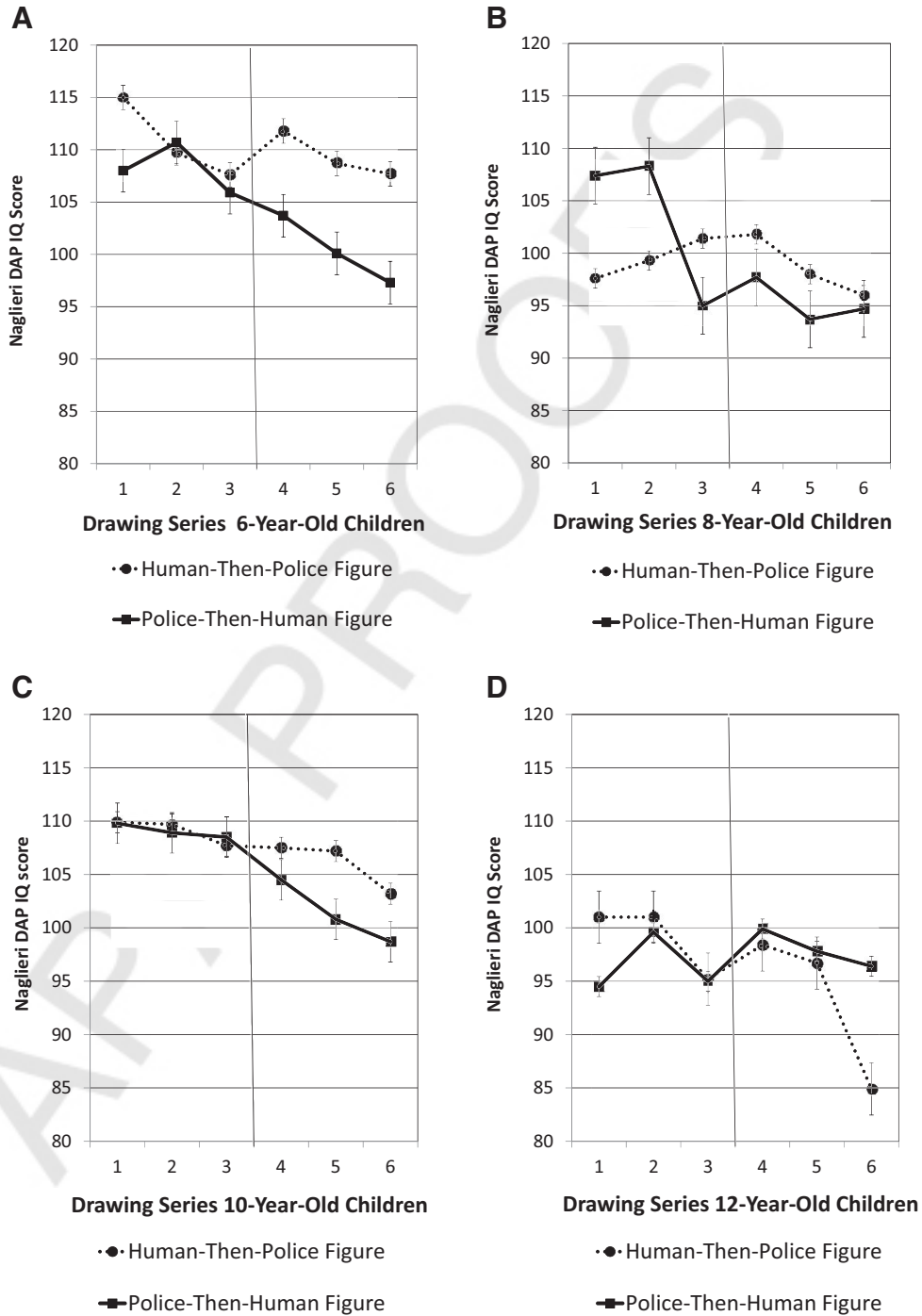


Figure 2. Continuous deterioration of Draw-A-Person (DAP) IQ scores in the police-then-human figure sequence condition occurred in all age groups except in the 12-year-old children.

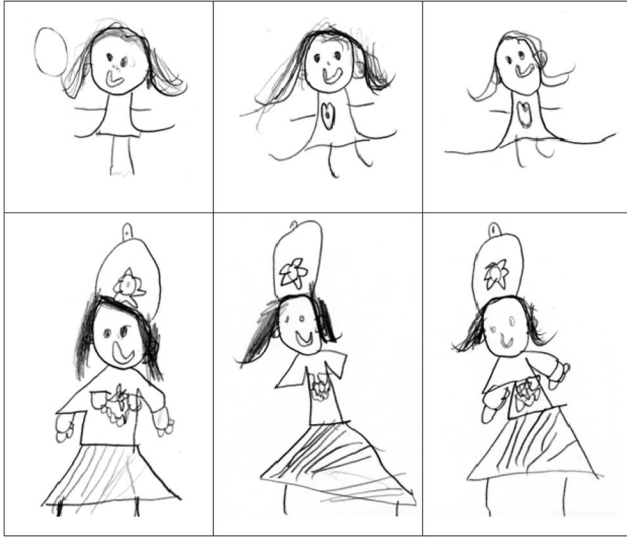


Figure 3. Seven-year-old girl, human-then-police figure sequence. Figure drawings per task instruction were very similar. This showed that the child had formed an action plan to which she was sticking for each instruction. Nevertheless, during the repetition some details were sometimes lost, such as nose or arms. The police figure is clearly more detailed than the plain human figure.

Furthermore, in the 7-year-old boy (Figure 4) a qualitative regression during the drawing series occurred, from drawing a relatively sophisticated and detailed likeness of a male person in the first half of the sequence, to drawing a tadpole figure in the second half of the drawing series. This did not occur in the same way in the drawing series of the 12-year-old boy (Figure 6). In this drawing series, the proportions changed from those of an adult man in the first half of the sequence to those of a more childlike person in the second half of the drawing series. Thus, both a 7-year-old boy and a 12-year-old boy showed a structural regression toward an earlier drawing stage (human-tadpole and adult-child).

Discussion

The study investigated with an extensive within-subjects assessment whether children can improve their DAP raw and IQ score with practice. It is commonly assumed even by experts in the field that practice helps improvement, and in particular, that practice in the DAP test would improve scores (Cox, 2005). However, the current study demonstrated that the opposite was the case: Practice of the DAP test produced loss of details and a score deterioration until the age of 10 years, and not an improvement in scores. Hence, this study on practice of the human figure drawing delivered an important correction of the commonly held belief that practice would help young children improve their drawing skills. However, the new “police” specification in the test instruction halfway in the drawing series recovered the score as predicted.

This suggested that although the DAP drawing task was easy and open-ended, the actual practice with this task tested children’s limits during the multiple repetitions within one session. Practice normally helps children to recover their performance level in the

same session (Lange-Küttner, Averbek, Hirsch, Wießner, & Lamba, 2012). Hence, that practice of the DAP test would lead to deterioration in IQ scores was not to be expected. In one individual case a regression in drawing style occurred even to the tadpole stage, which indicated a mental age several years below the chronological age of 7. Although one would not want to conclude that the deterioration of the DAP IQ scores indicated a true regression in intelligence, the repetitions could potentially produce a sort of apparent pseudodebility. In the majority of cases, however, the deterioration consisted only of some loss of detail that indicated less alertness and less visual attention. The new task instruction was very successful in recovering children’s interest in the task. Children of all ages updated their drawing plan and increased their drawing score after the higher specification police task. The switch to a different drawing task did not lead to a recovery of the same starting level as at the beginning of the drawing series, but to a small, yet significant temporary increase of the scores halfway through the task. In contrast, when the second task had a less specified task instruction than the first task, this did not cue children into drawing a human figure according to a more elaborate action plan, and the drawing score continuously deteriorated in a nearly linear fashion. As in previous research (Lange-Küttner, 2012; Lange-Küttner & Green, 2007), children increased their performance with a more challenging task. In any case, however, the drawing scores of both conditions converged in the last repetition.

Nevertheless, 12-year-old children did not need this halfway instruction to prevent the level of their drawing score from con-

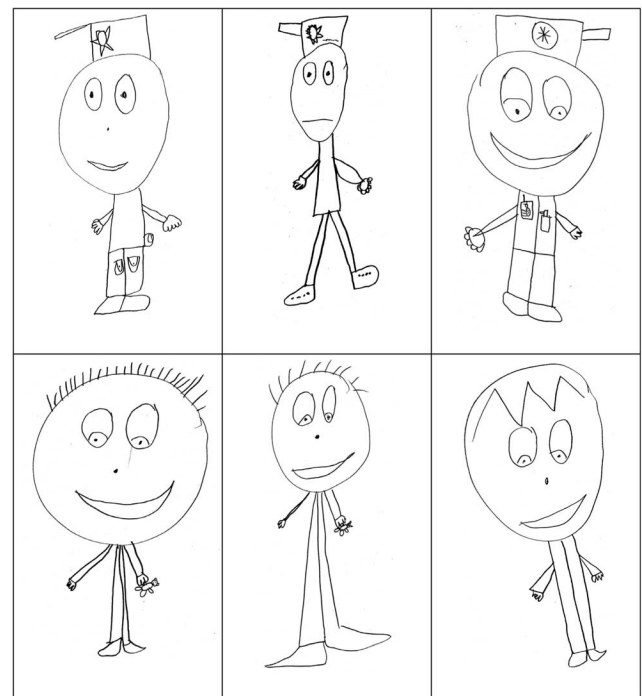


Figure 4. Seven-year-old boy, police-then-human figure sequence. Figure drawings per task instruction were very similar. As with the 7-year-old girl, the police figures were more detailed than the plain human figures, which in fact were tadpole figures equivalent of a much lower mental age.

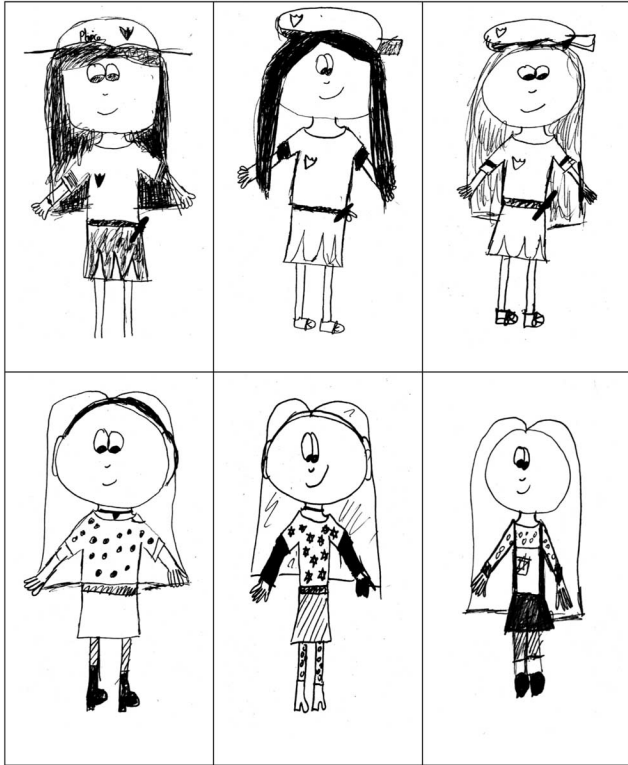


Figure 5. Twelve-year-old girl, police-then-human-figure sequence. Figure drawings per task were individual, for instance, showing different hair colors while wearing a similar police uniform. This showed that the child had formed a separate action plan for each drawing, even if the drawing instructions for the figures were identical. Furthermore, there is not much difference in the amount of detail between the two drawing tasks. This showed that she used an elaborate interpretation of the more sparse draw-a-person figure drawing instruction to enrich her drawings to the same level of detail as before in the higher specified police drawing instruction.

tinuous deterioration. Low, nonsignificant correlations between the events indicated that the 12-year-old children in the DAPP-DAP sequence actually interpreted the task instruction differently on each drawing occasion, and thus would have set up a different action plan for each figure drawing even if the adult’s instruction had stayed the same. Flexibility is also a topic in drawing research since some years (e.g., Barlow, Jolley, White, & Galbraith, 2003; Barrett & Eames, 1996; Berti & Freeman, 1997; Karmiloff-Smith, 1990; Picard & Vinter, 2007; Zhi, Thomas, & Robinson, 1997). The current study suggested that whereas the younger children would accept an identical instruction as a message to produce an identical drawing, the older children avoided to repeat themselves (see also Witt & Vinter, 2011) by not following the verbal instruction in a literal fashion. Hence, the study highlighted that the prevention of deterioration would not occur due to flexible attention to more detail, but with a focus on the production of original, distinguishable, and individual human figures. Originality as well as flexibility and elaboration is the hallmark of creativity (Kaufman, Kaufman, & Lichtenberger, 2011). It needs to be noticed, though, that at the end of the session, the 12-year-old children’s practice had maintained their level of performance, but they still

had not increased scores even at this age, defying the belief that children’s practice makes them perfect.

Regression in drawing is rarely mentioned (e.g., Freeman & Cox, 1985; Golomb, 1992; Goodnow, 1977; Lange-Küttner & Thomas, 1995; Lange-Küttner & Vinter, 2008; van Sommers, 1984; Willats, 2005). So far, in a longitudinal study, regression in drawing could only be observed in terms of sparseness of drawings (Lange-Küttner, 1994). Regression as psychological concept implies that a child who can already produce an age-appropriate human figure would then go back to an earlier drawing stage. In the DAP, only the amount of detail is counted, and drawing stages are not conceptualized. However, also in the details of the DAP, stage-defining criteria are scored (e.g., the classical feature of the tadpole stage is scored as an attachment of the legs directly at the head, with an omission of the trunk). Thus, a regression in the DAP was not defined as just a low score, but as an *emergent loss* during the repeated measurement that did not exist in the child’s output before and produced a large gap between chronological and mental age.

In the current study, it was especially interesting that in individual cases, practice in human figure drawings could produce some true structural regressions during the drawing series. The mental age regression back to the tadpole stage in a 7-year-old boy

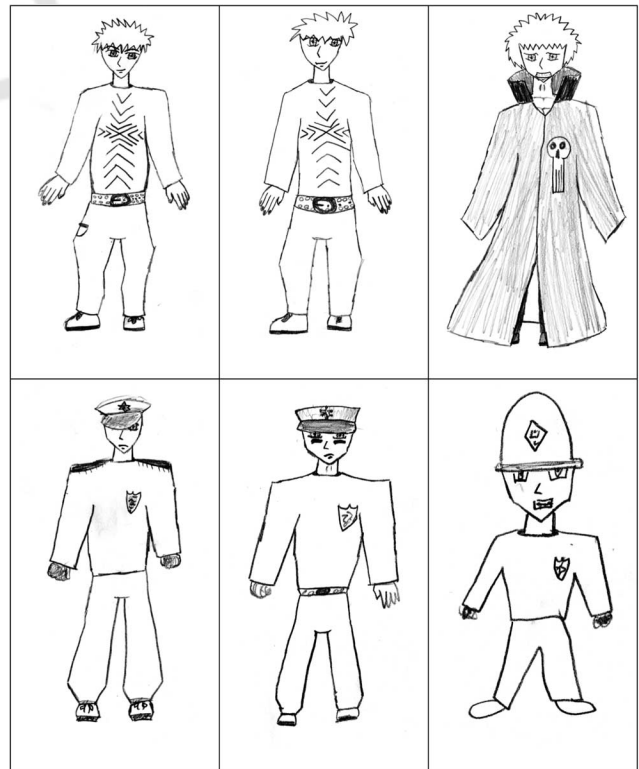


Figure 6. Twelve-year-old boy, human-then-police figure sequence. Figure drawings per task could be completely different, with a figure wearing a long coat instead of T-shirt and trousers. This showed that he had formed a separate updated action plan for the third drawing, even if the drawing instruction was the same. In both draw-a-person and draw-a-police-person tasks, though, the third figure in each sequence showed a loss of detail (e.g., hands were omitted), or the proportion had become more childlike.

involved not only loss of detail, as in almost all children up to the age of 10 in the current study, but also loss of organization of the graphic schemata. In contrast, the structural regression from depicting an adult to depicting a child in the drawing series of a 12-year-old boy may rather be evaluated as a variation in content than in drawing stage, as children at this age can intentionally draw the way they were drawing at earlier stages of their drawing development (Tryphon & Montangero, 1992). Younger children only know how an advanced or an immature drawing looks while lacking the ability to actually carry out the plan for a drawing that is different to their own style (Cox, 2000; Kosslyn, Heldmeyer, & Locklear, 1977).

Last but not least, the study showed significantly higher level DAP IQ scores in girls than in boys at all ages, except at age 12. There was no significant interaction of gender with practice, though. The female advantage in the human figure drawing test is well known since Goodenough (Goodenough, 1926; Goodenough & Harris, 1950; Lange-Küttner et al., 2002; Willcock et al., 2011). Girls also drew plenty of detail when they were copying cubes that had a regular surface pattern, whereas boys focused more on the silhouette (Lange-Küttner, 2011; Lange-Küttner & Ebersbach, 2013). However, Naglieri (1988) did not provide separate standardized scores for girls and boys because, unlike Goodenough and Harris, he did not obtain significant gender differences in a large survey sample of 2,622. But because gender differences may reach significance in smaller empirical samples, it may help future psychological assessments if separate norms for boys and girls were available.

In conclusion, one could say that this study on repetition of the DAP and a modified DAPP version showed that fatigue was an important factor in the assessment of children until age 10 (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999, Standard 10.6), which may explain some previous reports of low retest reliability. The study furthermore showed that the DAP IQ score could not be improved with practice (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999, Standard 1.9), which speaks to the validity of the test scores of this very popular test in the psychological assessment of children.

References

- Abell, S. C., Horkheimer, R., & Nguyen, S. E. (1998). Intellectual evaluations of adolescents via human figure drawings: An empirical comparison of two methods. *Journal of Clinical Psychology, 54*, 811–815. doi:10.1002/(SICI)1097-4679(199810)54:6<811::AID-JCLP8>3.0.CO;2-J
- Abell, S. C., Von Briesen, P. D., & Watz, L. S. (1996). Intellectual evaluations of children using human figure drawings: An empirical investigation of two methods. *Journal of Clinical Psychology, 52*, 67–74. doi:10.1002/(SICI)1097-4679(199601)52:1<67::AID-JCLP9>3.0.CO;2-T
- Abell, S. C., Wood, W., & Liebman, S. J. (2001). Children's human figure drawings as measures of intelligence: The comparative validity of three scoring systems. *Journal of Psychoeducational Assessment, 19*, 204–215. doi:10.1177/073428290101900301
- 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. doi:10.1002/1097-4679(199201)48:1<114::AID-JCLP2270480116>3.0.CO;2-Y
- Aldridge, J., Lamb, M. E., Sternberg, K. J., Orbach, Y., Esplin, P. W., & Bowler, L. (2004). Using a human figure drawing to elicit information from alleged victims of child sexual abuse. *Journal of Consulting and Clinical Psychology, 72*, 304–316. doi:10.1037/0022-006X.72.2.304
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for education and psychological testing*. Washington, DC: American Psychological Association.
- Bacchini, D., Amodeo, A. L., Vajro, S., & Licenziati, M. R. (2003). La valutazione psicologica del bambino obeso attraverso l'utilizzo del Test del Disegno della Figura Umana e del Bender Gestalt Test (The psychological assessment of obese children using the Human Figure Drawing Test and the Bender Gestalt Test). *Psicologia Clinica dello Sviluppo, 7*, 245–267.
- Barlow, C. M., Jolley, R. P., White, D. G., & Galbraith, D. (2003). Rigidity in children's drawings and its relation with representational change. *Journal of Experimental Child Psychology, 86*, 124–152. doi:10.1016/S0022-0965(03)00109-7
- Barrett, M., & Eames, K. (1996). Sequential development in children's human figure drawing. *British Journal of Developmental Psychology, 14*, 219–236. doi:10.1111/j.2044-835X.1996.tb00703.x
- Bekhit, N. S., Thomas, G. V., & Jolley, R. P. (2005). The use of drawing for psychological assessment in Britain: Survey findings. *Psychology and Psychotherapy: Theory, Research and Practice, 78*, 205–217.
- Berti, A. E., & Freeman, N. H. (1997). Representational change in resources for pictorial innovation: A three-component analysis. *Cognitive Development, 12*, 501–522. doi:10.1016/S0885-2014(97)90020-4
- Bombi, A. S. (1995). Rappresentazioni figurative della ricchezza e della povertà nei bambini dai sei agli undici anni (Graphic representations of wealth and poverty in children aged 6–11 yrs.). *Età Evolutiva, 50*, 3–18.
- Campbell, R. N., Duncan, P. A., Harrison, A. L., & Mathewson, L. C. (2008). Style and other factors affecting children's recognition of their own drawings. In C. Lange-Küttner & A. Vinter (Eds.), *Drawing and the non-verbal mind: A life-span perspective* (pp. 63–85). Cambridge, England: Cambridge University Press. doi:10.1017/CBO9780511489730.004
- Cannoni, E. (1993). La rappresentazione di somiglianze e differenze nei disegni infantili di figure umane (The representation of similarities and differences in children's human figure drawings). *Rassegna di Psicologia, 10*, 81–96.
- Chan, L., & Louie, L. (1992). Developmental trend of Chinese preschool children in drawing and writing. *Journal of Research in Childhood Education, 6*, 93–99. doi:10.1080/02568549209594826
- Chappell, P. A., & Steitz, J. A. (1993). Young children's human figure drawings and cognitive development. *Perceptual and Motor Skills, 76*, 611–617. doi:10.2466/pms.1993.76.2.611
- Cox, M. V. (1993). *Children's drawings of the human figure*. Hove, England: Erlbaum.
- Cox, M. V. (2000). Children's diachronic thinking in relation to developmental changes in their drawings of the human figure. *British Journal of Developmental Psychology, 18*, 13–24. doi:10.1348/026151000165544
- Cox, M. V. (2005). *The pictorial world of the child*. Cambridge, England: Cambridge University Press.
- Cox, M. V., & Cotgreave, S. (1996). The human figure drawings of normal children and those with mild learning difficulties. *Educational Psychology, 16*, 433–438. doi:10.1080/0144341960160407
- Cox, M. V., & Parkin, C. E. (1986). Young children's human figure drawing: Cross-sectional and longitudinal studies. *Educational Psychology, 6*, 353–368. doi:10.1080/0144341860060405
- Dandii, O. (2002). Assessment of the mental abilities of special school children in Mongolia. *Japanese Journal of Special Education, 39*, 83–89.
- Dennis, W. (1960). The human figure drawings of Bedouins. *Journal of Social Psychology, 52*, 209–219. doi:10.1080/00224545.1960.9922078

AQ: 4

- Fabry, J. J., & Bertinetti, J. F. (1990). A construct validation study of the Human Figure Drawing Test. *Perceptual and Motor Skills, 70*, 465–466. doi:10.2466/PMS.70.2.465-466
- Freeman, N. H. (1975). Do children draw men with arms coming out of the head? *Nature, 254*, 416–417. doi:10.1038/254416a0
- Freeman, N. H. (1980). *Strategies of representation in young children*. London, England: Academic Press.
- Freeman, N. H., & Cox, M. V. (Eds.). (1985). *Visual order: The nature and development of pictorial representation*. Cambridge, England: Cambridge University Press.
- Goldman, R. K., & Gilbert, D. C. (1992). Prediction of risk in kindergarten children. *Perceptual and Motor Skills, 75*, 1033–1034. doi:10.2466/pms.1992.75.3f.1033
- Golomb, C. (1977). Representational development of the human figure: A look at the neglected variables of SES, IQ, sex, and verbalization. *Journal of Genetic Psychology, 131*, 207–222. doi:10.1080/00221325.1977.10533293
- Golomb, C. (1992). *The child's creation of a pictorial world*. Berkeley, CA: University of California Press.
- Goodenough, F. L. (1926). *Measurement of intelligence by drawings*. Yonkers, NY: World Book.
- Goodenough, F. L., & Harris, D. B. (1950). Studies in the psychology of children's drawings. II 1928–1949. *Psychological Bulletin, 47*, 369–433. doi:10.1037/h0058368
- Goodnow, J. J. (1977). *Children's drawings*. London, England: Fontana.
- Gresham, F. M. (1993). "What's wrong in this picture?": Response to Motta et al.'s review of human figure drawings. *School Psychology Quarterly, 8*, 182–186. doi:10.1037/h0088269
- Harris, D. B. (1963). *Children's drawings as measures of intellectual maturity: A revision and extension of the Goodenough Draw-a-Man Test*. New York, NY: Harcourt, Brace & World.
- Holtzman, W. H. (1993). An unjustified, sweeping indictment by Motta et al. of human figure drawings for assessing psychological functioning. *School Psychology Quarterly, 8*, 189–190. doi:10.1037/h0088270
- Jing, J., Yuan, C., & Liu, J. (1999). Study of human figure drawings in learning disabilities. *Chinese Mental Health Journal, 13*, 133–134.
- Karmiloff-Smith, A. (1990). Constraints on representational change: Evidence from children's drawings. *Cognition, 34*, 57–83. doi:10.1016/0010-0277(90)90031-E
- Kaufman, J. C., Kaufman, S. B., & Lichtenberger, E. O. (2011). Finding creative potential on intelligence tests via divergent production. *Canadian Journal of School Psychology, 26*, 83–106.
- Kay, S. R. (1989). Cognitive battery for differential diagnosis of mental retardation vs. psychosis. *Research in Developmental Disabilities, 10*, 251–260. doi:10.1016/0891-4222(89)90014-0
- Kellogg, R. (1969). *Analyzing children's art*. Palo Alto, CA: Mayfield.
- Kifune, N., & Tachibana, H. (1991). Sex difference in drawings of the human figure by mentally retarded children. *Japanese Journal of Special Education, 29*, 1–6.
- Kirkham, N. Z., Cruess, L., & Diamond, A. (2003). Helping children apply their knowledge to their behavior on a dimension-switching task. *Developmental Science, 6*, 449–467. doi:10.1111/1467-7687.00300
- Kosslyn, S. M., Heldmeyer, K. H., & Locklear, E. P. (1977). Children's drawings as data about their internal representations. *Journal of Experimental Child Psychology, 23*, 191–211. doi:10.1016/0022-0965(77)90099-6
- Lange-Küttner, C. (1989). *Raubegriff und Objektbeziehungen beim Kind* (Space concept and object relations in the child). Frankfurt am Main, Germany: Lang.
- Lange-Küttner, C. (1994). *Gestalt und Konstruktion. Die Entwicklung der grafischen Kompetenz beim Kind* (Gestalt and construction. The development of graphic competence in the child). Bern, Switzerland: Huber.
- Lange-Küttner, C. (2009). Habitual size and projective size: The logic of spatial systems in children's drawings. *Developmental Psychology, 45*, 913–927. doi:10.1037/a0016133
- Lange-Küttner, C. (2011). Sex differences in visual realism in drawings of animate and inanimate objects. *Perceptual and Motor Skills, 113*, 439–453. doi:10.2466/04.10.24.PMS.113.5.439-453
- Lange-Küttner, C. (2012). The importance of reaction times for developmental science: What a difference milliseconds make. *International Journal of Developmental Science, 6*, 51–55.
- Lange-Küttner, C., Averbach, B. B., Hirsch, S. V., Wießner, I., & Lamba, N. (2012). Sequence learning under uncertainty in children: Self-reflection vs. self-assertion. *Frontiers in Psychology, 3*, 127. doi:10.3389/fpsyg.2012.00127
- Lange-Küttner, C., & Ebersbach, M. (2013). Girls in detail, boys in shape: Gender differences when drawing cubes in depth. *British Journal of Psychology, 104*, 413–437. doi:10.1111/bjop.12010
- Lange-Küttner, C., & Green, H. (2007). What is the age of mental rotation? In *Proceedings of the Sixth IEEE International Conference on Development and Learning* (pp. 259–263). Piscataway, NJ: IEEE Press. doi:10.1109/DEVLRN.2007.4354043
- Lange-Küttner, C., Kerzmann, A., & Heckhausen, J. (2002). The emergence of visually realistic contour in the drawing of the human figure. *British Journal of Developmental Psychology, 20*, 439–463. doi:10.1348/026151002320620415
- Lange-Küttner, C., & Thomas, G. V. (Eds.). (1995). *Drawing and looking: Theoretical approaches to pictorial representation in children*. Hemel Hempstead, England: Harvester Wheatsheaf.
- Lange-Küttner, C., & Vinter, A. (Eds.). (2008). *Drawing and the non-verbal mind. A life-span perspective*. Cambridge, England: Cambridge University Press. doi:10.1017/CBO9780511489730
- La Voy, S. K., Pedersen, W. C., Reitz, J. M., Brauch, A. A., Luxenberg, T. M., & Nofsinger, C. C. (2001). Children's drawings: A cross-cultural analysis from Japan and the United States. *School Psychology International, 22*, 53–63. doi:10.1177/0143034301221005
- Loxton, H., Mostert, J., & Moffatt, D. (2006). Screening of intellectual maturity: Exploring South African preschoolers' scores on the Goodenough–Harris Drawing Test and teachers' assessment. *Perceptual and Motor Skills, 103*, 515–525.
- Mack, W. (2007). Improving postswitch performance in the dimensional change card-sorting task: The importance of the switch and of pretraining by redescribing the test cards. *Journal of Experimental Child Psychology, 98*, 243–251. doi:10.1016/j.jecp.2007.05.004
- Martlew, M., & Connolly, K. J. (1996). Human figure drawings by schooled and unschooled children in Papua New Guinea. *Child Development, 67*, 2743–2762. doi:10.2307/1131750
- McCarthy, D. (1944). A study of the reliability of the Goodenough Drawing Test of Intelligence. *Journal of Psychology, 18*, 201–216. doi:10.1080/00223980.1944.10544119
- Motta, R. W., Little, S. G., & Tobin, M. I. (1993a). A picture is worth less than a thousand words: Response to reviewers. *School Psychology Quarterly, 8*, 197–199. doi:10.1037/h0088274
- Motta, R. W., Little, S. G., & Tobin, M. I. (1993b). The use and abuse of human figure drawings. *School Psychology Quarterly, 8*, 162–169. doi:10.1037/h0088273
- Naglieri, J. A. (1988). *Draw A Person: A Quantitative Scoring System*. San Antonio, TX: Psychological Corporation.
- Naglieri, J. A. (1993). Human figure drawings in perspective. *School Psychology Quarterly, 8*, 170–176. doi:10.1037/h0088275
- Naglieri, J. A., & Pfeiffer, S. I. (1992). Performance of disruptive behavior disordered and normal samples on the Draw A Person: Screening Procedure for Emotional Disturbance. *Psychological Assessment, 4*, 156–159. doi:10.1037/1040-3590.4.2.156

- Pfeffer, K., & Olowu, A. A. (1986). Perceptions of dress style in some Nigerian children's drawings. *Journal of Social Psychology, 126*, 287–290. doi:10.1080/00224545.1986.9713588
- Picard, D., & Vinter, A. (2007). Relationships between procedural rigidity and interrepresentational change in children's drawing behavior. *Child Development, 78*, 522–541. doi:10.1111/j.1467-8624.2007.01013.x
- Prewett, P. N., Bardos, A. N., & Naglieri, J. A. (1989). Assessment of mentally retarded children with the Matrix Analogies Test–Short Form, Draw A Person: A Quantitative Scoring System, and the Kaufman Test of Educational Achievement. *Psychology in the Schools, 26*, 254–260. doi:10.1002/1520-6807(198907)26:3<254::AID-PITS2310260306>3.0.CO;2-C
- Richter, L. M., Griesel, R. D., & Wortley, M. E. (1989). The Draw-a-Man Test: A 50-year perspective on drawings done by Black South African children. *South African Journal of Psychology, 19*, 1–5. doi:10.1177/008124638901900101
- Riggs, K. J., Jolley, R. P., & Simpson, A. (2013). The role of inhibitory control in the development of human figure drawing in young children. *Journal of Experimental Child Psychology, 114*, 537–542. doi:10.1016/j.jecp.2012.10.003
- Silk, A. M. J., & Thomas, G. V. (1988). The development of size scaling in children's figure drawings. *British Journal of Developmental Psychology, 6*, 285–299. doi:10.1111/j.2044-835X.1988.tb01101.x
- Sisto, F. F. (2000). Relationship of the Piagetian cognitive development to human figure drawing. *Child Study Journal, 30*, 225–232.
- Taborda de Velasco, R. A. (1993). Dibujo de la Figura Humana. Estudios comparativos de adaptaciones efectuadas en San Luis, Argentina (Human figure drawing: Comparative studies of adaptations carried out in San Luis, Argentina). *Interdisciplinaria Revista de Psicología y Ciencias Afines, 12*, 1–9.
- ter Laak, J., de Goede, M., Aleva, A., & van Rijswijk, P. (2005). The Draw-A-Person Test: An indicator of children's cognitive and socio-emotional adaptation? *Journal of Genetic Psychology, 166*, 77–93. doi:10.3200/GNTP.166.1.77-93
- Terman, L. M., & Merrill, M. A. (1960). *Stanford-Binet Intelligence Scale: Manual for the third revision, Form L-M*. Oxford, England: Houghton Mifflin.
- Thomas, G. V. (1995). The role of drawing strategies and skills. In C. Lange-Küttner & G. V. Thomas (Eds.), *Drawing and looking: Theoretical approaches to pictorial representation in children* (pp. 107–122). Hemel Hempstead, England: Harvester Wheatsheaf.
- Thomas, G. V., & Jolley, R. P. (1998). Drawing conclusions: A re-examination of empirical and conceptual bases for psychological evaluation of children from their drawings. *British Journal of Clinical Psychology, 37*, 127–139. doi:10.1111/j.2044-8260.1998.tb01289.x
- Thomas, G. V., & Tsalimi, A. (1988). Effects of order of drawing head and trunk on their relative sizes in children's human figure drawings. *British Journal of Developmental Psychology, 6*, 191–203. doi:10.1111/j.2044-835X.1988.tb01093.x
- Towse, J. N., Redbond, J., Houston-Price, C. M. T., & Cook, S. (2000). Understanding the dimensional change card sort: Perspectives from task success and failure. *Cognitive Development, 15*, 347–365. doi:10.1016/S0885-2014(00)00021-6
- Tryphon, A., & Montangero, J. (1992). The development of diachronic thinking in children: Children's ideas about changes in drawing skills. *International Journal of Behavioral Development, 15*, 411–424. doi:10.1177/016502549201500308
- van Sommers, P. (1984). *Drawing and cognition: Descriptive and experimental studies of graphic production processes*. Cambridge, England: Cambridge University Press. doi:10.1017/CBO9780511897672
- Vedder, P., van de Vijfeijken, K., & Kook, H. (2000). Borrowing norms for the human figure drawing test: How to validate this practice? *Scandinavian Journal of Educational Research, 44*, 393–404. doi:10.1080/713696680
- Walker, S. (2007). *Independent school inspection report: Hall School Wimbledon*. London, England: Ofsted. Retrieved from <http://www.ofsted.gov.uk/inspection-reports/find-inspection-report/provider/ELS/101086>
- Wilgosh, L. (1991). Underachievement in culturally different gifted children. *European Journal for High Ability, 2*, 166–173. doi:10.1080/0937445910020206
- Willats, J. (2005). *Making sense of children's drawings*. Mahwah, NJ: Erlbaum.
- Willcock, E., Imuta, K., & Hayne, H. (2011). Children's human figure drawings do not measure intellectual ability. *Journal of Experimental Child Psychology, 110*, 444–452. doi:10.1016/j.jecp.2011.04.013
- Wisniewski, J. J., & Naglieri, J. A. (1989). Validity of the Draw a Person: A Quantitative Scoring System with the WISC–R. *Journal of Psychoeducational Assessment, 7*, 346–351. doi:10.1177/073428298900700408
- Witt, A., & Vinter, A. (2011). Learning implicitly to produce avoided behaviours. *Quarterly Journal of Experimental Psychology, 64*, 1173–1186. doi:10.1080/17470218.2010.543283
- Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS): A method of assessing executive function in children. *Nature Protocols, 1*, 297–301.
- Zhi, Z., Thomas, G. V., & Robinson, E. J. (1997). Constraints on representational change: Drawing a man with two heads. *British Journal of Developmental Psychology, 15*, 275–290. doi:10.1111/j.2044-835X.1997.tb00521.x

AQ: 5

AQ: 6

Received June 16, 2012

Revision received August 19, 2013

Accepted August 21, 2013 ■

AUTHOR QUERIES

AUTHOR PLEASE ANSWER ALL QUERIES

1

AQau—Please confirm the given-names and surnames are identified properly by the colors.

■ = Given-Name, ■ = Surname

The colors are for proofing purposes only. The colors will not appear online or in print.

AQ1—Author: Please limit the number of keywords to five. Delete at least one.

AQ2—Author: Wechsler (1967) citation has no corresponding reference. Please provide. Check year (WISC, 1949; WISC-R, 1974). Rather than the Wechsler Intelligence Scale for Children, do mean instead the Wechsler Preschool and Primary Scale of Intelligence: “Wechsler, D. (1967). *Manual for the Wechsler Preschool and Primary Scale of Intelligence*. San Antonio, TX: Psychological Corporation”?

AQ3—Author: Please note insertion of missing cross-reference to Figure 2. Is placement OK?

AQ4—Author: Please confirm accuracy of Fabry and Bertinetti (1990) reference, as cited in text (one time). You did not provide one.

AQ5—Author: Please confirm accuracy of Terman and Merrill (1960) reference, as cited in text (one time). You did not provide one.

AQ6—Author: Please confirm accuracy of Zelazo (2006) reference, as cited in text (one time). You did not provide one.
