SCIENTIFIC ARTICLE

Inter- and Intrarater Reliability of the Thumb Grasp and Pinch Assessment for Children Following Index Pollicization for Congenital Thumb Hypoplasia

Wendy A. Tomhave, BA,* Kathleen M. Kollitz, MD,† Steven L. Moran, MD*†

Purpose The Thumb Grasp and Pinch (T-GAP) assessment quantifies functional hand use in children with congenital thumb hypoplasia by categorizing grasp and thumb use patterns during assessment activities that encourage a variety of grasp and pinch styles. This study aims to demonstrate interrater and intrarater reliability results of the T-GAP.

Methods A retrospective review was performed of children who had undergone index finger pollicization for congenital thumb hypoplasia and subsequent evaluation with videotaping of the T-GAP assessment. Following a training period, 4 occupational therapists scored 11 T-GAP videos on 2 separate occasions, separated by at least 2 weeks. Intraclass correlation coefficients (ICCs), standard error of measurements, minimum detectable change (MDC), and Pearson correlation coefficients were calculated.

Results The T-GAP raw scores were 16 to 55, demonstrating a range of mild to severe hand grasp differences. The ICCs for the interrater reliability trials were 0.887 and 0.901. Intrarater ICCs were all above 0.88. The MDC for each trial was 8.1 and 6.7 points. Pearson correlation coefficients calculated for each rater and each pair of raters were above 0.8 in all cases.

Conclusions Interrater and intrarater reliability testing results for the T-GAP were excellent in all cases; this strongly suggests that results from T-GAP assessments are reliable. The high ICCs suggest that raters can classify and score children's hand function consistently.

Clinical relevance This study, in conjunction with previous work, suggests that the T-GAP may be an ideal approach to assessing the outcomes of pollicization and provide a means of ongoing assessment of children's grip and pinch function. *(J Hand Surg Am. 2019;44(7):618.e1-e8.*

Key words Assessment, dexterity, pollicization, reliability, thumb.

Online Exclusive

*From the *Shriners Hospitals for Children—Twin Cities, Minneapolis, MN; and the †Mayo Clinic, Rochester, MN.*

Received for publication April 5, 2018; accepted in revised form September 18, 2018.

No benefits in any form have been received or will be received related directly or indirectly to the subject of this article.

Corresponding author: Steven L. Moran, MD, Mayo Clinic, 200 First St. SW, Rochester, MN 55905; e-mail: Moran.Steven@mayo.edu.

0363-5023/19/4407-0020$36.00/0
https://doi.org/10.1016/j.jhsa.2018.09.009

THE THUMB IS THOUGHT TO CONTRIBUTE up to 40% of the hand's usefulness[1] and is a specialized organ that cannot be replaced by any other digit.[2] The saddle-shaped carpometacarpal joint allows a wide range of movement that enables the thumb to form a variety of grasp and pinch styles for handling objects of different size and shapes. In contrast, children with congenital thumb hypoplasia, with an absent or deficient thumb, may use abnormal

grasp patterns, including a side-to-side scissors pinch[3,4] or finger flexion, when holding objects. To improve hand grasp, pollicization of the index finger is a well-established treatment to create an opposable new "thumb."[2,5,6] The pollicized thumb has been shown to lack range of motion (ROM), strength, and dexterity[7–18] and these differences are accentuated in children with more severe forms of radial longitudinal deficiency.

The goal of pollicization is to enable children to use a wider variety of grasp and pinch patterns and ultimately to improve dexterous hand use in daily activities; however, outcome measures to evaluate hand function are limited. In addition, it is difficult to follow improvements in grasp and pinch over time. Dexterity assessments to measure functional outcomes following index pollicization include the Jebsen-Taylor Test of Hand Function,[8,9,11,18] the Functional Dexterity Test,[8,11] the Box and Blocks Test,[16,19] and the Nine-Hole Peg Test.[20,21] Most require repetitive handling of blocks or pegs and performance is measured as timed speed.[22–28] Dexterity assessments do not measure the quality or method of grasp employed by the child to complete the task. Finally, many of these tests may be successfully completed without use of the thumb.

Several authors have created other activity-based outcomes that focus on how the thumb is used for handling large and small objects, peeling a sticker, or handling a tape measure.[9,11,17,29–31] None have been adopted across more than 3 subsequent studies because most are based on observation to evaluate performance or lack a well-developed scale. There are no standardized activity-based participation measures that classify grasp and thumb use patterns for children with radial longitudinal deficiency with absent or unstable thumbs.

The thumb grasp and pinch (T-GAP) assessment was developed to address the need for a standardized assessment that challenges the hand during goal-specific object handling activities and grades the level of hand dexterity based on methods of grasp and thumb use. The assessment does not focus on speed of task completion. The T-GAP is intended to grade hand grasp and pinch function using a hierarchical scale based on normal and abnormal motor development and identify if and how children who have undergone index pollicization use their new thumb for grasp and pinch during age-appropriate play. Previous publications[32] have shown concurrent validity of the T-GAP and significant correlation between the T-GAP scores and dexterity measures and construct validity between strength and ROM of the thumb and the T-GAP scores. The goal of the current study was to establish inter- and intraobserver reliability of the T-GAP in children following index pollicization.

METHODS

The T-GAP is a video-based evaluation intended for children ages 18 months to 18 years with congenital thumb hypoplasia. It includes 9 age-appropriate tasks based on developmental stages of hand function for 3 separate age groups (18 months–4 years, 5–7 years, and 8–18 years). The tasks were selected in order to evoke specific object manipulation routines: large grasp, medium grasp, small grasp, lateral pinch, tip pinch, tripod pinch, manipulation, resistance, school activities and activities of daily living (Table 1). Children were allowed to employ any grasp they liked in order to complete the task. Each task includes a score guideline written in italics, which identifies the specific component of the activity to be rated by the evaluator. For example, the activity of daily living task "put sock on over toes" for ages 18 months to 4 years, is rated by the evaluator specifically on "how the sock is opened" prior to putting the sock over the toes. A variety of grasp styles may be employed throughout the task, so it is necessary to focus on a specific aspect of each activity during evaluation. In the case of 2 grasp styles used equally, the higher score is awarded. The tasks are video recorded during a 5- to 10-minute semistructured play session and scored during a subsequent viewing.

Each grasp is assigned points using a 7-point hierarchical scale, ranging from the most primitive grasp styles, flexing or scissoring the fingers without inclusion of the thumb (0–3 points), to more sophisticated patterns that engage the thumb in distal patterns of radial digital opposition (4–7 points) (Table 2). If a child is unable to complete the portion of the task that is scored, the score is 0. The grasp and thumb-use scoring hierarchy of the T-GAP was modeled after principles of normal motor development of the hand. Palmar grips that immobilize an object are awarded fewer points than digital distal patterns that permit manipulation.[33,34] Ulnar grasps are less developed than radial patterns, which integrate the thumb in patterns of opposition for key pinch, tip pinch and fine radial digital precision.[35–38] The grasp and pinch style scoring hierarchy and photo examples of each grasp style can be found in Figure 1. The T-GAP has a point value range of 0 to 63, with a higher score indicating a greater amount of thumb opposition during tasks. Total score is calculated as the sum of all points awarded for the 9 tasks.

TABLE 1. T-GAP Activities

Activity	Ages 18 mo–4 y	Ages 5–7 y	Ages 8–18 y
Tip pinch	Pick up 3 Cheerios 1 at a time and release into a film container. *Score how the Cheerio is held.*	Pick up 3 pennies 1 at a time and release into a piggy bank. *Score how the penny is held.*	Thread 5 plastic beads onto a zip tie. *Score how the bead is held.*
Lateral key pinch	Open a zippered pencil case and remove 2 markers. *Score how the zipper tab is held.*	Turn a key to open a 1 3/16" Master Padlock. *Score how the key is held.*	Turn a key to open a 1 3/16" Master Padlock. *Score how the key is held.*
Small grasp	Pull cap off a large-diameter Crayola marker. *Score how the marker is held.*	Pull cap off a small-diameter Crayola marker. *Score how the marker is held.*	Remove cap from ballpoint pen. *Score how the pen is held.*
Medium grasp	Separate 5 Duplos that are stacked together. *Score how the Duplos are held.*	Turn end of kaleidoscope 3 times. *Score how the kaleidoscope is held.*	Make a telescope with a 5 × 7" sheet of paper and place rubber band over it. *Score how the paper tube is held.*
Large grasp	Open a 4 oz. jar of bubbles. *Score how the bottle is held.*	Twist cap from a 3"-diameter peanut butter jar. *Score how the container is held.*	Twist cap off from a 3"- diameter peanut butter jar. *Score how the container is held.*
Manipulation	Form Play-Doh into a bowl. *Score how the Play-Doh is held.*	Form Play-Doh into a bowl. *Score how the Play-Doh is held*	Rotate a pencil 3 times in a hand-held pencil sharpener- *Score how the pencil is held-*
Resistance	Open a drawstring bag-*Score how the bag is held when opened-*	Pull back the foam pull on slingshot. *Score how the foam pull is held.*	Pull back the foam pull on slingshot. *Score how the foam pull is held.*
School	Open a box of 8 crayons and remove 1. *Score how the crayon is held.*	Color inside a circle with a crayon. *Score how the crayon is held.*	Write name with a no. 2 pencil. *Score how the pencil is held.*
Activity of daily living	Put a sock on over the toes *Score how the sock is opened.*	Tie shoelaces into a knot. *Score how the laces are held.*	Tie shoelaces into a bow. *Score how the laces are held.*

Data collection

After institutional review board approval, charts were retrospectively reviewed for children who had been followed for pollicization between 2007 and 2014. Children were required to have at least 1 video assessment of the T-GAP on file for inclusion. A team of 4 certified occupational therapists reviewed and scored each video on 2 occasions a minimum of 2 weeks apart. A 2-hour training and discussion period was conducted prior to the first scoring session. For reliability testing, therapists scored each video independently in a private room free from distractions. Therapists were given age-appropriate T-GAP score forms and the T-GAP scoring hierarchy, which includes several photo examples of each grasp style. The therapist rated each task the child performed, and the points awarded for each of the 9 tasks were later summed to produce a final T-GAP score. For both scoring sessions, all videos were available, and the therapist chose the order when scoring the videos to minimize the effects of order on scoring.

Statistical analysis

Two-way random intraclass correlation coefficients (ICCs) were used to calculate interrater reliability for absolute agreement for each trial.[39,40] Pearson correlation coefficients (PCCs) were also calculated for each pair of raters. The standard error of measurement (SEM) and SEM percentage were reported.[41] Finally, the minimum detectable change (MDC), the magnitude of change necessary to exceed the measurement error for each trial, was calculated as the 95% confidence interval around the SEM.[42] Intrarater reliability was assessed using ICCs (1-way random) for absolute agreement for each rater between trial 1 and trial 2. The PCCs between trial 1 and trial 2 were reported for each rater. For both ICCs and PCCs, values greater than 0.9 were considered excellent, 0.75 to 0.9 were considered strong, values between 0.6 and 0.74 were considered moderate, values between 0.4 and 0.59 were considered fair, and values 0.39 or less were considered poorly correlated.

TABLE 2. T-GAP Scoring Hierarchy

Grasp/Pinch Type	Point Value
No Use of Thumb	
No grasp or pinch; passive stabilization with hand	0
Palmar grasp; fingers to palm	1
Ulnar scissor grasp; between little/ring fingers	2
Radial scissor grasp; between index/middle or middle/ring fingers	3
Use of Thumb	
Cylindrical grasp; opposed thumb to all fingers	4
Lateral key pinch; thumb to index finger	5
Tip pinch; thumb to index finger tip	6
Tripod pinch; thumb to distal index/middle fingers	7

RESULTS

Eleven children between 25 months and 13 years of age (average, 5.77 years) had a previously recorded T-GAP on file and were included for interrater and intrarater reliability testing. One video per child was scored, meaning 1 hand was excluded in bilaterally affected children. All children had congenital thumb hypoplasia, including Manske types IIIB (2), IV (4) and V (5), requiring index pollicization. Six had varying radius involvement and were classified as Bayne 1 (2), Bayne 2 (2), and Bayne 4 (2). Four underwent centralization of the wrist to improve hand to forearm alignment. One had a triphalangeal thumb. Three had bilateral pollicization. The right hand was evaluated in 3 children and the left in 8.

The T-GAP has a theoretical minimum score of 0 and maximum score of 63. The lowest score collected was 16 and the highest score collected was 55, indicating that there were no floor or ceiling effects. The raw data appear in Appendix A (available on the *Journal*'s Web site at www.jhandsurg.org) .

The ICC for interrater reliability for trial 1 was 0.887 (95% confidence interval [95% CI], 0.732−0.965). The SEM was 2.92, SEM% was 8.1%, and the MDC was 8.1 points. For trial 2, the ICC was nearly identical, at 0.901 (95% CI, 0.762−0.969). For trial 2, the SEM was 2.5, the SEM% was 7.1%, and the MDC was 7 points. The PCCs between each pair of raters are presented in Table 3. Intrarater reliability ICCs were similarly quite high, ranging from 0.881 to 0.945. Values by rater with 95% CIs and PCCs are presented in Table 4.

DISCUSSION

Knowledge about the qualitative development of hand movements in young children is scarce.[43] Presently, time performance and precision demands are the most common features of quantitative dexterity testing[44] and have been used as proxies for thumb functional outcomes following surgery for index pollicization. Qualitative hand use assessments have been described in the literature, but none have become standardized for this population. The T-GAP provides a new quantitative means of evaluating functional dexterity by measuring how the fingers and thumb work together to form grasp and pinch when challenged to manipulate objects of different sizes and shapes.

This study sought to evaluate the interrater and intrarater reliability of the T-GAP test. Our interrater results compared the judges, or raters, for each scoring session, using ICCs to understand how strongly units in the same group resemble each other. Our results showed strong correlation in the T-GAP scoring comparisons. Intrarater testing, which calculates the degree of agreement from one trial to another for each rater, also demonstrated strong agreement, with ICCs ranging from 0.878 to 0.915. These results suggest that the T-GAP has potential to be a reliable measure of thumb use in congenital thumb hypoplasia.

The MDC is the minimum amount of change in the rater's score that reflects a true change between the 2 trials rather than measurement error. The MDC for the 2 interrater reliability trials was 8.1 and 7.0 points. The small decrease in the trial 2 score may be a learning effect.

The PCCs between scores for each pair of raters ranged from 0.887 to 0.958 in trial 1 and 0.904 to 0.993 in trial 2. The small increase seen in trial 2 may be a practice effect. The PCCs between each pair of raters were above 0.881 in all cases in each trial. We interpreted ICCs and PCCs above 0.9 as excellent and above 0.75 as strong. All of our measured values surpassed 0.85.

There are several strengths of the T-GAP design that resulted in strong T-GAP reliability results. First, the test uses a simple scoring system. Second, all activities include a well-defined aspect to be rated, which is written in italics on the score form and which helps to focus the raters' attention on the same step in completion of the task. Finally, evaluator training included the review of both still photos and video examples. Raters were allowed to ask clarifying questions during the training, which likely resulted in a deeper understanding of the grasp patterns defined by the T-GAP.

Two other standardized assessments of pediatric hand function have also reported inter- and intraobserver reliability data comparable with the T-GAP results. The

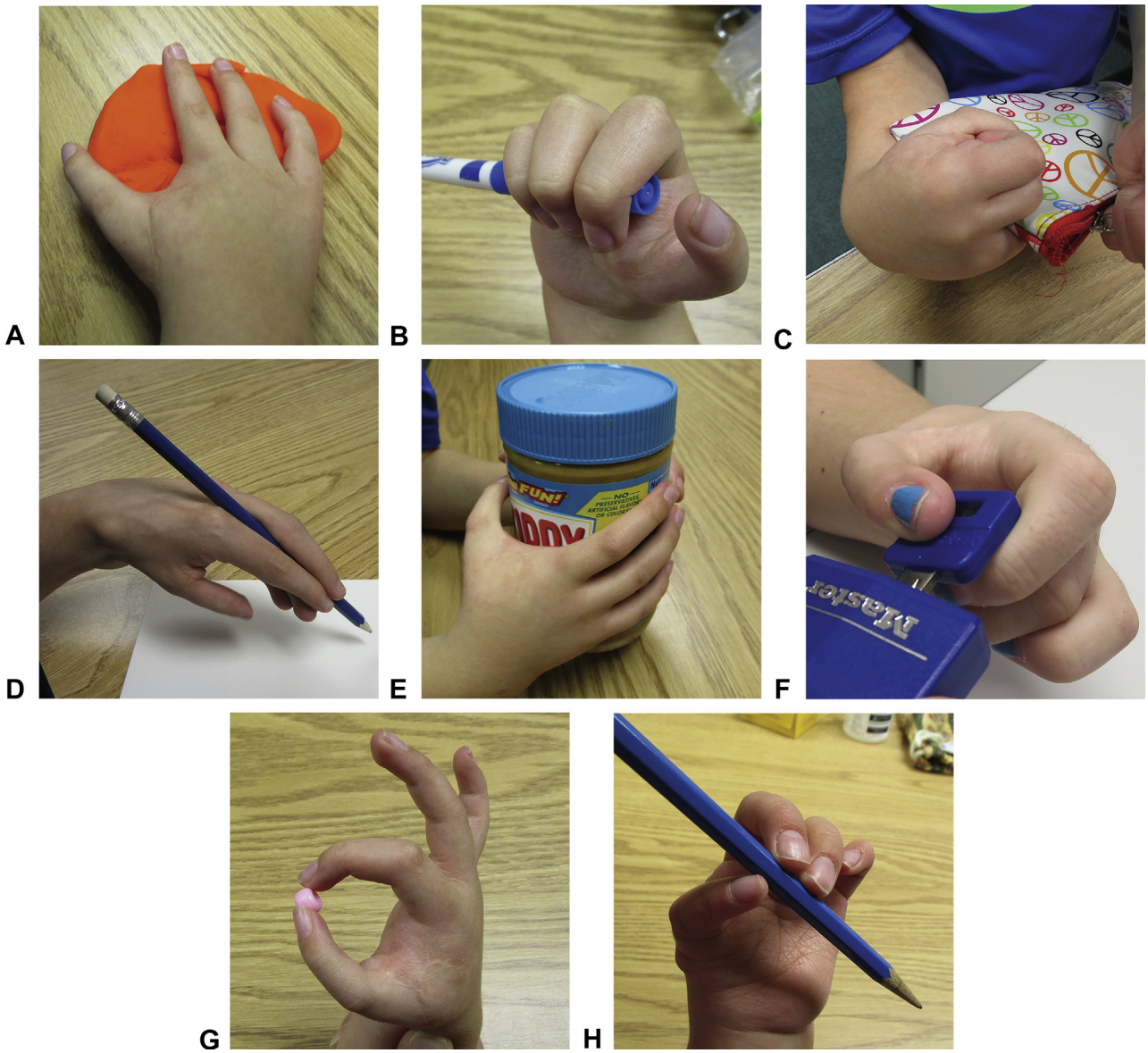

FIGURE 1: Depictions of the grasp patters defined in the T-GAP. **A** No grasp of the object; the patient is assigned 0 points if the task cannot be completed. **B** Palmar grasp (no thumb use) is assigned 1 point. **C** Ulnar scissor pinch between the 2 ulnar-most fingers, 2 points. **D** Radial scissor pinch between the 2 radial-most fingers, 3 points. **E** Cylindrical grasp is a large grasp that employs the thumb, 4 points. **F** Key pinch, in which the thumb is opposed to the lateral aspect of the neighboring finger, 5 points. **G** Tip pinch, in which the thumb is opposed to the tip of the adjacent finger in a static, nonmanipulative grasp, 6 points. **H** Tripod pinch between the thumb and the adjacent 1 or 2 fingers in a manipulative, dynamic grasp, 7 points.

Assisting Hand Assessment (AHA)[45] evaluates the effectiveness of the affected hand during bimanual activities in children with unilateral upper extremity impairment. Twenty-two test items are rated, using a 4-point scale, for general arm use, grasp-release, fine motor, coordination, and pace during a semistructured play session. Interrater and intrarater reliability were measured following AHA training and certification by 20 raters who then graded 8 children, with sum scores of 0.97 and 0.99. In addition, the Shriner's Hospital Upper Extremity Assessment (SHUEE)[43] is a video-based tool that includes dynamic positional analysis of upper extremity function in children with hemiplegic cerebral palsy. The SHUEE includes 16 tasks and a scoring scale from 0 to 3, with higher scores awarded for alignment toward extension of the elbow, wrist and fingers, forearm supination, and thumb abduction. Following a standardized training program, 4 raters graded 11 subjects with intraobserver reliability 0.98 and interobserver reliability 0.89.

Clinically, we use the T-GAP to track patients' function over time. We combine this score with other

TABLE 3. Pearson Correlations Coefficients Between Trial 1 and Trial 2 for Each Rater

Trial 1 Pearson Correlation Coefficients					
	Rater 1	Rater 2	Rater 3	Rater 4	Rater 5
Rater 1		0.943	0.960	0.929	0.983
Rater 2	0.943		0.948	0.974	0.970
Rater 3	0.960	0.948		0.922	0.964
Rater 4	0.929	0.974	0.922		0.943
Rater 5	0.983	0.970	0.964	0.943	
Trial 2 Pearson Correlation Coefficients					
	Rater 1	Rater 2	Rater 3	Rater 4	Rater 5
Rater 1		0.963	0.908	0.918	0.946
Rater 2	0.963		0.920	0.967	0.964
Rater 3	0.908	0.920		0.873	0.897
Rater 4	0.918	0.967	0.873		0.929
Rater 5	0.946	0.964	0.879	0.929	

TABLE 4. Intrarater Reliability by Rater Between Trial 1 and Trial 2

	Rater 1	Rater 2	Rater 3	Rater 4	Rater 5
ICC (95% CI)	0.965 (0.902–0.988)	0.960 (0.890–0.986)	0.919 (0.773–0.971)	0.940 (0.832–0.979)	0.954 (0.868–0.984)
PCC	0.932	0.952	0.895	0.879	0.907

measures of strength, ROM, and satisfaction with function. We have found the T-GAP to be a useful adjunct for clinical decisions when opponensplasty or thumb stabilization procedures are being considered as well as when nonsurgical therapies are recommended.

Our results are limited by the relatively small number of videotaped T-GAP assessments reviewed. This was primarily dictated by the available videos that met our inclusion criteria for a retrospective review and of children post-pollicization. Before the T-GAP could be widely adopted, it should be validated independently by another group to demonstrate not only that it is easy to learn and administer but also that similar levels of reliability and reproducibility can be achieved outside of our patient population and institution. Future areas of investigation can include use of the T-GAP in distinguishing changes in hand function before and after surgery for centralization, pollicization, and opponensplasty. In addition, the T-GAP may be considered to understand the impact of various surgeries and rehabilitation therapies with other hand conditions that affect grasp and pinch function.

In conclusion, the current study provides evidence for strong inter- and intrarater reliability of the T-GAP. Previous publications by the authors have suggested evidence for construct validity (that the test measures thumb use) and concurrent validity (agreement with other accepted measurements). Standardized tests are much needed for children with congenital thumb hypoplasia to measure thumb use and the impact of thumb absence or a poorly functioning thumb on hand function. Thumb use may be a new variable by which pollicization may be assessed and is more directly related to the goals of surgery than dexterity.[32] Understanding atypical grasp and thumb use patterns could potentially facilitate the choice of strategies in the therapeutic process, determine the success of surgical treatment, and aid in future surgical decisions.

ACKNOWLEDGMENTS

The authors wish to acknowledge Barb Knudson, OTR/L, Kae Crawford, OTR/L, Betsy Houser, OTR/L, and Donabelle Hanson, RPT, for their contributions to this study.

REFERENCES

1. Vekris MD, Beris AE, Lykissas MG, Soucacos PN. Index finger pollicization in the treatment of the congenitally deficient thumb. *Ann Plast Surg*. 2011;66(2):137–142.
2. Taghinia AH, Littler JW, Upton J. Refinements in pollicization: a 30-year experience. *Plast Reconstr Surg*. 2012;130(3):423e–433e.
3. Kozin SH, Weiss AA, Webber JB, Betz RR, Clancy M, Steel HH. Index finger pollicization for congenital aplasia or hypoplasia of the thumb. *J Hand Surg Am*. 1992;17(5):880–884.
4. Sykes PJ, Chandraprakasam T, Percival NJ. Pollicisation of the index finger in congenital anomalies. A retrospective analysis. *J Hand Surg Br*. 1991;16(2):144–147.
5. Buck-Gramcko D. Pollicization of the index finger. Method and results in aplasia and hypoplasia of the thumb. *J Bone Joint Surg Am*. 1971;53(8):1605–1617.
6. Foucher G, Medina J, Lorea P, Pivato G. Principalization of pollicization of the index finger in congenital absence of the thumb. *Tech Hand Up Extrem Surg*. 2005;9(2):96–104.
7. Goldfarb CA, Monroe E, Steffen J, Manske PR. Incidence and treatment of complications, suboptimal outcomes, and functional deficiencies after pollicization. *J Hand Surg Am*. 2009;34(7):1291–1297.
8. Aliu O, Netscher DT, Staines KG, Thornby J, Armenta A. A 5-year interval evaluation of function after pollicization for congenital thumb aplasia using multiple outcome measures. *Plast Reconstr Surg*. 2008;122(1):198–205.
9. Manske PR, Rotman MB, Dailey LA. Long-term functional results after pollicization for the congenitally deficient thumb. *J Hand Surg Am*. 1992;17(6):1064–1072.
10. Kollitz KM, Tomhave WA, Van Heest AE, Moran SL. Change in hand function and dexterity with age after index pollicization for congenital thumb hypoplasia. *Plast Reconstr Surg*. 2018;141(3):691–700.
11. de Kraker M, Selles RW, van Vooren J, Stam HJ, Hovius SE. Outcome after pollicization: comparison of patients with mild and severe longitudinal radial deficiency. *Plast Reconstr Surg*. 2013;131(4):544e–551e.
12. Netscher DT, Aliu O, Sandvall BK, et al. Functional outcomes of children with index pollicizations for thumb deficiency. *J Hand Surg Am*. 2013;38(2):250–257.
13. Staines KG, Majzoub R, Thornby J, Netscher DT. Functional outcome for children with thumb aplasia undergoing pollicization. *Plast Reconstr Surg*. 2005;116(5):1314–1315.
14. Roper BA, Turnbull TJ. Functional assessment after pollicisation. *J Hand Surg Br*. 1986;11(3):399–403.
15. Clark DI, Chell J, Davis TR. Pollicisation of the index finger. A 27-year follow-up study. *J Bone Joint Surg Br*. 1998;80(4):631–635.
16. Ceulemans L, Degreef I, Debeer P, De Smet L. Outcome of index finger pollicisation. *J Hand Microsurg*. 2010;2(1):13–17.
17. Ekblom AG, Dahlin LB, Rosberg H-E, Wiig M, Werner M, Arner M. Hand function in children with radial longitudinal deficiency. *BMC Musculoskelet Disord*. 2013;14:116.
18. Zlotolow DA, Tosti R, Ashworth S, Kozin SH, Abzug JM. Developing a pollicization outcomes measure. *J Hand Surg Am*. 2014;39(9):1784–1791.
19. Tonkin MA, Boyce DE, Fleming PP, Filan SL, Vigna N. The results of pollicization for congenital thumb hypoplasia. *J Hand Surg Eur Vol*. 2015;40(6):620–624.
20. Ekblom AG, Dahlin LB, Rosberg H-E, Wiig M, Werner M, Arner M. Hand function in adults with radial longitudinal deficiency. *J Bone Joint Surg Am*. 2014;96(14):1178–1184.
21. Waljee JF, Carlozzi N, Franzblau LE, Zhong L, Chung KC. Applying Patient-Reported Outcomes Measurement Iinformation System to assess upper extremity function among children with congenital hand differences. *Plast Reconstr Surg*. 2015;136(2):200e–207e.
22. Lightdale-Miric N, Mueske NM, Lawrence EL, et al. Long term functional outcomes after early childhood pollicization. *J Hand Ther*. 2015;28(2):158–166.
23. Poole JL, Burtner PA, Torres TA, et al. Measuring dexterity in children using the Nine-hole Peg Test. *J Hand Ther*. 2005;18(3):348–351.
24. Mathiowetz V, Volland G, Kashman N, Weber K. Adult norms for the Box and Block test of manual dexterity. *Am J Occup Ther*. 1985;39(6):386–391.
25. Jongbloed-Pereboom M, Nijhuis-van der Sanden MW, Steenbergen B. Norm scores of the box and block test for children ages 3-10 years. *Am J Occup Ther*. 2013;67(3):312–318.
26. Gogola GR, Velleman PF, Xu S, Morse AM, Lacy B, Aaron D. Hand dexterity in children: administration and normative values of the functional dexterity test. *J Hand Surg Am*. 2013;38(12):2426–2431.
27. Aaron DH, Jansen CW. Development of the Functional Dexterity Test (FDT): construction, validity, reliability, and normative data. *J Hand Ther*. 1992;16(1):12–21.
28. Jebsen RH, Taylor N, Trieschmann RB, Trotter MJ, Howard LA. An objective and standardized test of hand function. *Arch Phys Med Rehabil*. 1969;50(6):311–319.
29. Lee-Valkov PM, Aaron DH, Eladoumikdachi F, Thornby J, Netscher DT. Measuring normal hand dexterity values in normal 3-, 4-, and 5-year-old children and their relationship with grip and pinch strength. *J Hand Ther*. 2003;16(1):22–28.
30. Percival NJ, Sykes PJ, Chandraprakasam T. A method of assessment of pollicisation. *J Hand Surg Br*. 1991;16(2):141–143.
31. Goldfarb CA, Wustrack R, Pratt JA, Mender A, Manske PR. Thumb function and appearance in thrombocytopenia: absent radius syndrome. *J Hand Surg Am*. 2007;32(2):157–161.
32. Kollitz KM, Tomhave W, Van Heest AE, Moran SL. A new, direct measure of thumb use in children after index pollicization for congenital thumb hypoplasia. *J Hand Surg Am*. 2018;43(11):978–986.e1.
33. Napier JR. The prehensile movements of the human hand. *J Bone Joint Surg Br*. 1956;38-B(4):902–913.
34. Elliott JM, Connolly KJ. A classification of manipulative hand movements. *Dev Med Child Neurol*. 1984;26(3):283–296.
35. Erhardt RP. *Developmental Hand Dysfunction: Theory, Assessment, and Treatment*. 2nd ed. Austin, TX: Pro-Ed Inc.; 1994.
36. Folio MR, Fewell RR. *Peabody Developmental Motor Scales, PDMS-2*. 2nd ed. Austin, TX: Pro-Ed Inc.; 2000.
37. Landsmeer JM. Power grip and precision handling. *Ann Rheum Dis*. 1962;21:164–170.
38. Pehoski C. Object manipulation in infants and children. *Hand Function in the Child: Foundations for Remediation*. 2nd ed. St. Louis: Mosby Inc.; 1995:136–153.
39. Saito Y, Sozu T, Hamada C, Yoshimura I. Effective number of subjects and number of raters for inter-rater reliability studies. *Stat Med*. 2006;25(9):1547–1560.
40. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15(2):155–163.
41. de Vet HC, Terwee CB, Knol D, Bouter LM. When to use agreement versus reliability measures. *J Clin Epidemiol*. 2006;59(10):1033–1039.
42. Beckerman H, Roebroeck ME, Lankhorst GJ, Becher JG, Bezemer PD, Verbeek AL. Smallest real difference, a link between reproducibility and responsiveness. *Qual Life Res*. 2001;10(7):571–578.
43. Davids JR, Peace LC, Wagner LV, Gidewall MA, Blackhurst DW, Roberson WM. Validation of the Shriners Hospital for Children Upper Extremity Evaluation (SHUEE) for children with hemiplegic cerebral palsy. *J Bone Joint Surg Am*. 2006;88(2):326–333.
44. Duff SV, Aaron DH, Gogola GR, Valero-Cuevas FJ. Innovative evaluation of dexterity in pediatrics. *J Hand Ther*. 2015;28(2):144–150.
45. Krumlinde-Sundholm L, Holmefur M, Kottorp A, Eliasson AC. The Assisting Hand Assessment: current evidence of validity, reliability, and responsiveness to change. *Dev Med Child Neurol*. 2007;49(4):259–264.

APPENDIX A. T-GAP Interrater Intrarater Data

Subject	Rater 1, Trial 1	Rater 2, Trial 1	Rater 3, Trial 1	Rater 4, Trial 1	Rater 1, Trial 2	Rater 2, Trial 2	Rater 3, Trial 2	Rater 4, Trial 2
1	35	37	36	31	32	38	34	35
2	31	37	33	28	36	38	40	33
3	30	28	29	31	25	34	30	30
4	31	33	36	28	34	35	34	32
5	36	38	32	34	35	35	34	32
6	16	21	21	18	18	21	16	16
7	35	39	36	36	37	39	41	36
8	44	49	48	43	38	46	46	45
9	33	40	38	42	34	37	31	34
10	46	53	51	55	49	46	52	46
11	42	47	39	40	42	45	40	42