International Journal of Psychology and Psychological Therapy, 2020, 20, 2, 177-188
20th ANNIVERSARY EDITION // EDICIÓN XX ANIVERSARIO

Psychometric Properties and Measurement Invariance of the Generalized Tracking Questionnaire-Children

Cristhian A Espitia Pereira, Jessica M Ayala Ávila, María B García Martín*
Fundación Universitaria Konrad Lorenz, Colombia
Juan C Suárez Falcón
Universidad Nacional de Educación a Distancia, Madrid, España
Francisco J Ruiz
Fundación Universitaria Konrad Lorenz, Colombia

Abstract

Self-report measures of generalized patterns of rule-following are being developed in the last years including the Generalized Pliance Questionnaire (GPQ) and the Generalized Tracking Questionnaire (GTQ). Specifically, the GTQ is a 11-item self-report developed for adults that has shown excellent psychometric properties, a one-factor structure, and criterion validity in view of its positive correlations with executive functions tests. This study aims to develop the children's version of the GTQ (i.e., Generalized Tracking Questionnaire-Children, GTQ-C). In so doing, we adapted the vocabulary of some of the GTQ items and added an additional item (i.e., the GTQ-C consisted of 12 items). Afterward, the GTQ for children (i.e., GTQ-C) was administered to a sample of 730 Colombian children and adolescents from 7 to 17 years. A cross-validation study was conducted to analyze the factor structure of the questionnaire. The analysis showed that a one-factor structure showed a good fit to the data. All items of the GTQ-C showed good discrimination indexes and the whole questionnaire showed adequate internal consistency. The GTQ-C showed measurement invariance across gender and age group. In conclusion, the GTQ-C seems to be a valid and reliable measure of generalized tracking for children that might be used to analyze the developmental trajectories of tracking and its relationship with other relevant behavioral processes.

Key words: tracking, rule-governed behavior, Relational Frame Theory, ACT, children.

How to cite this paper: Espitia-Pereira CA, Ayala-Ávila JM, García-Martín MB, Suárez-Falcón JC, & Ruiz FJ (2020). Psychometric Properties and Measurement Invariance of the Generalized Tracking Questionnaire-Children. *International Journal of Psychology & Psychological Therapy, 20, 2,* 177-188.

Novelty and Significance

What is already known about the topic?

- Generalized tracking involves the skill to establish functional relationships among behaviors and their consequences.
- The Generalized Tracking Questionnaire (GTQ) has been recently developed to measure generalized tracking in adults.

What this paper adds?

- Adaptation of the GTQ to children (GTQ-C).
- The GTQ-C showed acceptable internal consistency and a one-factor structure.
- The GTQ-C showed measurement invariance across gender and group age.

Rule-governed behavior (RGB) is a key concept to explain complex human behavior from the standpoint of behavior analysis and contextual behavioral science (Hayes, Barnes-Holmes, & Roche, 2001; Hayes, Barnes-Holmes, & Wilson, 2012; Skinner, 1966; Vaughan, 1989). The term was coined by Skinner (1966), who differentiated it from contingency-shaped behavior (CSB). CSB is present in human and nonhuman animals and consists of behavior directly controlled by its history of consequences (Catania, Shimoff, & Matthews, 1989). Conversely, RGB is exclusive of human beings and can be

* Correspondence: María B García Martín, Fundación Universitaria Konrad Lorenz, Carrera 9 bis, 62-43, Bogotá, Cundinamarca, Colombia. Email: mariabelengarciamartin2@gmail.com

defined as behavior controlled by antecedent verbal stimuli provided by another person or by the own individual (Hayes & Hayes, 1989; Skinner, 1957).

By its very definition, CSB is sensitive to changes in reinforcement schedules. For instance, if a pigeon's behavior of pecking is controlled by continuous reinforcement and the reinforcement schedule is changed for a fixed-interval one, the pigeon will soon begin to peck almost only when the interval of time is finishing. Importantly, this sensitivity to the change in contingencies has been observed in preverbal children (e.g., Lowe, Beasty, & Bentall, 1983), but verbal human participants tend to perform differently. Specifically, verbal human participants usually show the persistence of behavior patterns when the schedule conditions change, even at the cost of losing access to reinforcers (e.g., Hayes, Brownstein, Zettle, Rosenfarb, & Korn, 1986; Lowe, Harzem, & Hughes, 1978; Weiner, 1969). This phenomenon is usually called insensitivity to contingencies (e.g., Shimoff, Catania, & Matthews, 1981) and was soon attributed to the effect of the experimental instructions and/or the participants' self-talk in the form of descriptions of contingencies (Bijou, 1958; Catania, 2006; Lowe *et alia*, 1978). In other words, verbal human participants might be more sensitive to the antecedent verbal functions of rules, which are linked to their learning history of RGB, than to the actual contingencies contacted when performing the behavior (Hayes, Zettle, & Rosenfarb, 1989).

Several behavior-analytic accounts of RGB have been developed during the last decades (e.g., Schlinger, 1990; Zettle & Hayes, 1982). Among them, relational frame theory (RFT; Hayes *et alia*, 2001) provides the most complete approach because it explains the core characteristics and behavioral processes involved in RGB, including the conceptualization of verbal stimuli, the generation, meaning and understanding of rules, and rule-following (Hayes & Hayes, 1989; Hayes *et alia*, 1989; Luciano, Valdivia, Cabello, & Hernández, 2009).

Briefly, RFT defines verbal stimuli as events that acquire functions for the listener through derived relational responding (or relational framing). RFT defines multiple forms of relational framing such as coordination (i.e., A is the same as B), comparative (i.e., A is more than B, A is less than B), hierarchical (i.e., A belongs to B, A contains B), temporal (i.e., A goes before B, A goes after B), causal (i.e., if A, then B), and deictic ("I-you," "here-there," "now-then") (Barnes-Holmes *et alia*, 2001; Hayes & Hayes, 1989). These relational framings have three properties: mutual entailment, combinatorial entailment, and transformation of functions. Mutual entailment implies the bidirectionality of stimuli relations: if a verbal-competent child is told that A is related to B; then he or she will derive that B is related to A (e.g., if car is the same as coche, then coche is the same as car). Combinatorial entailment involves combining two or more mutual entailments: if the child is told that A is related to B, and B is related to C; he or she will derive that A is related to C, and C is related to A (e.g., if car is the same as coche, and coche is the same as automobile; then car is the same as automobile and automobile is the same as car). Lastly, transformation of functions means that the functions of a stimulus will provoke a change in the functions of the stimuli related to it: If C acquires a reinforcing function, this will provoke a change in function in stimuli A and B (e.g., if the child has a reinforcing experience with an automobile, this function will be transferred to the stimuli "car" and "coche").

According to RFT, the higher the fluency in relational framing, the more the child will understand and produce new rules that have never been reinforced. However, although rule-understanding is a condition for rule-following, both terms refer to different behavioral processes. For instance, a child might understand the meaning of

a rule because of being fluent in the relational framings involved in the rule, but not follow it because the speaker is not a credible source of reinforcement (Barnes-Holmes *et alia*, 2001; Hayes *et alia*, 1989).

With respect to rule-following, RFT adopted the functional classification initially suggested by Zettle and Hayes (1982): pliance, tracking, and augmenting. For the sake of brevity, in this paper, we will focus only on pliance and tracking since augmenting is usually linked to them.

Pliance is the first type of rule-following developed and is due to a history of multiple exemplars in which the speaker reinforced the listener's behavior when it corresponded with the rule content (Hayes *et alia*, 1989; Hayes, Gifford, & Hayes, 1998; Luciano, Valdivia, & Ruiz, 2012). Tracking is usually developed after some practice with pliance and is due to a history of multiple exemplars in which acting according to the rule content was reinforced by the natural consequences that are derived from the way the world is arranged (Hayes *et alia*, 1989; Luciano *et alia*, 2012). Thus, the main difference between pliance and tracking is the apparent source of reinforcement for rule-following: social or arbitrary in the case of pliance and nonarbitrary in the case of tracking (Hayes *et alia*, 1989; Luciano *et alia*, 2012; Zettle & Hayes, 1982). Note that the word "apparent" emphasizes that rules are antecedent events and, as such, a rule should be classified as an instance of pliance (i.e., ply) or tracking (i.e., track) only based on the functions actualized on the listener when performing it and not according to the actual consequences contacted after following the rule. In other words, the function of the rule is determined by the listener's history in rule-following, which implies that speakers cannot reliably produce plys or tracks (Hayes *et alia*, 1989).

It is also important to highlight that plys or tracks are not defined by the content of the rule. For instance, the fact that a rule might specify a natural or nonarbitrary contingency (e.g., "Brush your teeth so that you can feel them clean and smooth") does not mean that the rule would be a track. If the rule would be a track or not would only depend on the actualized function for the listener (e.g., if the child follows the rule to obtain those sensations on the mouth, the rule would be a track; however, if the child follows the rule just to please the mother and obtain her approval, the rule would be a ply). In conclusion, confounding the content of the rule with the functions actualized in the listener is just another form of confounding the stimulus object and the stimulus functions, which is one of the main conceptual difficulties to provide functional-analytic explanations of complex human behavior (Hayes & Hayes, 1989; Kantor, 1924; Parrott, 1984). According to the latter points, the difficulties found in the experimental analysis of the functional classes of rule-following are understandable (Kissi *et alia*, 2017).

The ontogenetic origin of rule-following might explain the phenomenon of insensitivity to contingencies. A learning history mainly based on pliance will lead the individual to follow rules because of their antecedent verbal functions, linked to reinforcement provided by the speaker, without contacting the natural consequences of his or her behavior. For instance, a child might follow the instructions to solve a math problem to please the teacher, but without contacting the natural contingencies involved in the solution. If this occurs, even a slight change in the nature of the problem might lead the child to solve it incorrectly. Let's also consider a typical experimental example. If the experimental instructions actualize pliance functions in the participant, he or she will behave under the control of the apparent social reinforcement (e.g., "She said that I have to press the "P" key repeatedly" or "If I don't press the "P" key, she might get disappointed") without contacting the actual change in the experimental contingencies.

As commented above, tracking develops after some practice with pliance and when multiple exemplars are provided in which the speaker helps the child to verbally contact the natural consequences of his or her behavior (Luciano et alia, 2009). Consequently, if these interactions are lacked or the child does not have enough fluency in relational framing to contact the natural consequences, pliance will become the predominant type of rule-following for the child (Luciano *et alia*, 2012; Törneke, Luciano, & Valdivia-Salas, 2008). This context is the breeding ground for social approval to become the main source of reinforcement for the individual, which is a pattern of rule-following called generalized pliance (see a detailed explanation of the development of generalized pliance in Ruiz, Suárez Falcón, Barbero Rubio, & Flórez, 2019).

Generalized pliance, also called excessive pliance regulation, has been associated with psychopathology and behavioral ineffectiveness (Hayes, Strosahl, & Wilson, 1999; Törneke *et alia*, 2008). The empirical evidence of the relationships between generalized pliance, psychopathology, and insensitivity to contingencies is recently growing thanks to the development of the Generalized Pliance Questionnaire (GPQ; Ruiz *et alia*, 2019) and the Generalized Pliance Questionnaire-Children (GPQ-C; Salazar, Ruiz, Flórez, & Suárez Falcón, 2018). Specifically, scores on the GPQ showed positive correlations with experiential avoidance, cognitive fusion, obstruction in values, dysfunctional schemas, and emotional symptoms; and negative correlations with progress in values, life satisfaction, and mindfulness skills (Ruiz *et alia*, 2019), with the correlations obtained by the GPQ-C being very similar. Importantly, O'Connor, Byrne, Ruiz, and McHugh (2019) showed that the GPQ has criterion validity according to the strong correlations found with tests of cognitive flexibility and insensitivity to contingencies such as the Wisconsin Card Sorting Test (WCST; Grant & Berg, 1948). Similarly, Salazar, Ruiz, García Martín, and Bedoya Valderrama (2020) showed that the GPQ-C can discriminate between children with learning difficulties and control participants and that GPQ-C scores showed a strong correlation with the performance on the WCST.

Contrary to pliance, tracking is more likely to be sensitive to direct contingencies because rule-following is due to the apparent conditional relationship between the actual behavior and the consequences contacted. Furthermore, an individual with high fluency in relational framing who has also been exposed to multiple interactions in which he or she has been guided to observe and describe functional relationships among events will develop a pattern of rule-following that we call generalized tracking (Ruiz *et alia*, 2020). This generalized pattern of rule-following involves the motivation and skill to establish functional relationships among behaviors and their consequences and to adjust behavior according to them. The term generalized tracking has not been used in behavior analysis or contextual behavioral literature, although mentioning tracking as a skill has been relatively frequent (Luciano *et alia*, 2009, 2012; Luciano, Gómez, & Valdivia, 2002; Villatte, Villatte, & Hayes, 2015). It is also important to note that generalized tracking involves the individual behaving both as the speaker who derive the rule and the listener who follows it (Skinner, 1957).

Recently, a self-report measure of generalized tracking has been developed for adults. In so doing, Ruiz *et alia* (2020) conducted three studies, with a total of 1155 participants. The final version of the GTQ consists of 11 items and showed excellent psychometric properties, a one-factor structure, and measurement invariance across gender and clinical and nonclinical participants. The scores on the GTQ showed moderate to strong negative correlations with measures of generalized pliance, experiential avoidance, cognitive fusion, repetitive negative thinking, and emotional symptoms; and strong positive

correlations with valued living, general self-efficacy, and life satisfaction. Additionally, following the functional-analytic account of executive functions provided by Hayes, Gifford, and Ruckstuhl (1996) as a subset of rule-governed behavior characterized by rule flexibility, the correlations of the GTQ scores with the performance on a wide range of executive function tests were explored. The GTQ showed theoretically coherent and moderate correlations with all the tests administered, which evaluated semantic classification, different types of working memory, verbal fluency, sequential planning, and inhibitory control. These results indicate that the GTQ showed criterion validity.

The aim of the current study is to develop a children's version of the GTQ (i.e., GTQ-C) and examine its factor structure and psychometric properties. Some items of the original GTQ were adapted to be used with children and an additional item was included. Afterward, the GTQ was administered to a sample of 730 children and adolescents. A cross-validation study was conducted to assess the factor structure of the questionnaire. Measurement invariance across age group and gender was also explored. Lastly, internal consistency was explored and the mean scores on the GTQ-C across gender and group age were compared.

Method

Participants

The sample consisted of 730 participants (52.6% females) with ages ranging from 7 and 17 years (M= 12.29, SD= 2.29) and from second to eleventh grade: 1.5% in 2nd grade, 4.5% in 3rd grade, 8.2% in 4th, 14.9% in 5th, 11.4% in 6th, 12.3% in 7th, 17.3% in 8th, 16.4 in 9th, 6.3% in 10th, and 7.1% in 11th. All participants were Colombian and attended private (75.1%) or public schools (24.9%).

Instrument

Generalized Tracking Questionnaire-Children (GTQ-C). GTQ-C is the result of adapting some items of the GTQ for adults (Ruiz *et alia* 2020), adding one more item, and reducing the Likert-type scale from 7 to 5 points (5= always true, 1= never true). Higher scores reflect more generalized tracking. Items 4, 7, 8, and 11 of the GTQ were adapted by changing the wording to facilitate children's understanding. The Spanish and English versions of the GTQ-C can be seen in Appendix A and B, respectively.

Procedure

The procedure was approved by the institutional Ethics Committee. Participants were recruited from public and private schools from Bogotá (Colombia). Parents were given an informed consent form to allow children's participation. Data collection was group-based and was conducted in a regular class in the schools by a trained psychologist. First, the children signed the assent forms and were then given the questionnaire package, which included the GTQ-C and other instruments that are not relevant for the current study. The administration of the questionnaire package took approximately 15 minutes. Participants were allowed to stop participating at any given time.

Data Analysis

Two random samples of approximately equal size were obtained through the SPSS 25 to conduct a cross-validation study to analyze the factor structure of the GTQ-C.

In the first random sample, we conducted an Exploratory Factor Analysis (EFA) with the Factor 10.9.02 software (Lorenzo Seva & Ferrando, 2006). We selected the robust diagonally weighted least square estimation method (Robust DWLS) with Direct Oblimin rotation using polychoric correlations. The number of factors to retain from the EFA was determined by means of the optimal implementation of parallel analysis based on minimum rank factor analysis (PA; Timmerman & Lorenzo Seva, 2011). An assessment of essential unidimensionality was conducted by computing Unidimensional Congruence (*UniCo*), Explained Common Variance (*ECV*), and Mean of Item Residual Absolute Loadings (*MIREAL*) indexes. Values larger than .95 and .85 in *UniCo* and *ECV*, respectively, suggest that data can be treated as essentially unidimensional; whereas for the *MIREAL*, a value lower than .30 suggests unidimensionality (Ferrando & Lorenzo-Seva, 2018).

A Robust *DWLS* estimation method using polychoric correlations was adopted to conduct the Confirmatory Factor Analysis (CFA) using LISREL (version 8.71, Jöreskog & Sörbom, 1999). We computed the Satorra-Bentler chi-square test and the following goodness-of-fit indexes for the one-factor model: (a) the Root Mean Square Error of Approximation (*RMSEA*), (b) the Comparative Fit Index (*CFI*), (c) the Non-Normed Fit Index (*NNFI*), and (d) the Standardized Root Mean Square Residual (*SRMR*). According to Hu and Bentler (1999), *RMSEA* values of 0.08 represent a good fit, and values below 0.05 represent a very good fit to the data. For the *SRMR*, values below 0.08 represent a reasonable fit, and values below 0.05 indicate a good fit. With respect to the *CFI* and *NNFI*, values above 0.90 indicate well-fitting models, and values above 0.95 represent a very good fit to the data.

Additional CFAs were performed with the whole sample (N= 730) to test for metric and scalar invariances across gender and age group, following Jöreskog (2005), and Millsap and Yun-Tein (2004). In other words, we analyzed whether the item factor loadings and item intercepts are invariant across boys and girls and age (7-12 years and 13-17 years). In so doing, the relative fits of three increasingly restrictive models were compared: the multiple-group baseline model, the metric invariance model, and the scalar invariance model. The multiple-group baseline model allowed the twelve unstandardized factor loadings to vary across gender and age group. The metric invariance model, which was nested within the multiple-group baseline model, placed equality constraints (i.e., invariance) on those loadings across groups. Lastly, the scalar invariance model, which was nested within the metric invariance model, is tested by constraining the factor loadings and items' intercepts to be the same across groups. Equality constraints were not placed on estimates of the factor variances because these are known to vary across groups even when the indicators are measuring the same construct in a similar manner (Kline, 2005). For the model comparison, the *RMSEA*, *CFI*, and *NNFI* indexes between nested models were compared. The more constrained model was selected (i.e., second model versus first model, and third model versus second model) if the following criteria suggested by Cheung and Rensvold (2002) and Chen (2007) were met: (a) the difference in *RMSEA* (*ΔRMSEA*) was lower than .01; (b) the differences in *CFI* (*ΔCFI*) and *NNFI* (*ΔNNFI*) were equal to or greater than -.01.

Cronbach's alpha and McDonald's omega coefficients were computed providing percentile bootstrap 95% confidence intervals (*CI*), for the overall sample and each of the random samples, with the MBESS package in R (Kelley & Lai, 2012; Kelley & Pornprasertmanit, 2016). Descriptive data were calculated with SPSS25. A two-way analysis of variance (ANOVA) was computed to analyze differences in GTQ-C scores

across gender and age. The effect size for the ANOVA, partial eta-squared (η^2), was computed to analyze the magnitude of the differences across the independent variables. The values of η^2 can be interpreted according to the following guidelines: .01 small, .09 medium, and .25 large

RESULTS

The first random sample consisted of 357 participants and was used to conduct the EFA. Bartlett's statistic was statistically significant (605.4(66), p <.001) and the result of the Kaiser-Meyer-Olkin (*KMO*) test was fair (.79). The PA suggested extracting only one factor, which accounted for 35.7% of the variance. Table 1 shows the items of the GTQ-C, their English translation, and the factor loadings obtained in the EFA. Factor loadings were adequate for all items: from .38 (Item 5) to .59 (Item 2). Values of *UniCo* (.96) and *MIREAL* (.21) supported the essential unidimensionality of the GTQ-C, whereas the result of the *ECV* (.80) did not support that conclusion. In conclusion, the results of the conducted EFAs suggested that the GTQ-C can be treated as an essentially unidimensional measure.

As the EFA previously conducted indicated that the GTQ-C seems to be a unidimensional measure, the CFA was conducted with the second random sample (N= 373) to analyze the fit of a one-factor model. The overall fit of the one-factor model in the GPQ-C was adequate: $S\text{-}B\chi^2(54)$= 130.03, p < .05; *RMSEA*= 0.062, 90% *CI* [0.048, 0.075], *CFI*= 0.95, *NNFI*= 0.94, *SRMR*= 0.063. Modification indices recommended allowing error terms between Items 4-7 to correlate. When doing so, the fit of the one-factor model improved slightly (*RMSEA*= 0.053, *NNFI*= 0.96, *CFI*= 0.97, *SRMR*= 0.057).

Table 2 shows the results of the metric and scalar invariance analyses. Parameter invariance was supported at both the metric and scalar levels across gender and age (7-12 and 13-17 years old) because changes in *RMSEA*, *CFI*, and *NNFI* were lower than .01.

Table 1. Item Description of the GTQ-C, English Translation, and Factor Loadings from EFA with the First Random Sample.

Items	Factor loadings (N= 357)
1. Cuando veo que algo no está funcionando, intento algo diferente [When I realize something is not working, I try something different]	.52
2. Disfruto descubriendo cómo funcionan las cosas y llegando a mis propias conclusiones [I enjoy discovering how things work and reaching my own conclusions]	.59
3. Me adapto fácilmente a los cambios [I adapt to changes easily]	.42
4. Tengo facilidad para encontrar soluciones nuevas cuando me surge un problema [I am good at finding new solutions when a problem surface]	.50
5. Tomo decisiones basándome en lo que yo creo y no en lo que los demás dicen [I make decisions based on what I think and not in what other people say]	.38
6. Me gusta probar diferentes maneras de hacer las cosas para ver cuál es mejor [I like to explore different ways to do things to realize what is the best]	.50
7. Soy bueno encontrando nuevas formas de resolver los problemas que me surgen [I am good at finding new ways of solving the problems that show up]	.55
8. Si veo que algo no funciona, cambio mi forma de actuar rápidamente [If I see something that does not work, I change my form of acting quickly]	.47
9. Aprendo de las consecuencias de mis acciones con facilidad [I learn from the consequences of my actions easily]	.53
10. Cuando me doy cuenta de que estoy equivocado, cambio mi forma de pensar y actuar [When I realize I am wrong, I change my way of thinking and acting]	.40
11. Hago las cosas de acuerdo con los resultados que he obtenido en otras situaciones [I do the things according to the results I have obtained in other situations]	.51
12. Me gusta explorar y descubrir cosas diferentes [I like to explore and discover different things]	.51

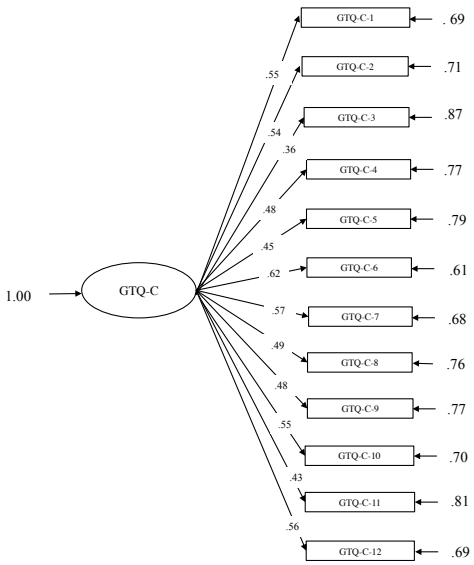

Figure 1. Completely standardized solution of the one-factor model of the GTQ-C.

Table 2. Metric and Scalar Invariance of the GTQ-C across Gender and Age.

	Model	RMSEA	ΔRMSEA	CFI	ΔCFI	NNFI	ΔNNFI
Measurement invariance across gender	MG Baseline model	.0497	--	.963	--	.954	--
	Metric invariance	.0500	-.0003	.958	-.005	.954	.000
	Scalar invariance	.0502	-.0002	.954	-.004	.953	.001
Measurement invariance across age	MG Baseline model	.0565	--	.959	--	.950	--
	Metric invariance	.0551	.0014	.957	-.002	.952	.002
	Scalar invariance	.0542	.0009	.955	-.002	.954	.002

In the overall sample, all items showed good discrimination, with corrected item-total correlations ranging from .33 (Item 3) to .45 (Item 7). Coefficient alpha (.76, 95% *CI* [.73, .78]) and coefficient omega (.76, 95% *CI* [.73, .78]) were acceptable. Similar results were obtained in the two random samples. Specifically, in the first random sample, corrected-item total correlations ranged from .30 (Item 5) to .45 (Item 2), with a coefficient alpha of .75 (95% *CI* [.71, .79]) and a coefficient omega of .75 (95% *CI* [.70, .79]). In the second random sample, corrected-item total correlations ranged from .30 (Item 3) to .48 (Item 6), with a coefficient alpha of .76 (95% *CI* [.73, .80]) and a coefficient omega of .76 (95% *CI* [.72, .80]).

Descriptive data on the GTQ-C are presented in Table 3. The two-way ANOVA revealed a small-size statistically significant effect of group age ($F(1)= 11.88$, $p= .001$, $\eta^2= .016$), with the group of children scoring higher than the group of adolescents. Neither the factor gender was statistically significant ($F(1)= 0.58$, $p= .44$, $\eta^2= .001$), nor the interaction between group age and gender ($F(1)= 1.27$, $p= .26$, $\eta^2= .002$).

DISCUSSION

Table 3. Descriptive Data of the GTQ-C.

Gender	Age	*M (SD)*
Boys	7-12 years	44.39 (7.33)
	13-17 years	41.95 (6.62)
Girls	7-12 years	43.39 (7.86)
	13-17 years	42.15 (6.94)
Overall	Overall	42.91 (7.24)

Research on generalized patterns of rule-following is being conducted in recent years thanks to the development of self-report measures (e.g., O'Connor *et alia*, 2019; Ruiz *et alia*, 2019, 2020; Salazar *et alia*, 2018, 2020; Waldeck, Pancani, & Tyndall, 2019). A measure of generalized pliance for adults (i.e., GPQ) was first developed (Ruiz *et alia*, 2019), which was subsequently adapted for children (GPQ-C; Salazar *et alia*, 2018). Both versions of the GPQ have shown criterion validity to the extent that their scores correlated with insensitivity to contingencies tasks and executive functions tests (O'Connor *et alia*, 2019; Salazar *et alia*, 2020). More recently, Ruiz *et alia* (2020) have developed a measure of generalized tracking (i.e., GTQ) that showed excellent internal consistency and a one-factor structure. Importantly, the GTQ showed significant positive correlations with a wide range of executive functions tests.

Due to the absence of a children version of the GTQ, this study aimed to develop the GTQ-C and preliminarily analyze its psychometric properties and factor structure. In so doing, we adapted the wording of some of the items of the adult version and added one more item. A cross-validation study was conducted to analyze the factor structure of the GTQ-C. The EFA computed showed that the GTQ-C seemed to have a one-factor structure and that all items had appropriate factor loadings. In conclusion, the 12 items of the GTQ-C were retained. The CFA supported the one-factor model of the GTQ-C. Additional CFAs showed the measurement invariance of the GTQ-C across gender and group. This is important because measurement invariance is a condition to compare the mean scores across groups of participants (Greiff & Scherer, 2018).

The GTQ-C showed acceptable internal consistency, with coefficient alphas and omega of .76. All items showed adequate corrected item-total correlations. No differences across gender were found on the GTQ-C, but the cohort of children scored slightly higher than the cohort of adolescents. This could be related to the social pressures usually faced by adolescents, which might lead them to be relatively more sensitive to pliance than tracking in this period. However, the cross-sectional comparison cannot control for multiple sources of influence within the cohorts that might explain the difference in scores. Accordingly, longitudinal analyses should be conducted to shed more light in this respect.

Some limitations of the current study are worth mentioning. Firstly, this study did not provide evidence of the convergent construct validity of the GTQ-C. In this regard, the original GTQ showed medium to strong negative correlations with the GPQ, experiential avoidance, cognitive fusion, repetitive negative thinking, and emotional symptoms; and strong positive correlations with valued living and life satisfaction (Ruiz *et alia,* 2020). It is intriguing what type of correlations might the GTQ-C show with these constructs across the development. For instance, as tracking develops after some practice with pliance, it seems plausible that the GPQ-C and GTQ-C would show positive correlations in children and that this correlation would turn negative during the adolescence. A similar pattern of results might be found with other relevant constructs such as psychological inflexibility and emotional symptoms. Accordingly, further studies should systematically analyze the correlations of the GTQ-C with this type of

measures. Secondly, we did not correlate the GTQ-C with behavioral measures such as the executive function tests as in Ruiz *et alia* (2020). Future studies should analyze the criterion validity of the GTQ-C having into account the developmental characteristics of pliance and tracking commented above. Thirdly, the GTQ-C provides a measure of the tendency to engage in tracking as averaged across contexts. However, a measure of tracking contextualized to specific settings (e.g., school, family, friendship, etc.) might be more adequate when working in these contexts. Lastly, further studies should analyze the psychometric properties and factor structure of the GTQ-C in other Spanish-speaking countries and other languages.

In conclusion, this study presented the adaptation of the GTQ for children (i.e., GTQ-C) and it showed good psychometric properties and a one-factor structure. The GTQ-C might promote research on functional classes of RGB in applied settings and might be used to establish the developmental trajectories of generalized patterns of rule-following such as generalized pliance and generalized tracking.

References

Barnes-Holmes D, O'Hora D, Roche B, Hayes SC, Bissett RT, & Liddy F (2001). Understanding and verbal regulation. In SC Hayes, D Barnes-Holmes, & B Roche (Eds.), *Relational Frame Theory. A post-Skinnerian account of human language and cognition* (pp. 103-117). New York: Plenum Press.

Bijou S (1958). Operant extinction after fixed-interval schedules with young children. *Journal of the Experimental Analysis of Behavior, 1*, 25-29.

Catania AC (2006). *Learning* (4th Ed.). New York: Sloan.

Catania AC, Shimoff E, & Matthews BA (1989). An experimental analysis of rule-governed behavior. In SC Hayes (Ed.), *Rule-governed Behavior* (pp. 119-150). Boston, MA: Springer.

Chen FF (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 14*, 464-504. Doi: 10.1080/10705510701301834

Cheung GW & Rensvold RB (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 9*, 233-255. Doi: 10.1207/S15328007SEM0902_5

Ferrando PJ & Lorenzo Seva U (2018). Assessing the quality and appropriateness of factor solutions and factor score estimates in exploratory item factor analysis. *Educational and Psychological Measurement, 78*, 762-780. Doi: 10.1177/0013164417719308

Grant DA & Berg EA (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of Experimental Psychology*, *38*, 404-411. Doi: 10.1037/h0059831

Greiff S & Scherer R (2018). Still comparing apples with oranges? Some thoughts on the principles and practices of measurement invariance testing. *European Journal of Psychological Assessment, 34*, 141-144. Doi: 10.1027/1015-5759/a000487

Hayes SC, Barnes-Holmes D, & Roche B (2001). *Relational Frame Theory: A post-Skinnerian account of human language and cognition*. New York: Plenum Press.

Hayes SC, Barnes-Holmes D, & Wilson KG (2012). Contextual behavioral science: Creating a science more adequate to the challenge of the human condition. *Journal of Contextual Behavioral Science*, *1*, 1-16. Doi: 10.1016/j.jcbs.2012.09.004

Hayes SC, Brownstein AJ, Zettle RD, Rosenfarb I, & Korn Z (1986). Rule-governed behavior and sensitivity to changing consequences of responding. *Journal of Experimental Analysis of Behavior, 45*, 237-256. Doi: 10.1901/jeab.1986.45-237

Hayes SC, Gifford EV, & Hayes GJ (1998). Moral behavior and the development of verbal regulation. *The Behavior Analyst, 21*, 253-279.

Hayes SC, Gifford EV, & Ruckstuhl LE (1996). Relational frame theory and executive function: A behavioral approach. In GR Lyon & NA Krasnegor (Eds.), *Attention, Memory, and Executive Function* (pp. 279-305). Baltimore, MD: Paul H Brookes Publishing.

Hayes SC & Hayes LJ (1989). The verbal action of the listener as a basis for rule-governance. In SC Hayes (Ed.), *Rule-Governed Behavior* (pp. 153-190). New York: Plenum Press.

Hayes SC, Strosahl KD, & Wilson KG (1999). *Acceptance and Commitment Therapy: An experiential approach to behavior change*. New York: Guilford Press.

Hayes SC, Zettle R, & Rosenfarb I (1989). Rule-following. In SC Hayes (Ed.), *Rule Governed Behavior* (pp. 191-220). New York: Plenum Press.

Hu L & Bentler PM (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1-55. Doi: 10.1080/10705519909540118

Jöreskog KG (2005). *Structural equation modeling with ordinal variables using LISREL*. Lincolnwood, IL: Technical report, Scientific Software International.

Jöreskog KG, & Sörbom D (1999). *LISREL 8.30*. Chicago, IL: Scientific Software International.

Kantor JR (1924). *Principles of psychology Vol. I*. Chicago, IL: Principia Press.
Kelley K & Lai K (2012). *MBESS: MBESS. R package version 3.3.2*. Retrieved from http://CRAN.R-project.org/package=MBESS.
Kelley K & Pornprasertmanit S (2016). Confidence intervals for population reliability coefficients: Evaluation of methods, recommendations, and software for composite measures. *Psychological Methods, 21*, 69-92. Doi: 10.1037/a0040086
Kissi A, Hughes S, Mertens G, Barnes-Holmes D, De Houwer J, & Crombez G (2017). A systematic review of pliance, tracking, and augmenting. *Behavior Modification, 41*, 683-707. Doi: 10.1177/0145445517693811
Kline RB (2005). *Principles and practice of structural equation modeling*. New York: Guilford Press.
Lorenzo Seva U & Ferrando PJ (2006). FACTOR: A computer program to fit the exploratory factor analysis model. *Behavioral Research Methods, 38*, 88-91. Doi: 10.3758/BF03192753
Lowe CF, Beasty A, & Bentall RP (1983). The role of verbal behavior in human learning: Infant performance on fixed-interval schedules. *Journal of the Experimental Analysis of Behavior, 39*, 157-164. Doi: 10.1901/jeab.1983.39-157
Lowe CF, Harzem P, & Hughes S (1978). Determinants of operant behaviour in humans: Some differences from animals. *The Quarterly Journal of Experimental Psychology, 30*, 373-386. Doi: 10.1080/14640747808400684
Luciano C, Gómez I, & Valdivia S (2002). Consideraciones acerca del desarrollo de la personalidad desde un marco funcional-contextual [Considerations about the development of personality from a functional-contextual standpoint]. *International Journal of Psychology and Psychological Therapy, 2*, 173-197.
Luciano C, Valdivia S, Cabello F, & Hernández M (2009). Developing self-directed rules. In RA Rehfeldt, & Y Barnes-Holmes (Eds.), *Derived relational responding. Applications for learners with autism and other developmental disabilities* (pp. 335-352). Oakland, CA: New Harbinger.
Luciano C, Valdivia S, & Ruiz FJ (2012). The self as the context for rule-governed behavior. In L McHugh, & I Stewart (Eds.), *The self and perspective taking: Research and applications* (pp. 143-160). Oakland, CA: Context Press.
Millsap RE & Yun-Tein J (2004). Assessing factorial invariance in ordered-categorical measures. *Multivariate Behavioral Research, 39*, 479-515. Doi: 10.1207/S15327906MBR3903_4
O'Connor M, Byrne P, Ruiz FJ, & McHugh L (2019). Generalized pliance in relation to contingency insensitivity and mindfulness. *Mindfulness, 10*, 833-840. Doi: 10.1007/s12671-018-1046-5
Parrott LJ (1984). Listening and understanding. *The Behavior Analyst, 7*, 29-39.
Ruiz FJ, Suárez Falcón JC, Barbero Rubio A, & Flórez CL (2019). Development and initial validation of the Generalized Pliance Questionnaire. *Journal of Contextual Behavioral Science, 12*, 189-198. Doi: 10.1016/j.jcbs.2018.03.003
Ruiz FJ, García Martín MB, Suárez Falcón JC, Bedoya Valderrama L, Segura Vargas MA, PeñaVargas A, Henao AM, & Ávila Campos JE (2020). *Development and initial validation of the Generalized Tracking Questionnaire*. Unpublished manuscript. Fundación Universitaria Konrad Lorenz.
Salazar DM, Ruiz FJ, Flórez CL, & Suárez-Falcón JC (2018). Psychometric Properties of the Generalized Pliance Questionnaire-Children. *International Journal of Psychology and Psychological Therapy, 18*, 273-287.
Salazar DM, Ruiz FJ, & García Martín MB, & Bedoya Valderrama L (2020). *Generalized pliance and insensitivity to contingencies in children with learning difficulties*. Unpublished manuscript. Fundación Universitaria Konrad Lorenz.
Schlinger HD (1990). A reply to behavior analysts writing about rules and rule-governed behavior. *The Analysis of Verbal Behavior, 8*, 77-82. Doi: 10.1007/BF03392849
Shimoff E, Catania AC, & Matthews BA (1981). Uninstructed human responding: Sensitivity of low-rate performance to schedule contingencies. *Journal of the Experimental Analysis of Behavior, 36*, 207-220. Doi: 10.1901/jeab.1981.36-207
Skinner BF (1957). *Verbal behavior*. New York: Appleton-Century-Crofts.
Skinner BF (1966). An operant analysis of problem solving. In B Kleinmuntz (Ed.), *Problem Solving: Research, Method and Theory* (pp. 133-171). New York: John Wiley & Sons.
Timmerman ME & Lorenzo Seva U (2011). Dimensionality assessment of ordered polytomous items with parallel analysis. *Psychological Methods, 16*, 209-220. Doi: 10.1037/a0023353
Törneke N, Luciano C, & Valdivia S (2008). Rule-governed behavior and psychological problems. *International Journal of Psychology and Psychological Therapy, 8*, 141-156.
Vaughan M (1989). Rule-governed behavior in behavior analysis. In SC Hayes (Ed.), *Rule-governed behavior: Cognition, contingencies, and instructional control* (pp. 97-118). New York: Plenum Press.
Villatte M, Villatte JL, & Hayes SC (2015). *Mastering the clinical conversation: Language as intervention*. New York: Guilford.
Waldeck D, Pancani L, & Tyndall I (2019). An examination of the construct validity of the Generalized Pliance Questionnaire. *Journal of Contextual Behavioral Science, 11*, 50-54.
Weiner H (1969). Controlling human fixed-interval performance. *Journal of the Experimental Analysis of Behavior, 12*, 349-373.
Zettle RD & Hayes SC (1982). Rule-governed behavior: A potential theoretical framework for cognitive-behavior therapy. *Advances in Cognitive-Behavioral Research and Therapy, 1*, 73-118

Received, April 5, 2020
Final Acceptance, April 30, 2020

APPENDIX A
SPANISH VERSION OF THE GTQ-C

Debajo encontrará una lista de afirmaciones. Por favor, puntúe en qué grado cada afirmación ES VERDAD PARA USTED haciendo un círculo en los números de al lado. Utilice la siguiente escala para hacer su elección.

	Nunca	Pocas veces	A veces	Frecuente-mente	Siempre
Cuando veo que algo no está funcionando, intento algo diferente.	1	2	3	4	5
Disfruto descubriendo cómo funcionan las cosas y llegando a mis propias conclusiones.	1	2	3	4	5
Me adapto fácilmente a los cambios.	1	2	3	4	5
Tengo facilidad para encontrar soluciones nuevas cuando me surge un problema.	1	2	3	4	5
Tomo decisiones basándome en mi experiencia y no en lo que los demás dicen.	1	2	3	4	5
Me gusta probar distintas maneras de hacer las cosas para ver cuál es mejor.	1	2	3	4	5
Soy bueno encontrando nuevas formas de resolver los problemas que me surgen.	1	2	3	4	5
Si veo que algo no funciona, cambio mi forma de actuar rápidamente.	1	2	3	4	5
Aprendo de las consecuencias de mis acciones con facilidad.	1	2	3	4	5
Cuando me doy cuenta de que estoy equivocado, cambio mi forma de pensar y actuar.	1	2	3	4	5
Hago las cosas de acuerdo con los resultados que he obtenido en otras situaciones.	1	2	3	4	5
Me gusta explorar y descubrir cosas diferentes.	1	2	3	4	5

APPENDIX B
ENGLISH VERSION OF THE GTQ-C

Below you will find a list of statements. Please rate how true each statement is for you by circling a number next to it. Use the scale below to make your choice.

	Never true	Seldom true	Sometimes true	Frequently true	Always true
When I realize something is not working, I try something different	1	2	3	4	5
I enjoy discovering how things work and reaching my own conclusions	1	2	3	4	5
I adapt to changes easily	1	2	3	4	5
I am good at finding new solutions when a problem surface	1	2	3	4	5
I make decisions based on what I think and not in what other people say	1	2	3	4	5
I like to explore different ways to do things to realize what is the best	1	2	3	4	5
I am good at finding new ways of solving the problems that show up	1	2	3	4	5
If I see something that does not work, I change my form of acting quickly	1	2	3	4	5
I learn from the consequences of my actions easily	1	2	3	4	5
When I realize I am wrong, I change my way of thinking and acting	1	2	3	4	5
I do the things according to the results I have obtained in other situations	1	2	3	4	5
I like to explore and discover different things	1	2	3	4	5