TAMSAAL
ISSN Print: 3007-5939
ISSN Online: 3007-5947
Volume 4, Number 11, 2026, Pages 01 – 21
Journal Home Page
https://tamsaal.pk/index.php/tamsaal/index

Human Figure Drawing in a Pakistani Adult Clinical Sample: Base Rates, Demographic Correlates, and Presenting-Problem Differentiation

Syed Aftab Shah[1], Dr Anila Amber Malik[2]

[1]*PhD Scholar, Department of Psychology, University of Karachi Pakistan. Email: saftab123@gmail.com*
[2]*Department of Psychology, University of Karachi. Email: aamalik@uok.edu.pk*

ARTICLE INFO

Article History:

Received:	*September*	*25, 2026*
Revised:	*October*	*04, 2026*
Accepted:	*October*	*26, 2026*
Available Online:	*November*	*01, 2026*

Keywords:

Draw-a-Person; human figure drawing; projective assessment; base rates; Pakistan; cross-cultural assessment; forensic psychology

Corresponding Author:
Syed Aftab Shah
Email:
saftab123@gmail.com

DOI:
https://doi.org/10.59075/tamsaal.v4i11.251

ABSTRACT

Background. *The Draw-a-Person procedure remains among the most widely administered instruments in clinical practice worldwide, yet almost the entire empirical literature on its indicators derives from Western child samples. No published base rates exist for adult human figure drawings (HFD) in Pakistan, and interpretive cut-points in routine use were established in populations whose drawing conventions, dress, and visual culture differ substantially from those of South Asian clinical attenders.* ***Objectives.*** *To establish base rates for 52 HFD indicators in a Pakistani adult clinical sample, to test whether age, sex, and education predict drawing characteristics, and to determine whether presenting problem differentiates drawings once multiplicity is controlled.* ***Method.*** *One hundred and one consecutive adult clinical attenders (60 female, 40 male, 1 other; M age = 28.53, SD = 8.77) completed a single-figure Draw-a-Person task. Drawings were scored on a 52-indicator protocol spanning structural, completeness, proportionality, special-feature, and global-rating domains. Presenting problems were coded from clinical records into six non-exclusive categories. Analyses comprised base-rate estimation with Wilson confidence intervals, Welch t-tests with Hedges-corrected effect sizes, Fisher exact tests, and multiple regression, with Benjamini–Hochberg control of the false discovery rate across the differentiation analyses.* ***Results.*** *Twenty-one protocols (20.8%) contained no torso, arms, or legs and are described here as a head-only response class; these drawings occupied the page as fully as full figures ($d = -0.02$) but scored 8.6 points lower on completeness ($d = -2.50$). Ninety-eight per cent of figures were drawn without a baseline. Hats were absent from every protocol; pockets appeared in 9.9% and buttons in 18.8%, against hair in 91.1%. Age and education were unrelated to figure size, completeness, and global ratings. Women drew substantially smaller figures than men ($d = -0.75$, $p = .0008$), an effect that survived adjustment for anxiety, age, and figure type. Anxiety was associated with smaller figures ($d = -0.58$) and trauma with less complete ones ($d = -0.69$), but no presenting-problem effect survived false-discovery-rate correction across 25 tests (minimum $q = .076$).* ***Conclusions.*** *HFD indicators in this population depart sharply from published Western expectation. The floating figure, treated in the classical literature as pathognomonic of insecurity, is here the population norm and carries no discriminative information. A substantial minority of adults respond to the Draw-a-Person instruction by drawing a head rather than a person, which renders proportionality indices undefined and depresses completeness for reasons unrelated to psychopathology. Sex is the only demographic variable that predicts drawing characteristics. Presenting-problem differentiation is at best weak and does not survive control for multiple comparisons. These base rates provide the first normative reference for adult HFD interpretation in Pakistan and argue against the inferential use of individual indicators in clinical and, particularly, forensic contexts.*

© 2026 The Authors, Published by ULLR. This is an Open Access Article under the Creative Common Attribution (CC-BY 4.0)

Introduction

The human figure drawing is one of the most enduring of all the psychological instruments, existing on the flimsi of only a little evidence. The Draw-a-Person procedure has been used to assess personality for nearly 80 years now, since Machover, and is still one of the most used projective personality assessment techniques in clinical settings, yet is one of the most criticised in the research literature. However, the discrepancy between these two facts is the issue this paper seeks to resolve and it only increases when the instrument is applied to samples outside of the West.

The Draw-a-Person tradition and its interpretive claims

Figure drawings are used clinically as a developmental tool, and not as a projective tool. Goodenough's Draw-a-Man test, which was published in 1926, assigned a score to the presence and articulation of body parts as a measure of intellectual maturity, because it was believed that the degree of a child's development of the concept of the human body was reflected in the degree of the presence and articulation of body parts in his drawing of the man. Harris' revision in 1963 extended and re-normalized this scoring and the developmental tradition has continued with Naglieri's Draw A Person: A Quantitative Scoring System and its successors, which are still in print and in use. The projective turn was introduced by Machover, who theorized that the personality of the person who drew the figure is reflected in the figure itself, that is, that the figure is not just an aspect of cognitive maturity, but is actually a projection of the person's body image; and that the graphic characteristics of the figure reflect structure and conflict of the person's personality (Ahmad et al., 2025). The system that was developed after it was the interpretive system, based on psychoanalytical reasoning and clinical observation, not on empirical association; large eyes were thought to represent suspiciousness or paranoid fear of being observed, wide hips a sexual preoccupation, etc., and so on, through a comprehensive list of signs. Buck's House-Tree-Person, Urban's Draw-A-Person Catalogue, and Koppitz's system of 30 emotional clues to kids' drawings each extended this logic, and each provided sign-to-meaning mappings for individual protocols. Three aspects of this interpretive framework are of special significance to this study. First, it is item level: inference is drawn from the presence or absence of a specific graphic feature to a specific psychological attribution. Second, it is norm-free, in the sense that the classical texts do not indicate the frequency of the sign in unselected populations, so a clinician can't make a principled judgment about whether a noted feature is unheard-of. Third, it is culturally unmarked: the sign catalogues are culturally unmarked and do not accommodate any variation in dress conventions, drawing instruction, or visual culture, and are written as if the human figure were a universal object drawn under universal conditions.

The empirical critique

The response to Machover's system was quick and in general has not changed in the past sixty years. In the 1950s and 1960s Swensen reviewed the hypotheses in Machover's 1948–61 studies and discovered that most of the specific sign hypotheses failed to be replicated, and global judgments of drawings were more likely to replicate than the individual ones were. The findings of Roback's review of its usefulness in personality assessment were similar in 1968, as were those of Kahill's survey of the literature from 1967 to 1982. Much sharper was the review by Motta, Little and Tobin in the early 1990s where they stated that the continued use of the projective use of figure drawings in clinical practice was not defensible in the light of the

evidence available. The most thorough and up-to-date evaluation of the modern status of the projective technique is by Lilienfeld, Wood, and Garb in their monograph on the scientific status of projective techniques. The authors examined the psychometric characteristics of the Rorschach, Thematic Apperception Test, and human figure drawings, and found the evidence for the validity of the human figure drawings was less than that for either of the other two tests, and that projective indexes have, with a few exceptions, not shown incremental validity in relation to other psychometric evidence. Their accompanying meta-analysis on the detection of child sexual abuse showed that some projective instruments were more useful than chance, but that there were essentially no findings that were replicated by independent research teams, and that effect sizes from published studies were significantly greater than those from unpublished ones (a clear effect size file-drawer effect in the literature for the first time). Syeda Masooma Naqvi et al., 2025 the branch of development has not gone much better when examined. Recent research on the Draw A Person: Intellectual Ability Test (DAP-IAT) indicated low to moderate correlation in children and adults with Wechsler scores, and high false positive and false negative rates for screening for borderline intellectual functioning and superior intellectual functioning, respectively; the authors concluded that figure drawings should not be used as a projective assessment of intellectual functioning. The same was found in a separate evaluation of the Quantitative Scoring System which concluded that the system could not provide a "psychometrically sound" evaluation of children's intellectual functioning and that practitioners should not use it. Two possible reasons have been suggested as to why clinical confidence is maintained in the face of this record? Two explanations have been offered as to why clinical confidence is even maintained in the face of this record? The first is illusory correlation – Chapman and Chapman showed that observers can reliably perceive associations between drawing signs and clinical characteristics that are not actually part of the data and the perceived associations follow the semantic similarity between the sign and the symptom, not the empirical covariation. The second is the art quality confound which is discussed later.

Graphic competence as a confounding variable

One striking and persistent result is that clinicians judging the psychopathology of figure drawings are greatly affected by drawing ability of the figure. As early as 1958, Feldman and Hunt demonstrated that ratings based on figure drawings parallel the graphic difficulty of the elements being rated. The general factor found in the Draw-a-Person scores by Nichols and Strümpfer was a general factor for drawing, not a personality factor. Cressen discovered that, to a significant extent, judges' evaluations of the Draw-a-Person were based on artistic quality. The most direct test was made by Feher, Vandecreek and Teglasi who tried to guess which of 64 drawings was drawn by a psychiatric patient, half of the guessers being specifically told that the drawings were not of high art quality. The warned judges' hit rate of 58% was statistically above chance, as was the hit rate of the naive judges, but was of little practical use; they were guided by art quality. The authors found some limited validity in the figure drawing tests, and at the same time found that figure drawing tests invariably measure artistic ability in some degree and that this was very hard to control (Ahmad et al., 2022). This is important for base-rate work as most structural HFD indices are arithmetically mixed with graphic competence. Completeness scores are scores of any body parts, while proportionality indices are the ratio of body parts. Both are not defined or "null" for drawings that fail to include the corresponding parts, and both are mechanically better with drawing ability. If these sources are not separated in a base-rate table, the competence variation will be treated as clinical variation.

The contemporary revival

In the past decade, with this backdrop, figure drawing research has experienced a resurgence in activity, largely in response to the interest in the use of non-verbal screening instruments that are less vulnerable to social-desirability distortion than self-report assessments. After participants were divided into major depressive disorder and non-depressed control groups of 113 and 97, respectively, and age and gender were taken into account, three of the drawing features were linked to depression: shaded eyes, detailed clothing and drawing of extra personal objects. The receiver-operating-characteristic analysis they conducted, however, indicated little discrimination for any single feature of the drawing, and the authors claim it should be viewed as a whole, rather than in terms of individual features. It is a significant qualification and echoes Swensen's half a century earlier who had concluded that global judgement is superior to local inference. There is another more rewarding track of self-figure drawing that has grown in trauma populations, which is parallel. Lev-Wiesel's programme focused on a small number of indicators in the adult's own drawings of childhood sexual abuse (face line double, hollow or shaded; eyes drawn as dots, omissions, shading or hollow; hands and arms cut off, attached or detached; and the shading or blocking out of the genitals from the rest of the body) and, subsequently, studies have been undertaken that have extended this framework to such physical and emotional experiences as eating disorders, self-injury, and suicidality. Compared with the classical tradition, this work is methodologically more solid in that the indicators are well determined in advance and are then evaluated in relation to criteria that have been measured independently, although the number of the persons involved is usually limited (between 20 and 94). It is also noteworthy that these studies employ a slightly different version of the draw yourself task as opposed to the classical draw a person task, a substantially different task.

The cross-cultural problem

The evidence on the systematic differences in drawing conventions across cultures is now considerable, and more telling than any of the internal-validity critiques, in undermining the transportability of catalogues of drawing conventions derived from the West. Comparative studies in the ecocultural tradition have demonstrated that both figure size and the number of facial features as well as facial expression are culturally variable. In the comparison made between children from relatively collectivist settings (Turkey and Cameroon) and children from urban Western settings, the authors found that children who came from relatively collectivist backgrounds tended to be more likely to leave out facial features or to show less overt emotional expression. The size of figures has been correlated with the emphasis on independence and uniqueness in the culture, with taller figures being associated with a greater emphasis on independence and uniqueness. According to Fatima et al., 2025 a cross-continental study of children in England, Iran and Brazil revealed that the figures drawn were significantly smaller in Brazil than in England and Iran, and that the children in this study were aware of their figures being smaller, which led to concerns that emotional cues, including the size of the drawing, shading, the use of a profile versus front view, and the level of detail, might be influenced by a cultural factor and thus misinterpreted. The same was found in formal instruction of drawing in Papua New Guinea, as shown by Martlew and Connolly's comparison between schooled and unschooled children.

In the case of Muslim-majority populations, there is yet another factor that the assessment literature has neglected to take into account. There is a documented resistance to depictions of

animate beings in the Islamic visual tradition, which is rooted in the hadiths but not in the Quran itself and is often read much more liberally over time and within communities. It has traditionally been enforced most rigorously with regard to religious art and architecture; and, in almost every Islamic culture, figural representation was most robust in secular settings. The practice of aniconism in Islamic societies nowadays is mostly confined to religious contexts (Fatima et al., 2025). It is not that adults in Pakistan will not do it, it is that the number of previous examples of human figure that they have drawn would not necessarily be the same as those of a Western normative sample. This concern, though expressed informally by clinicians in Pakistan, was that those who had previously drawn portraits of human figures would have a definite impact on the instrument's output that would not be observed otherwise. This concern has not, to our knowledge, been tested against data.

The absent base rates

The result of all of the above is a particular and correctable deficiency. When dealing with a single figure drawing one needs to have an understanding of what is normal, "normal," or common, among the population from which the drawing was drawn. The lack of a ground line in a given protocol is devoid of information; this is true even if the classical literature claims they are insecure and unstable, provided that ninety-eight per cent of drawings in the population of drawings of a given clinical population lack a ground line. For one hundred people to draw hats without an adult being present, then a scoring inventory which assigns an item to the hat is measuring nothing. Base rates are not a prerequisite for interpretation, they are a precondition for interpretation.

These base rates are not available for clinical populations in Pakistan. The HFD literature is dominated by a literature on children, which reviewers repeatedly mention and which is borne out by the fact that even recent adult literature refers to the need to fill a gap. Overall, the literature on cross-cultural issues is also skewed towards children and towards comparisons between Western and non-Western groups, including a small number of comparisons between Eastern and Western Asians, African and Western groups, and Latin America and other groups; South Asian groups are underrepresented, and Pakistan is virtually absent (Malik et al., 2025). In Pakistan, where figure drawings are commonly used in clinical, educational, and forensic contexts, the norms practitioners use to interpret drawings are drawn from a different population, different age group, and different visual culture.

The present study

This paper presents base rates for 52 indicators of human figure drawing in a consecutive sample of 101 adult clinical attenders in Karachi, and addresses three sequential questions that can be solved using the base rates.

The first is descriptive: What is the clinical figure drawing of an adult Pakistani? What indicators are almost everywhere; what indicators are virtually nonexistent; and what indicators are sufficiently variable to be useful at any level? This question comes before all other interpretations and has not been previously answered for this population.

The second is related to demographic confounding: do characteristics of drawing differ as a function of age, sex, and education? If they do, they must be interpreted on the basis of them; if

they do not, the source of potential artefact is significantly reduced. The education question is particularly interesting, because of the art quality literature; the most obvious proxy for drawing practice available in routine clinical records is formal education.

The third relates to clinical differentiation: after controlling for demographic factors, are drawing characteristics predicted by presenting problem, and do such effects remain following appropriate multiple comparisons? Where the classical interpretive claims reside is where the field's replication record is poor – the differentiation question. It's addressed here with pre-specified indices, effect sizes with confidence intervals and false-discovery-rate correction, so that whatever comes out can be judged on its magnitude, not whether it has an asterisk or not.

A fourth question arose from the data, rather than being a question that had been formulated in advance, and is presented accordingly. A large proportion of respondents (36%) followed the draw-a-person task by drawing a head only. This response class was unexpected, did not have a "home" in the scoring literature, and has significant measurement implications. The description of it is provided above, and it is included as a covariate in the analyses.

Method

Design and setting

This was a cross-sectional observational study using clinically collected data. Participants were consecutive adult attenders at a private psychological practice in Karachi, Pakistan, seen between the practice's establishment and the data-freeze point for the present analysis. The Draw-a-Person task was administered as part of routine initial assessment, and all drawings produced during the accrual window were included; there was no selection on presentation, drawing quality, or subsequent outcome.

Participants

The analysed sample comprised 101 adults. Sixty (59.4%) were women, 40 (39.6%) were men, and one participant recorded a gender other than male or female. Mean age was 28.53 years (SD = 8.77, range 18–65); women were somewhat younger than men (M = 27.23 vs 30.00), though not significantly so, $t(96.8) = -1.53$, $p = .132$, $d = -0.32$. Educational attainment was high and narrowly distributed: 76 participants (75.2%) were graduates, 17 (16.8%) had completed intermediate education, and 8 (7.9%) held postgraduate qualifications. The restricted range on education is a characteristic of the private urban practice setting and constrains the education analyses reported below.

Presenting problems were abstracted from clinical records into six non-exclusive categories, so that comorbidity was recorded rather than forced into a single classification. Anxiety was present in 83 participants (82.2%), mood or depressive difficulties in 55 (54.5%), relational difficulties in 44 (43.6%), trauma or post-traumatic presentations in 32 (31.7%), and behavioural difficulties in 9 (8.9%); a residual 'other' category was recorded for 82 (81.2%). Participants carried a mean of 3.02 problem categories (SD = 1.09, range 1–6), and only five presented with a single category. Mood or depressive difficulties were more frequently recorded in women than men (65.0% vs 37.5%, Fisher $p = .008$); no other presenting-problem category differed by sex.

Ethical considerations

The drawings and clinical records were obtained during the normal clinical practice. The data were anonymized at source, participants were coded sequentially only, no contact details, identifying clinical narrative were kept in the analysis data set. Since the material has been gathered for clinical purposes, not for research, the consent to use the material for secondary research was covered by means of the practice's retrospective consent procedure, and no drawing is reproduced in this report.

The drawing task

Participants were given a blank A4 sheet in portrait orientation and a pencil with an eraser, and asked to draw a person. No further specification was given: the instruction did not name a sex, an age, an activity, or a level of completeness, and did not indicate that the figure should represent the participant. No time limit was imposed. This is the classical single-figure Draw-a-Person administration, and it is distinct from the draw-yourself instruction used in the self-figure trauma literature.

Scoring protocol

Drawings were scored according to a specially designed 52 indicator protocol divided into 6 sections. Demographic and clinical context was collected in Section A. Structural characteristics recorded in Section B were: figure height, figure width, figure height as a percentage of page height, categorical size, horizontal and vertical placement, line pressure, line continuity, and symmetry. The completeness of the items scored in Section C represented facial features (head, face outline, eyes, nose, mouth, ears, hair, eyebrows or lashes, and chin), trunk (neck, shoulders, torso, and waist), upper limbs (arms, hands, and fingers), lower limbs (legs, feet, and toes), and clothing (presence, buttons or zipper, pockets, hat, and shoes). Proportionality was measured as head height, arm length, leg height, shoulder width, and hip width (cm) in section D with ratios of head-to-body, arm-to-body, leg-to-body, shoulder-to-hip, and hip-to-body calculated and categorical classifications provided at pre-specified cut-points. The special features noted in Section E were shading, erasure, line reinforcement, transparency, ground line, floating figure, profile/back orientation, figure sex relative to the drawer, and facial expression. Four-point scales were used to rate global levels of integration and detail level in Section F.

All completeness items were coded 1 (clearly present), 0 (absent), or P (partially drawn or ambiguous) and the partial code was assigned with great caution when the coder was not sure whether the item was present or not. Partial credit was recorded separately and items scored a "1" only counted for a composite score. Gardez et al., 2024 derived proportions were fixed in advance and stored in an editable rules table, to ensure the same proportion being derived from the same measurements on every specimen. Cut-points were set so that figures smaller than 25% of the page height were classified small, and those greater than 75% were classified large; head-to-body ratios less than 0.125 were classified disproportionately small, and those more than 0.25 were classified disproportionately large; arm-to-body ratios less than 0.30 were classified short, and those greater than 0.50 were classified elongated; leg-to-body ratios less than 0.35 were classified short, those greater than 0.55 were classified long, and the intervening area was roughly equal; shoulder-to-hip ratios less than 0.90 were classified hips-wider, and those more than 1.10 were classified shoulders-wider, with the area in between treated as roughly equal.

Structural absence: clearly noted contrast to missing data. For any missing body part, the measurement was omitted and the derived category returned a classification that was explicitly not-drawn instead of a blank, therefore the two categories could not be combined for analysis. This separation was significant and is reported in the Results.

Statistical analysis

Base rates for indicators of binary variable are reported as percentages and presented with a Wilson 95% confidence interval, which has the expected properties for indicators with extreme base rates, such as several of those mentioned above. All continuous indices are summarized as means, standard deviations, medians, and ranges.

For continuous measures, group comparisons were conducted using the Welch t-test, which is not based on the assumption of equal variances and suitable for the very uneven cell sizes created by the presenting-problem categories. The reported effect sizes are expressed as Cohen's d and include 95% confidence intervals; every t-test is also accompanied by a distribution-free test (Mann–Whitney U test), which are reported when the two tests disagree and are noted in the text. Fisher's exact test was used to evaluate associations of categorical variables. Spearman's correlations and Pearson's correlations (with Fisher transformed confidence intervals) are reported for associations between continuous variables.

The 25 tests (five problem categories × five indices) in the presenting-problem analyses are included. Given its history of unreplicable conclusions in figure-drawing literature, the false discovery rate for these 25 tests was controlled by the Benjamini–Hochberg procedure, and the FDR-corrected (FDR) q-values are presented alongside uncorrected p-values. A parallel family of 70 item-level Fishers were corrected in the same fashion. Uncorrected results are also provided in order to allow readers to make independent judgments about magnitude as compared to the correction; however, no claim for differentiation is made with the use of an uncorrected p-value alone.

The principal univariate effects were tested for mutual adjustment by multiple regression. Sensitivity analyses are conducted in cases where the head-only response class could reasonably explain an effect, and are reported for retesting the major comparisons within the full-figure subsample. The analysis was done in Python 3.12 in the packages pandas, NumPy, and SciPy.

Results

A head-only response class

Twenty-one of the 101 protocols (20.8%, 95% CI [14.0, 29.8]) contained no torso, no arms, and no legs. Coder notes for these cases were explicit and consistent, recording variants of a floating head, a large face drawing, or in one case a full face occupying the entire page. These are not incomplete or abandoned drawings: they are a distinct response to the instruction, in which the participant drew a head rather than a person.

This class is demographically unremarkable. It did not differ by sex (12 of 60 women, 9 of 40 men; Fisher $p = .805$), by age ($M = 28.71$ vs 28.49, $d = +0.03$, $p = .924$), or by education (χ^2 $p = .898$). It was not associated with mood, anxiety, relational, or trauma presentations (all Fisher $p > .11$). Most strikingly, head-only drawings occupied the page as fully as full figures: mean page

height utilization was 52.67% against 53.07%, d = −0.02, p = .953. These participants used the paper; they simply drew a different object.

The measurement consequences are severe. Head-only protocols scored 7.62 on completeness against 16.26 for full figures, a difference of 8.6 points and d = −2.50, 95% CI [−3.09, −1.91]. Head-to-body ratio averaged 0.609 against 0.227, which is to say that the ratio in these cases is not measuring head proportion at all but the arithmetic consequence of a head that constitutes most of the figure. Arm, leg, and shoulder-to-hip ratios are undefined. Global ratings were also lower (integration d = −0.64, p = .006; detail d = −0.72, p = .0002), though facial detail was equivalent (C1 Face Score 6.48 vs 6.85, d = −0.16, p = .631) — these participants drew faces as fully as anyone else.

Because completeness and the proportionality indices behave differently in this group for reasons that are structural rather than clinical, figure type is carried as a covariate through the analyses that follow, and sensitivity analyses within the full-figure subsample are reported wherever it could account for an effect.

Base rates

Table 1 reports base rates for all 34 binary indicators with Wilson confidence intervals and partial-credit counts. Several features of the distribution merit comment.

The most striking single finding is that 98 of 100 code able protocols (98.0%, 95% CI [93.0, 99.4]) depicted a figure without a ground line. In the classical interpretive literature the floating figure is read as a marker of insecurity, instability, or poor reality contact. In this population it is the norm, and its absence two cases in a hundred is the rare event. No interpretive weight can be placed on it. The same logic applies at the other extreme to back-view drawings (0%), ground lines (2.0%), profile orientation (3.0%), toes (5.0%), and transparency (6.0%): each is too rare to support inference from a single protocol, and their absence is uninformative.

The clothing items are the clearest instance of an inventory mismatched to its population. Not one participant in 101 drew a hat. Pockets appeared in 9.9% of protocols and buttons or zippers in 18.8%, against shoes at 54.5% and clothing of any kind at 62.4%. Hair, by contrast, was drawn in 91.1% of protocols a rate exceeded only by the head itself (97.0%) and the face outline (93.1%). An item pool that allocates three of its five clothing items to buttons, pockets, and hats, and none to hair, is scoring garments this sample does not draw while ignoring the feature it draws most consistently.

Facial features showed high but graded presence: head 97.0%, face outline 93.1%, hair 91.1%, nose 75.2%, eyes 72.3%, eyebrows or lashes 71.3%, chin 68.3%, mouth 62.4%, ears 46.5%. The partial-credit counts on the central facial features are substantial and warrant note: mouth was coded partial in 35 protocols, eyes in 25, and nose in 21. Roughly a third of protocols were therefore ambiguous on the three features that most influence facial scoring, which is a limitation on the precision of any facial index in this sample.

Body-part completeness declined predictably down the figure: neck 85.0%, shoulders 78.0%, arms 77.2%, torso 72.0%, waist 67.0%, legs 66.3%, feet 62.4%, hands 58.4%, fingers 55.4%, toes 5.0%. Among the special features, line reinforcement (67.3%) and shading (59.4%) were

common enough to be uninformative in isolation, whereas erasure (13.9%) and extra objects (11.9%) were relatively uncommon.

Categorical distributions were similarly concentrated. Three-quarters of figures fell in the medium size band (76.2%), 88.1% were placed centrally on the horizontal axis and 78.2% in the middle vertically, 91.1% were drawn with moderate line pressure, and 87.1% were rated symmetrical. Eighty-one per cent of participants drew a figure of the same sex as themselves. Facial expression was rated positive in 50.5% and neutral in 42.6%, with only four negative ratings and three protocols in which no face was drawn. Continuous indices are summarized in Table 2.

Demographic correlates

Age was unrelated to every principal index. Correlations with page height utilization ($r = +.16$, 95% CI [−.04, .34], $p = .111$), completeness ($r = −.16$, $p = .099$), integration ($r = −.10$, $p = .309$), and detail level ($r = −.08$, $p = .419$) were all small and non-significant, and every corresponding Spearman coefficient was smaller still (all $|\rho| \leq .09$, all $p > .36$). A weak association with facial detail ($r = −.20$, $p = .044$) did not survive as a rank correlation ($\rho = −.09$, $p = .365$) and is not interpreted.

Education was likewise unrelated to drawing characteristics. Spearman correlations across the ordinal education variable were negligible for page height utilisation ($\rho = −.01$, $p = .949$), completeness ($\rho = +.02$, $p = .867$), facial detail ($\rho = −.01$, $p = .906$), integration ($\rho = −.12$, $p = .231$), and detail level ($\rho = −.16$, $p = .123$). Kruskal–Wallis comparisons across intermediate, graduate, and postgraduate groups were non-significant for four of the five indices; the exception, integration, reached $H(2) = 6.05$, $p = .049$ with $\varepsilon^2 = .041$, driven by the eight postgraduate participants scoring lowest of the three groups — a direction opposite to any drawing-practice account and resting on a very small cell. Given the restricted educational range of this sample, the appropriate conclusion is that these data provide no evidence of an education effect rather than that none exists.

Sex, by contrast, produced a substantial and specific effect. Women drew markedly smaller figures than men: mean page height utilisation 46.85% (SD = 18.47) against 62.21% (SD = 23.26), $t(70.4) = −3.50$, $p = .0008$, $d = −0.75$, 95% CI [−1.16, −0.34], with the Mann–Whitney test agreeing ($p = .0006$). The effect was confined to size. Completeness ($d = +0.07$, $p = .738$), facial detail ($d = +0.27$, $p = .201$), integration ($d = +0.16$, $p = .404$), and detail level ($d = +0.23$, $p = .232$) showed no sex difference, and none of nine binary indicators tested — shading, erasure, reinforcement, floating, clothing, hair, hands, feet, shoes — differed by sex (all Fisher $p > .29$). Women in this sample did not draw less; they drew smaller.

Presenting-problem differentiation

Table 3 reports all 25 presenting-problem comparisons with effect sizes, confidence intervals, and Benjamini–Hochberg adjusted q-values. The headline result is straightforward: no effect survived correction. The two largest, both involving trauma, reached $q = .076$; under Bonferroni correction ($\alpha = .002$) no test was significant.

The uncorrected pattern is nonetheless worth describing, on the understanding that it is hypothesis-generating rather than confirmatory. Trauma or post-traumatic presentation was

associated with less complete figures (12.25 vs 15.49, d = −0.69, 95% CI [−1.12, −0.26], p = .005) and with reduced facial detail (5.62 vs 7.30, d = −0.75, 95% CI [−1.18, −0.32], p = .006), together with weaker integration (d = −0.42, p = .042). Anxiety was associated with smaller figures (50.79% vs 63.13%, d = −0.58, 95% CI [−1.10, −0.07], p = .041; Mann–Whitney p = .015), though the comparison rests on only 18 participants without anxiety. Behavioural presentation showed a large but imprecisely estimated completeness effect (d = −0.71, 95% CI [−1.40, −0.02], p = .094) based on nine cases. Mood or depressive presentation — the category with the strongest classical expectations attached to it — differentiated nothing: all five effect sizes fell between d = −0.24 and d = +0.12, with confidence intervals spanning zero. Relational presentation was likewise undifferentiating.

At item level, 70 Fisher tests across five presenting-problem categories and 17 indicators produced no result surviving false-discovery-rate correction (minimum q = .627). The largest uncorrected differences again involved trauma, whose protocols less often included eyes (56.2% vs 79.7%, p = .018), nose (59.4% vs 82.6%, p = .024), and hair (81.2% vs 95.7%, p = .027).

Multivariable models

Two regression models were fitted to test whether the principal univariate effects survived mutual adjustment. Predicting completeness from trauma, figure type, sex, and age produced a well-fitting model (R^2 = .565, adjusted R^2 = .547, n = 101) in which figure type was overwhelmingly dominant (b = −8.26, SE = 0.83, p < .001) while trauma retained an independent though much smaller contribution (b = −1.83, SE = 0.73, p = .014). Age contributed marginally (b = −0.078, p = .048) and sex not at all (b = +0.07, p = .918). Slightly more than half the variance in completeness in this sample is attributable to whether the participant drew a person or a head.

Predicting page height utilisation from sex, anxiety, age, and figure type produced a weaker model (R^2 = .165, n = 101) in which sex was the strongest predictor (b = −14.41, SE = 4.16, p = .0008) and anxiety retained an independent contribution (b = −11.09, SE = 5.42, p = .043). Neither age (p = .580) nor figure type (p = .736) contributed, confirming that head-only drawings are not smaller drawings.

Sensitivity analyses

Restricting analysis to the 80 full-figure protocols left the sex effect on size essentially unchanged (47.69% vs 61.43%, d = −0.75, 95% CI [−1.21, −0.28], p = .003). The anxiety effect on size was similarly stable (d = −0.53, p = .014), as was the trauma effect on completeness (d = −0.65, 95% CI [−1.16, −0.15], p = .039). The floating-figure base rate within full figures was 97.5%, indistinguishable from the whole-sample estimate. The principal findings are therefore not artefacts of the head-only class, even though that class dominates the completeness variance.

Discussion

This study provides the first base-rate reference for adult human figure drawing indicators in a Pakistani clinical population. Three findings organise the discussion: the population norms depart sharply from Western interpretive expectation; a substantial minority of adults answer the

draw-a-person instruction with a head rather than a person; and clinical differentiation, where it appears at all, is weak and does not survive control for multiplicity.

Base rates overturn classical sign interpretation in this population

The floating figure is the clearest case. Classical interpretation treats the absence of a ground line as evidence of insecurity, instability, or impaired reality contact. In this sample 98% of figures float. A sign present in 98 of 100 clinical protocols cannot discriminate anything within that population; it is a description of how people in this setting draw. The two participants who did draw a ground line are, statistically, the anomalies — though with n = 2 nothing can be said about them either.

This is not an argument that the classical interpretation was wrong in the population where it was formulated. It is an argument that base rates are population-specific and that transporting a sign catalogue without them produces systematic error in a predictable direction: every protocol looks pathological, because the local norm has been coded as a pathological sign. The same reasoning applies with equal force to back views, profile orientation, ground lines, transparency, and toes, each too rare here to license inference, and to shading and line reinforcement, each too common.

The clothing findings extend the point from interpretation to instrument design. The complete absence of hats across 101 protocols, alongside pockets at 9.9% and buttons at 18.8%, indicates that the clothing sub-inventory imported from Western scoring systems is measuring a garment repertoire this sample does not draw. Meanwhile hair — present in 91.1% of protocols, and by a wide margin the most consistently rendered non-structural feature — receives no dedicated place in the classical clothing or detail inventories. Where an item pool has this little variance, it contributes noise and coder burden without discriminative return. This observation is developed further in a companion paper on the cultural situatedness of the item pool.

The head-only response class

One participant in five drew a head. This is a substantial and, so far as we are aware, undescribed phenomenon in the adult figure drawing literature, and its properties are informative. It is not a demographic artefact: it is unrelated to sex, age, and education. It is not a clinical marker: it is unrelated to every presenting-problem category tested. It is not an index of disengagement or task refusal: these participants used as much of the page as everyone else and drew faces of equivalent detail. It appears instead to be a legitimate alternative reading of an ambiguous instruction. Asked to draw a person, roughly a fifth of adults in this sample drew the part of a person that identifies them.

The measurement implications are more important than the phenomenon itself. Completeness scores in a mixed sample of head-only and full-figure protocols are dominated by which kind of drawing was produced — figure type alone accounted for the bulk of the explained variance in the completeness model — and any correlation computed across the pooled sample is a mixture statistic rather than a clinical one. Proportionality indices are worse: head-to-body ratio in a head-only protocol approaches unity by construction, and arm, leg, and shoulder-to-hip ratios do not exist. A study that treats these as missing at random, or that pools the two response types, will generate associations that are arithmetic rather than psychological. This is a specific and checkable mechanism by which figure drawing research can produce findings that fail to

replicate, and it is one that the existing literature does not control for because it has not been described.

The practical recommendation is simple: figure type should be recorded as a classification variable in every adult HFD protocol, structural absence should be coded distinctly from missing measurement, and proportionality indices should be computed and reported only within full-figure protocols.

Demographic invariance and the sex effect

Age and education did not predict drawing characteristics in this sample. The education result deserves emphasis given the art-quality literature, which would predict that formal education, as a proxy for drawing practice, should track graphic competence and hence most structural indices. It did not — though the restricted educational range here means the finding is better stated as an absence of evidence than as evidence of absence, and it cannot be extended to samples with substantial primary-educated or unschooled representation, where Martlew and Connolly's schooling effects would be expected to appear.

Sex produced the only reliable demographic effect: women drew figures roughly a quarter smaller than men, $d = -0.75$, an effect that survived adjustment for anxiety, age, and figure type, and that was unchanged in the full-figure subsample. Its specificity is notable. Women did not draw less complete figures, less detailed faces, or more poorly integrated ones; they drew smaller ones. This rules out a general graphic-competence explanation, under which a sex difference in drawing skill should have expressed itself across completeness and global ratings as well.

Two readings are available and the present data cannot adjudicate between them. The cross-cultural literature associates larger figures with cultural emphasis on autonomy and self-assertion, and a sex difference in figure size within a single cultural setting is consistent with differential internalization of those norms. Alternatively, figure size may index something closer to spatial assertiveness in the immediate task situation. What can be said is that the effect is large enough to matter for interpretation: an interpretive scheme that reads small figures as evidence of withdrawal or low self-esteem, applied without sex-specific norms, will systematically over-pathologise women in this population. Sex-stratified norms are therefore a minimum requirement for any clinical use of figure size here.

Clinical differentiation is weak

The differentiation results are, on their face, negative, and this is the finding most likely to be misread. No presenting-problem effect survived false-discovery-rate correction across 25 tests, and none came close under Bonferroni. Reporting these analyses without correction would have produced four significant results and a considerably more attractive paper; it would also have reproduced exactly the methodological pattern that Lilienfeld and colleagues identified as the source of the field's file-drawer problem.

That said, a negative result under multiplicity correction with $n = 101$ is not the same as an absence of effect. The two trauma findings — reduced completeness ($d = -0.69$) and reduced facial detail ($d = -0.75$) — are of moderate magnitude with confidence intervals excluding zero, they survive adjustment in the regression model, they persist in the full-figure subsample, and they cohere with the item-level pattern in which trauma protocols less often included eyes, nose,

and hair. They also converge with the self-figure drawing literature, in which omission and disruption of facial features is among the more consistently replicated indicators in trauma populations, albeit under a different instruction. The honest reading is that trauma is the most promising differentiation signal in these data and requires confirmatory testing in an adequately powered sample with a pre-registered hypothesis, not that it has been established here.

The anxiety effect on figure size is more fragile. It is of similar magnitude ($d = -0.58$) and retains an independent contribution in the regression model, but it rests on a comparison in which only 18 participants lacked an anxiety presentation — a consequence of anxiety being present in 82% of this clinical sample. Its confidence interval extends nearly to zero. It should be regarded as suggestive at best.

The mood and depression results deserve separate comment because the classical literature attaches such strong expectations to them, and because the most directly comparable contemporary study found detectable depression-related features. Here, mood presentation differentiated nothing: five effect sizes between $d = -0.24$ and $d = +0.12$, all confidence intervals spanning zero. This is a genuine non-replication, and the most likely explanation is design rather than population. Deng and colleagues compared diagnosed patients against healthy controls; the present study compares clinical attenders with a mood presentation against clinical attenders without one. Within-clinical contrasts are inherently less powerful than clinical-versus-community contrasts, and the absence of a non-clinical comparison group is the principal design limitation of this study.

Implications for practice

For routine clinical use, the results support a narrow and cautious role for the figure drawing. As an engagement task, a conversation opener, or a source of hypotheses to be tested against other information, it retains the practical advantages that have always recommended it. As a source of inference about a specific individual's psychological state on the basis of specific graphic signs, it is not supported by these data in this population — not because the effects are absent, but because they are small relative to the variation introduced by response type, sex, and graphic competence, and because the base rates show that many classical signs carry no information here at all.

The forensic implication is sharper and should be stated plainly. Figure drawings are used in Pakistani legal and quasi-legal contexts, including child custody, competence, and abuse assessments. The present data indicate that an interpretive scheme applied without local base rates will identify the population norm as pathology in a high proportion of cases; that a fifth of adult protocols cannot be scored on proportionality indices at all; and that the strongest clinical association observed here does not survive elementary correction for multiple comparisons. None of this supports the use of figure drawing indicators as evidence about an individual in a forensic setting, and practitioners asked to do so should be prepared to explain why the instrument does not bear that weight.

Limitations

The most serious limitation is the absence of inter-rater reliability data. All protocols were scored by a single rater, and while the coding manual specifies operational definitions and fixed cut-

points for every derived variable, the reliability of the indicator-level judgements is unestablished. This constrains the base-rate estimates in particular, since a systematic scoring tendency would propagate directly into them. Double-scoring of a randomly selected subsample with kappa reported per indicator is the necessary next step and is planned.

The sample is a clinical convenience sample from a single private urban practice, with a narrow educational range and a marked over-representation of anxiety presentations. It cannot be taken as representative of Pakistani adults generally, of Pakistani clinical populations generally, or of populations with substantial rural, low-literacy, or unschooled representation, where the drawing-practice considerations discussed in the introduction would be expected to operate much more strongly.

There is no non-clinical comparison group. The base rates reported here are clinical base rates, and the question of how they differ from community norms in the same population cannot be answered from these data. This limits the differentiation analyses to within-clinical contrasts, which are the least powerful design available.

Presenting problems were abstracted from clinical records rather than established by structured diagnostic interview, and were recorded as non-exclusive categories with a mean of three per participant. The trauma category in particular aggregates heterogeneous presentations. Diagnostic imprecision attenuates effect sizes and is a plausible partial explanation for the weak differentiation results.

A small number of measurement anomalies remain in the dataset and are reported for transparency. Two protocols recorded figure heights exceeding A4 page height, yielding page utilisation values above 100%; these are most plausibly explained by landscape orientation or by measurement of a figure extending beyond the sheet, and were retained rather than excluded. One protocol carried a head height exceeding its recorded figure height and has been flagged for re-measurement. Two protocols recorded a shoulder width of zero. The affected cases are few and their exclusion does not alter any reported result, but the raw measurements have not been retrospectively adjusted.

Finally, the head-only class was identified in the course of analysis rather than anticipated, and its treatment here — as a covariate and as the basis for sensitivity analysis — is therefore post hoc. The class is defined by an objective and reproducible rule and its properties are described rather than tested against a hypothesis, but confirmation in an independent sample is required before it can be regarded as an established feature of adult HFD response.

Conclusion

Human figure drawings in this Pakistani adult clinical sample look substantially unlike the figure the classical interpretive literature describes. Almost every figure floats, no figure wears a hat, nearly all have hair, and one in five is a head rather than a person. Age and education predict nothing; sex predicts size and only size. Presenting problem predicts little, and nothing that survives correction for multiple comparisons, though trauma emerges as the signal most worth testing in a properly powered confirmatory study. These base rates are offered as the minimum reference material required before any interpretive claim about an individual drawing can be

made in this population, and as an argument that such claims should currently be made sparingly, if at all.

Tables

Table 1. ***Base rates of 34 binary human figure drawing indicators (N = 101)***

Indicator	n	Present	%	95% CI	Partial
Facial features					
C1 Head	101	98	97.0	[91.6, 99.0]	3
C1 Face	101	94	93.1	[86.4, 96.6]	6
C1 Eyes	101	73	72.3	[62.9, 80.1]	25
C1 Nose	101	76	75.2	[66.0, 82.6]	21
C1 Mouth	101	63	62.4	[52.6, 71.2]	35
C1 Ears	101	47	46.5	[37.1, 56.2]	11
C1 Hair	101	92	91.1	[83.9, 95.2]	6
C1 Eyebrows/Lashes	101	72	71.3	[61.8, 79.2]	4
C1 Chin	101	69	68.3	[58.7, 76.6]	13
Trunk					
C2 Neck	100	85	85.0	[76.7, 90.7]	2
C2 Shoulders	100	78	78.0	[68.9, 85.0]	9
C2 Torso	100	72	72.0	[62.5, 79.9]	9
C2 Waist	100	67	67.0	[57.3, 75.4]	5
Upper limbs					
C3 Arms	101	78	77.2	[68.1, 84.3]	3
C3 Hands	101	59	58.4	[48.7, 67.5]	4
C3 Fingers	101	56	55.4	[45.7, 64.8]	3
Lower limbs					
C4 Legs	101	67	66.3	[56.7, 74.8]	6
C4 Feet	101	63	62.4	[52.6, 71.2]	1
C4 Toes	100	5	5.0	[2.2, 11.2]	0
Clothing					
C5 Clothing Present	101	63	62.4	[52.6, 71.2]	5
C5 Buttons/Zipper	101	19	18.8	[12.4, 27.5]	0
C5 Pockets	101	10	9.9	[5.5, 17.3]	0
C5 Hat	101	0	0.0	[0.0, 3.7]	0
C5 Shoes	101	55	54.5	[44.8, 63.8]	1
Special features					
E Shading	101	60	59.4	[49.7, 68.5]	0
E Erasures	101	14	13.9	[8.4, 21.9]	0
E Reinforced Lines	101	68	67.3	[57.7, 75.7]	0
E Transparency	100	6	6.0	[2.8, 12.5]	0
E Baseline	100	2	2.0	[0.6, 7.0]	0
E Figure Floating	100	98	98.0	[93.0, 99.4]	0
E Profile View	100	3	3.0	[1.0, 8.5]	0
E Back View	100	0	0.0	[0.0, 3.7]	0
E Extra Objects	101	12	11.9	[6.9, 19.6]	0

Note. Percentages are of protocols with a codeable score for that indicator. Confidence intervals are Wilson intervals. "Partial" counts protocols scored P (partially drawn or ambiguous); partial scores never contribute to composite totals.

Table 2. Descriptive statistics for continuous drawing indices

Index	N	M	SD	Mdn	Range
Figure height (cm)	101	15.74	6.43	15.00	4.90–37.00
Figure width (cm)	101	8.99	4.81	7.70	2.00–21.00
Page height utilisation (%)	101	52.99	21.66	50.51	16.50–124.58
Completeness total (of 24)	101	14.47	4.93	16.00	0–22
Completeness partial count	101	1.70	2.49	1.00	0–16
C1 face score (of 9)	101	6.77	2.37	8.00	0–9
Special features total (of 9)	100	2.60	1.00	3.00	0–5
Head/body ratio	101	0.339	0.345	0.247	0.105–3.151
Arm/body ratio	82	0.276	0.082	0.267	0.146–0.667
Leg/body ratio	68	0.343	0.099	0.337	0.177–0.622
Shoulder/hip ratio	76	1.170	0.396	1.086	0.000–2.361
Integration rating (1–4)	101	2.71	0.86	3.00	1–4
Detail level rating (1–4)	101	2.71	0.78	3.00	1–4

Note. Reduced n on the arm, leg, and shoulder/hip ratios reflects structural absence — the relevant body part was not drawn — rather than missing measurement. Head/body ratios above 1.0 arise in head-only protocols and in one case flagged for re-measurement.

Table 3. Presenting-problem differentiation across five drawing indices, with false-discovery-rate correction

Index	Present M (SD)	Absent M (SD)	t	p	d [95% CI]	q
Mood / depression						
Page height %	54.22 (19.25)	51.52 (24.36)	0.61	0.543	+0.12 [-0.27, +0.52]	0.766
Completeness	14.20 (4.91)	14.78 (4.98)	-0.59	0.557	-0.12 [-0.51, +0.27]	0.766
Face score	6.67 (2.37)	6.89 (2.39)	-0.46	0.647	-0.09 [-0.48, +0.30]	0.809
Integration	2.62 (0.91)	2.83 (0.80)	-1.22	0.225	-0.24 [-0.63, +0.15]	0.624
Detail level	2.65 (0.82)	2.78 (0.73)	-0.83	0.408	-0.16 [-0.56, +0.23]	0.766
Anxiety						
Page height %	50.79 (21.04)	63.13 (22.17)	-2.16	0.041	-0.58 [-1.10, -0.07]	0.259
Completeness	14.39 (4.91)	14.83 (5.16)	-0.34	0.739	-0.09 [-0.60, +0.42]	0.840
Face score	6.67 (2.35)	7.22 (2.46)	-0.86	0.397	-0.23 [-0.74, +0.28]	0.766
Integration	2.69 (0.84)	2.83 (0.99)	-0.59	0.563	-0.17 [-0.68, +0.34]	0.766
Detail level	2.70 (0.78)	2.78 (0.81)	-0.38	0.708	-0.10 [-0.61, +0.41]	0.840
Trauma / PTSD						
Page height %	52.21 (20.35)	53.35 (22.37)	-0.25	0.800	-0.05 [-0.47, +0.37]	0.870
Completeness	12.25 (5.53)	15.49 (4.29)	-2.93	0.005	-0.69 [-1.12, -0.26]	0.076
Face score	5.63 (3.06)	7.30 (1.76)	-2.89	0.006	-0.75 [-1.18, -0.32]	0.076
Integration	2.47 (0.76)	2.83 (0.89)	-2.08	0.041	-0.42 [-0.84, +0.00]	0.259

Detail level	2.53 (0.62)	2.80 (0.83)	-1.79	0.078	-0.34 [-0.77, +0.08]	0.388
Relational						
Page height %	51.64 (21.11)	54.03 (22.20)	-0.55	0.582	-0.11 [-0.50, +0.28]	0.766
Completeness	14.43 (4.36)	14.49 (5.36)	-0.06	0.951	-0.01 [-0.41, +0.38]	0.951
Face score	6.93 (1.97)	6.65 (2.65)	0.62	0.540	+0.12 [-0.27, +0.51]	0.766
Integration	2.61 (0.84)	2.79 (0.88)	-1.02	0.310	-0.20 [-0.60, +0.19]	0.742
Detail level	2.59 (0.76)	2.81 (0.79)	-1.40	0.166	-0.28 [-0.67, +0.12]	0.518
Behavioural						
Page height %	51.55 (22.71)	53.13 (21.67)	-0.20	0.846	-0.07 [-0.76, +0.61]	0.881
Completeness	11.33 (5.32)	14.77 (4.81)	-1.87	0.094	-0.71 [-1.40, -0.02]	0.390
Face score	5.78 (3.07)	6.87 (2.29)	-1.04	0.327	-0.46 [-1.15, +0.23]	0.742
Integration	2.56 (0.73)	2.73 (0.88)	-0.67	0.519	-0.20 [-0.88, +0.49]	0.766
Detail level	2.44 (0.53)	2.74 (0.80)	-1.52	0.155	-0.38 [-1.07, +0.31]	0.518

Note. Welch t-tests. q = Benjamini–Hochberg adjusted p-value across all 25 tests. No comparison reached q < .05. Cell sizes: mood 55/46, anxiety 83/18, trauma 32/69, relational 44/57, behavioural 9/92.

Table 4. Head-only versus full-figure protocols

Variable	Head-only (n = 21)	Full figure (n = 80)	d	95% CI	p
Page height utilisation (%)	52.67 (29.47)	53.07 (19.33)	−0.02	[−0.50, +0.46]	.953
Completeness total	7.62 (3.50)	16.26 (3.45)	−2.50	[−3.09, −1.91]	< .001
C1 face score	6.48 (3.36)	6.85 (2.06)	−0.16	[−0.64, +0.32]	.631
Integration rating	2.29 (0.72)	2.83 (0.87)	−0.64	[−1.13, −0.15]	.006
Detail level rating	2.29 (0.46)	2.83 (0.81)	−0.72	[−1.21, −0.23]	< .001
Head/body ratio	0.609	0.227	—	—	—
Age (years)	28.71 (9.81)	28.49 (8.54)	+0.03	[−0.45, +0.51]	.924

Note. Head-only protocols are those in which no torso, arms, or legs were drawn. The two groups did not differ by sex (Fisher p = .805), education (χ^2 p = .898), or any presenting-problem category (all Fisher p > .11).

Table 5. Multiple regression models for completeness and figure size

Predictor	B	SE	t	P
Model 1: Completeness total (R^2 = .565, n = 101)				
Head-only figure type	−8.26	0.83	−9.98	< .001
Trauma / PTSD	−1.83	0.73	−2.51	.014
Age	−0.08	0.04	−2.00	.048
Female	+0.07	0.69	+0.10	.918
Model 2: Page height utilisation (R^2 = .165, n = 101)				
Female	−14.41	4.16	−3.46	< .001
Anxiety	−11.09	5.42	−2.05	.043
Age	+0.13	0.24	+0.56	.580
Head-only figure type	−1.68	4.97	−0.34	.736

Note. Ordinary least squares. Predictors entered simultaneously.

References

1. Amir, G., & Lev-Wiesel, R. (2007). Dissociative identity disorder as reflected in drawings of sexual abuse survivors. The Arts in Psychotherapy, 34(2), 114-123.
2. Ahmad, A., Malik, I. A., Samar Bakhtawar, Ghafoor, A. A., & Mirza, J. (2025). Exploring the Therapeutic Potential of Cognitive Behavioral Therapy in Delusional Disorder: A Single- Case Study Investigation. *The Critical Review of Social Sciences Studies*, *3*(1), 77–84. https://doi.org/10.59075/q3hhwg72
3. Ahmad, A., Malik, I. A., Bakhtawar, S., &Butt, Z. S. (2022). Evaluating the Efficacy of Dialectical Behavior Therapy (DBT) in Managing Major depressive (MDD) with Comorbid Self-harming Behaviors: A Case Study. The Critical Review of Social Sciences Studies, 2(2), 1330–1336.
4. Ballus, E., Comelles, M. C., Pasto, M. T., & Benedico, P. (2023). Children's drawings as a projective tool to explore and prevent experiences of mistreatment and/or sexual abuse. Frontiers in Psychology, 14, 1002864.
5. Baluch, B., Duffy, L. J., Badami, R., & Pereira, E. C. A. (2017). A cross-continental study on children's drawings of football players: Implications for understanding key issues and controversies in human figure drawings. Europe's Journal of Psychology, 13(3), 455-471.
6. Buck, J. N. (1948). The H-T-P technique: A qualitative and quantitative scoring manual. Journal of Clinical Psychology, 4(4), 317-396.
7. Chapman, L. J., & Chapman, J. P. (1967). Genesis of popular but erroneous psychodiagnostic observations. Journal of Abnormal Psychology, 72(3), 193-204.
8. Cressen, R. (1975). Artistic quality of drawings and judges' evaluations of the DAP. Journal of Personality Assessment, 39(2), 132-137.
9. Deng, X., Mu, T., Wang, Y., & Xie, Y. (2022). The application of human figure drawing as a supplementary tool for depression screening. Frontiers in Psychology, 13, 865206.
10. Fatima, A., Malik, I. A., & Bhatti, S. A. (2025). Stress, Procrastination, and Government Employees: A Gender based Analysis. *Social Science Review Archives*, *3*(2), 2011-2016.
11. Fatima, A., Malik, I. A., Ahmad, A., & Aziz, T. (2025). The impact of online learning on university students' mental health during the COVID-19 pandemic: A study of academic stress, uncertainty tolerance, and psychological wellbeing. Research Journal of Psychology, 3(1). https://doi.org/10.59075/rjs.v3i1.70
12. Feher, E., Vandecreek, L., & Teglasi, H. (1983). The problem of art quality in the use of human figure drawing tests. Journal of Clinical Psychology, 39(2), 268-275.
13. Feldman, M. J., & Hunt, R. G. (1958). A relation of difficulty in drawing to ratings of adjustment based on human figure drawings. Journal of Consulting Psychology, 22(3), 217-219.
14. Gernhardt, A., Rubeling, H., & Keller, H. (2013). This is my family: Differences in children's family drawings across cultures. Journal of Cross-Cultural Psychology, 44(7), 1166-1183.
15. Gernhardt, A., Rubeling, H., & Keller, H. (2015). Cultural perspectives on children's tadpole drawings: At the interface between representation and production. Frontiers in Psychology, 6, 812.
16. Goodenough, F. L. (1926). Measurement of intelligence by drawings. World Book.
17. Guo, H., Feng, B., Ma, Y., Zhang, X., Fan, H., Dong, Z., Chen, T., & Gong, Q. (2023). Analysis of the screening and predicting characteristics of the house-tree-person drawing

test for mental disorders: A systematic review and meta-analysis. Frontiers in Psychiatry, 13, 1041770.
18. Harris, D. B. (1963). Children's drawings as measures of intellectual maturity. Harcourt, Brace & World.
19. Imuta, K., Scarf, D., Pharo, H., & Hayne, H. (2013). Drawing a close to the use of human figure drawings as a projective measure of intelligence. PLoS ONE, 8(3), e58991.
20. Jaroenkajornkij, N., Lev-Wiesel, R., & Binson, B. (2022). Use of self-figure drawing as an assessment tool for child abuse: Differentiating between sexual, physical, and emotional abuse. Children, 9(6), 868.
21. Kahill, S. (1984). Human figure drawing in adults: An update of the empirical evidence, 1967-1982. Canadian Psychology, 25(4), 269-292.
22. Kissos, L., Goldner, L., Butman, M., Eliyahu, N., & Lev-Wiesel, R. (2020). Can artificial intelligence achieve human-level performance? A pilot study of childhood sexual abuse detection in self-figure drawings. Child Abuse & Neglect, 109, 104755.
23. Koppitz, E. M. (1968). Psychological evaluation of children's human figure drawings. Grune & Stratton.
24. Lev-Wiesel, R. (1999). The use of the Machover Draw-A-Person test in detecting adult survivors of sexual abuse: A pilot study. American Journal of Art Therapy, 37(4), 106-112.
25. Lev-Wiesel, R., Trebaticka, J., Girish, M., Matzova, S., Goldner, L., Vatral, M., Weiner, N., Vasaryova, D., Oryan, I., Zaksh, Y., Hana, Z., & Hadar, D. (2024). Use of self-figure drawing as a screening tool to identify child sexual abuse, eating disorders, and suicidality. International Journal of Adolescence and Youth, 29(1), 2307501.
26. Lilienfeld, S. O., Wood, J. M., & Garb, H. N. (2000). The scientific status of projective techniques. Psychological Science in the Public Interest, 1(2), 27-66.
27. Machover, K. (1949). Personality projection in the drawing of the human figure. Charles C. Thomas.
28. Malik, I. A. (2024). MEDIATING ROLE OF PERCEIVED CRITICISM IN PREDICTINGMARITAL SATISFACTION FROM EMOTIONAL GRANULARITY. Pakistan Journal of Humanities and Social Sciences Research, 7(1).https://ojs.wumardan.edu.pk/index.php/pjhssr/article/view/273
29. Malik, I. A., Bhatti, S. A., & Butt, Z. S. (2026). Exploring the Relationship between Life Satisfaction and Anxiety in Elite Karachi Couples: A Gender-Based Study. The Critical Review of Social Sciences Studies, 4(1), 32-40.
30. Malik, I. A., Dr. Haroon-ur-Rashid, & Butt, Z. S. (2025). Investigating Power Motivationin Law Enforcement Aspirants: A Study of Pakistani University Students using Human Figure Drawing Technique. The Critical Review of Social Sciences Studies, 3(1), 85–92.https://doi.org/10.59075/1hyast36
31. Malik, I. A., & Malik, A. A. (2026). Psychological Correlates of Academic Performance: Examining Stress, Emotion Regulation, and Coping Strategies in University Students. Volume No. 04(Issue No. 01), 3201–3215. https://doi.org/10.59075/bg3nq636
32. Martlew, M., & Connolly, K. J. (1996). Human figure drawings by schooled and unschooled children in Papua New Guinea. Child Development, 67(6), 2743-2762.
33. Syeda Masooma Naqvi, Imran Asharaf Malik, & Syeda Hafsa Zainab (2025).Understanding Burnout in Remote Work Environments: Psychological

Predictors, Workplace Stressors, and Preventive Interventions. Research Journal of Psychology, 3(4),213-230. https://doi.org/10.59075/rjs.v3i4.263

34. Mirza, J., Ahmad, A., & Malik, I. A. (2024). CBT as a therapeutic approach for mitigating anxiety and panic symptoms: A case study. Pakistan Social Sciences Review, 8(4), 361–369.
35. Motta, R. W., Little, S. G., & Tobin, M. I. (1993). The use and abuse of human figure drawings. School Psychology Quarterly, 8(3), 162-169.
36. Naglieri, J. A. (1988). Draw A Person: A quantitative scoring system. Psychological Corporation.
37. Naglieri, J. A., McNeish, T. J., & Bardos, A. N. (1991). Draw A Person: Screening procedure for emotional disturbance. Pro-Ed.
38. Gardez, S. N. J., Malik, I. A., Rehman, A., & Rafique, S. (2024). Determining the prevalence and socio-demographiccorrelates of hirsutism among female residents of Muzaffarabad City, AJK. The Research of Medical ScienceReview, 2(3), 494-502.
39. Nichols, R. C., & Strumpfer, D. J. (1962). A factor analysis of Draw-a-Person test scores. Journal of Consulting Psychology, 26(2), 156-161.
40. Roback, H. B. (1968). Human figure drawings: Their utility in the clinical psychologist's armamentarium for personality assessment. Psychological Bulletin, 70(1), 1-19.
41. Rubeling, H., Keller, H., Yovsi, R. D., Lenk, M., Schwarzer, S., & Kuhne, N. (2011). Children's drawings of the self as an expression of cultural conceptions of the self. Journal of Cross-Cultural Psychology, 42(3), 406-424.
42. Scheidecker, G., Gernhardt, A., Rubeling, H., Holtmannspotter, J., & Keller, H. (2019). How young adolescents draw themselves: A comparison across three ecosocial contexts in southern Madagascar. Cross-Cultural Research, 53(3), 259-287.
43. Swensen, C. H. (1957). Empirical evaluations of human figure drawings. Psychological Bulletin, 54(6), 431-466.
44. Swensen, C. H. (1968). Empirical evaluations of human figure drawings: 1957-1966. Psychological Bulletin, 70(1), 20-44.
45. 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(2), 127-139.
46. Urban, W. H. (1963). The Draw-A-Person catalogue for interpretive analysis. Western Psychological Services.
47. Yousaf, G., Malik, I. A., & Butt, Z. S. (2025). Emotional Intelligence and Life Satisfaction: A Comparative Study of Male and Female Traffic Wardens in Lahore. Journal for Social Science Archives, 3(1), 374–386. https://doi.org/10.59075/jssa.v3i1.124